13th International Conference on Spatial Information Theory

COSIT 2017, September 4–8, 2017, L'Aquila, Italy

Edited by

Eliseo Clementini
Maureen Donnelly
May Yuan
Christian Kray
Paolo Fogliaroni
Andrea Ballatore

LIPICS

Editors

Eliseo Clementini	Maureen Donnelly	May Yuan
University of L'Aquila	University at Buffalo	The University of Texas at Dallas
Italy	USA	USA
Christian Kray	Paolo Fogliaroni	Andrea Ballatore
Münster University	TU Vienna	Birkbeck, University of London
Germany	Austria	UK

ACM Classification 1998
A.0 [General] Conference Proceedings, I.2.1 Applications and Expert Systems, I.2.4 Knowledge Representation Formalisms and Methods, J.2 [Physical Sciences and Engineering] Earth and Atmospheric Sciences

ISBN 978-3-95977-043-9

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at http://www.dagstuhl.de/dagpub/978-3-95977-043-9.

Publication date
September, 2017

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Digital Object Identifier: 10.4230/LIPIcs.COSIT.2017.0

LIPIcs – Leibniz International Proceedings in Informatics

LIPIcs is a series of high-quality conference proceedings across all fields in informatics. LIPIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

ISSN 1868-8969

http://www.dagstuhl.de/lipics

Contents

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore
LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Preface

COSIT 2017 is the latest edition of the conference series on spatial information theory that has been bringing together leading researchers of the field for more than twenty years now. On its trajectory through time and space, the conference has returned to where it started: Italy. This year, it is held at L'Aquila, the capital city of the Abruzzo region, in Central Italy. The beauty of the city and the region is matched by the diversity and quality of the paper selected for presentation. In total, we received 51 submissions that were then reviewed by at least three members of the program committee. Based on the reviews, 22 papers were selected to be presented at the main conference and are included in this volume.

A quick look at the table of contents is enough to appreciate the breadth and diversity of the topics covered by the articles included in the proceedings. In addition to well-established topics, such as qualitative reasoning, spatial semantics, and wayfinding, the trend towards tackling the fundamental theoretical issues inherent to crowd-sourced spatial information continues from previous COSIT conferences. As it happens at COSIT, there are atypical topics being discussed, such as the interesting foray into a quantum theory applied to geographic fields. Overall, the program provides rich contributions for researchers in the key sub-domains of spatial information theory while at the same time extending the scientific scope of the field.

As in previous years, the main single-track program is complemented by satellite events. First of all, an intensive poster session gives the opportunity especially to young researchers to present their work to the community in a designated session. Preceding the single-track program, COSIT 2017 offers five workshops covering hot and emerging topics in spatial information theory, such as 'rethinking wayfinding support systems' and 'future directions in geospatial natural language research'. In addition, two tutorials are offered as well as the doctoral colloquium. The proceedings of the satellite events are available as a separate publication. Three keynotes complement the technical program: by Sang Ah Lee on a neuroscientist point of view on spatial cognition, by Stefano Borgo on the formalization of spatial environments of artificial agents, and by Bin Jiang on scaling and order in geographic space. Last but not least, COSIT 2017 also hosts several social events facilitating informal exchanges.

Organizing an event such as COSIT and making it a success is only possible with the help and commitment of many people. The program committee plays a pivotal role in ensuring a quality program, and we would like to thank all reviewers for their time and for the thorough reviews they produced. We would like to thank the University of L'Aquila that hosts the conference, offering the location and technical and logistic support, in addition to a financial contribution. For the first time, this year COSIT is an IFIP (International Federation for Information Processing) supported conference and an AICA (Italian Association for Informatics and Automatic Calculus) supported conference with no financial involvement of IFIP and AICA. Finally, we would like to thank all who attended COSIT 2017 to present their work, to discuss the work showcased at the conference and beyond, and to advance the state of the art in the field of spatial information theory.

July 2017

Eliseo Clementini, Maureen Donnelly,
May Yuan, Christian Kray,
Paolo Fogliaroni, Andrea Ballatore

Organization

General Chairs

- Eliseo Clementini, University of L'Aquila
- Maureen Donnelly, SUNY Buffalo

Program Chairs

- May Yuan, University of Texas at Dallas
- Christian Kray, University of Münster

Satellite Event Chairs

- Paolo Fogliaroni, Technical University Vienna
- Andrea Ballatore, Birkbeck, University of London

COSIT Steering Committee

- Christophe Claramunt, Naval Academy Research Institute, France
- Anthony Cohn, University of Leeds
- Claire Davies, University of Winchester
- Matt Duckham, RMIT University
- Max Egenhofer, University of Maine
- Sara Fabrikant, University of Zürich
- Andrew Frank, Technical University Vienna
- Christian Freksa, University of Bremen
- Nicholas Giudice, University of Maine
- Stephen Hirtle, University of Pittsburgh (Chair)
- Werner Kuhn, University of California, Santa Barbara
- David Mark, SUNY Buffalo
- Dan Montello, University of California, Santa Barbara
- Reinhardt Moratz, University of Maine
- John Stell, University of Leeds
- Kathleen Stewart, University of Iowa
- Thora Tenbrink, Bangor University
- Sabine Timpf, Augsburg University
- Stephan Winter, University of Melbourne
- Zena Wood, University of Greenwich
- Michael Worboys, University of Greenwich

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore
LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Program Committee

- Benjamin Adams, University of Canterbury
- Thomas Barkowsky, University of Bremen
- Kate Beard, University of Maine
- Scott Bell, University of Sasktchewan
- Alberto Belussi, University of Verona
- Itzhak Benenson, Tel Aviv University
- Moulin Bernard, Laval University
- Sven Bertel, Flensburg University of Applied Sciences
- Michela Bertolotto, University College Dublin
- Mehul Bhatt, University of Bremen
- Roland Billen, University of Liege
- Thomas Bittner, SUNY Buffalo
- Stefano Borgo, ISTC-CNR, Italy
- Gilberto Camara, INPE, Brasil
- Christophe Claramunt, Naval Academy Research Institute, France
- Alexis Comber, University of Leeds
- Helen Couclelis, University of California, Santa Barbara
- Kenny Coventry, University of East Anglia
- Matteo Cristani, University of Verona
- Maria Luisa Damiani, University of Milan
- Clare Davies, University of Winchester
- Leila De Floriani, University of Genoa
- Auriol Degbelo, University of Münster
- Matt Duckham, RMIT University
- Geoffrey Edwards, Université Laval
- Max Egenhofer, University of Maine
- Sara Irina Fabrikant, University of Zurich
- Zoe Falomir, University of Bremen
- Andrew Frank, Technical University Vienna
- Christian Freksa, University of Bremen
- Scott Freundschuh, University of New Mexico
- Mark Gahegan, University of Auckland
- Antony Galton, University of Exeter
- Nicholas Giudice, University of Maine
- Björn Gottfried, University of Bremen
- Amy Griffin, University of New South Wales
- Nicola Guarino, ISTC-CNR, Italy
- Stephen Hirtle, University of Pittsburgh
- Toru Ishikawa, University of Tokyo
- Bin Jiang, University of Gävle
- Christopher Jones, Cardiff University
- Marinos Kavouras, National Technical University of Athens
- Carsten Keßler, City University of New York
- Peter Kiefer, ETH Zurich
- Alexander Klippel, Pennsylvania State University
- Jakub Krukar, University of Münster
- Werner Kuhn, University of California, Santa Barbara

- Patrick Laube, Zurich University of Applied Sciences
- Sang Ah Lee, Korea Advanced Institute of Science and Technology
- William Mackaness, University of Edinburgh
- David Mark, SUNY Buffalo
- Daniel R. Montello, University of California, Santa Barbara
- Reinhard Moratz, University of Maine
- Gerhard Navratil, Technical University Vienna
- Marta Olivetti Belardinelli, Sapienza University of Rome
- Ana-Maria Olteteanu, University of Bremen
- Ross Purves, University of Zurich
- Martin Raubal, ETH Zurich
- Jochen Renz, Australian National University
- Kai-Florian Richter, Umea University
- Christoph Schlieder, University of Bamberg
- Angela Schwering, University of Münster
- Takeshi Shirabe, Royal Institute of Technology, Sweden
- Spiros Skiadopoulos, University of Peloponnese
- Barry Smith, SUNY Buffalo
- John Stell, University of Leeds
- Kathleen Stewart, University of Maryland
- Kristin Stock, Massey University
- Sabine Timpf, University of Augsburg
- Nico Van de Weghe, Ghent University
- Maria Vasardani, University of Melbourne
- Jan Oliver Wallgrün, Pennsylvania State University
- Robert Weibel, University of Zurich
- Stephan Winter, University of Melbourne
- Diedrich Wolter, University of Bamberg
- Zena Wood, University of Greenwich
- Michael Worboys, University of Greenwich
- Sisi Zlatanova, Technical University of Delft

Sponsors

- University of L'Aquila
- Department of Industrial and Information Engineering and Economics

Partners

- International Federation for Information Processing (IFIP)
- Italian Association for Informatics and Automatic Calculus (AICA)

List of Authors

Bahar Aameri
Benjamin Adams
Nikki Amos
Cecilio Angulo
Lucy Athersuch
Cristina Robles Bahm
Brandon Bennett
Mehul Bhatt
Thomas Bittner
Guido Boella
Stefano Borgo
Alessia Calafiore
Olga Chesnokova
Clare Davies
Zoe Falomir
Rob Feick
Robert Firth
Ioannis Giannopoulos
Martin Giese
Lucia Gomez Alvarez
Michael Grüninger
Nicola Guarino
Torsten Hahmann
Stephen C. Hirtle
Shelan Jeawak
Christopher Jones
David Jonietz
Leif Harald Karlsen
Markus Kattenbeck
Peter Kiefer
Vasiliki Kondyli
Thomas Liebig
Liel Luko
Grant McKenzie
Chiara Meneghetti
Veronica Muffato
Mario Nowak
Avi Parush
Ross S. Purves
Martin Raubal
Colin Robertson
Georgios Sarlas
Steven Schockaert
Carl Schultz
Giulia Sindoni
Maurice Sotzny
Lisa Stähli
John G. Stell
Shirly Stephen

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore
LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The Logic of Discrete Qualitative Relations

Giulia Sindoni[1] and John G. Stell[2]

1 **School of Computing, University of Leeds, Leeds, UK**
scgsi@leeds.ac.uk
2 **School of Computing, University of Leeds, Leeds, UK**
j.g.stell@leeds.ac.uk

Abstract

We consider a modal logic based on mathematical morphology which allows the expression of mereotopological relations between subgraphs. A specific form of topological closure between graphs is expressible in this logic, both as a combination of the negation $\neg$ and its dual $\lrcorner$, and as modality, using the stable relation Q, which describes the incidence structure of the graph. This allows to define qualitative spatial relations between discrete regions, and to compare them with earlier works in mereotopology, both in the discrete and in the continuous space.

1998 ACM Subject Classification I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Modal logic, Qualitative spatial reasoning, Discrete space

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.1

1 Introduction

Qualitative spatial relations have a long history with two major strands: the Region-Connection Calculus (RCC) of Randell et al. [10] and the 9-intersection approach of Egenhofer et al. [5]. These were initially intended to model 'continuous', or more precisely 'dense', space that can be subdivided indefinitely often. The RCC is defined in terms of first order axioms based on a primitive predicate of connection. The intersection-based theories, on the other hand, evaluate spatial scenarios through a matrix of statements that pairs of features of two regions have non-empty intersection. In the 9-intersection case a pair of regions A, B is given a spatial relationship by considering the interior, boundary, and exterior of each region and obtaining a matrix of truth values from the emptiness or non-emptiness of the 9 possible intersections between the three features of A and of B.

The mathematical discipline of *topology* provides one theory of space, related to the qualitative approach in various ways [17]. Topology is different from geometry as the former studies the properties of the space that are preserved under continuous deformation, while the latter includes shapes, relative positions and sizes of figures.

The qualitative spatial approach of the RCC is based on mereotopology, which includes mereology [13], the theory of parts and wholes.

Mereology alone is not expressive enough to be useful in Qualitative Spatial Reasoning. Besides the generic notion of parthood one needs also to be able to distinguish between central and peripheral parthood. Other desirable notions are those of connection and apartness. To express them a primitive relation of connection is usually introduced, stipulated to be symmetric and reflexive. Using parthood and connection as primitive, other important additional relations can be expressed. They are: 'X is disconnected from Y', 'X externally connected to Y', 'X is tangential part of Y', and 'X is a non-tangential part of Y'. This gives rise to what is known as RCC8. The only two relations taken as primitive are parthood and connection, being all the other relations reducible to logical formulae containing these

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 1; pp. 1:1–1:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

two. Systems of this sort are known as *mereotopologies*. A comprehensive analysis of models of mereotopological theories in terms of a relation of connection interpreted in a topological space has been presented by Cohn and Varzi [3].

Another direction in the modelling of qualitative relations in continuous space was initiated by Bloch [2], who combined modal logic with the image processing techniques of mathematical morphology [9]. Bloch demonstrated that a modal logic associated with mathematical morphology could be used to express qualitative spatial relations. More recent developments in mathematical morphology have seen much interest in applying the techniques in discrete spaces. These spaces generally consist of graphs in the sense of a set of nodes together with a binary relation of adjacency between the nodes. This kind of discrete space is exactly that investigated by Galton [6], [7], who considered how mereotopological notions could be developed for discrete space. Galton's notion of discrete space is the one of adjacency space: sets of nodes linked by a reflexive and symmetric relation of adjacency $\sim$. This is, in turn, based on the work on *digital topology* of Rosenfeld [12]. The main concern of digital topology is the study of topological properties of (subsets of) digital pictures, arrays of lattice points having positive integers coordinates (x, y). Here, given a point of coordinates (x, y), one can consider its *orthogonal adjacencies*, so those points sharing one of the coordinates with the point considered. Or one can also consider its *orthodiagonal adjacencies*, consisting of its orthogonal adjacencies with its four diagonal adjacent point. These constructions are the adjacency spaces $(\mathbb{Z}, \sim_4)$ and $(\mathbb{Z}, \sim_8)$.

Discrete space presents some notable challenges for mereotopology. For example, the usual definition of part in terms of connection leads inescapably, in the presence of atomic regions, to the conclusion that some regions will be parts of their complements. This may call for a different understanding of notions such as complement and part, or alternatively for novel techniques for developing mereotopological theories. In this paper we present new results on the mereotopology of discrete space using a recently developed modal logic [15] with a semantics based on morphological operations on graphs. This allows us use the approach suggested by Bloch for expressing qualitative relationships but in the very different setting of discrete space. Our results are related to the algebraic approach advocated by Stell and Worboys at COSIT twenty years ago [16]. This work took the bi-intuitionistic algebra of subgraphs that Lawvere [8] noted and showed its relevance to qualitative relations in discrete space. The current paper extends this significantly through its use of the modal logic [15] thus allowing us to adapt Bloch's insights about the use of morphological operations to the discrete case.

Both the mereotopological work of Galton [6] and the morphological investigations of Cousty et al. [4] take place in a setting where space consists of nodes which may or may not be linked by edges. Galton's adjacency spaces can be regarded as graphs. Anyway there is a notable difference between theory of adjacency space and graph theory, as Galton underlines [7]. A substructure of an adjacency space can be specified just in terms of nodes, two nodes being connected by only one edge, or relation of adjacency. This is not true in the general setting of a graph, where multiple edges may occur between two nodes, and, therefore, different subgraphs sharing the same set of nodes may be considered. Cousty et al. find indeed that edges need to play a more central role, and make the key observation that sets of nodes which differ only in their edges need to be regarded as distinct. The logic used in the present paper takes its semantics in a setting where regions in a graph are more general still. We allow graphs to have multiple edges between the same pair of nodes, thus using a structure sometimes called a multi-graph. This generality appears important in practical examples, such as needing to model two distinct roads between the same endpoints, or distinct rail connections between the same two stations.

The contribution of the present paper is thus to develop the interaction between modality and morphology identified by Bloch but in the discrete setting. In doing so we are able to show how this relates to earlier work in mereotopology both in the discrete and in the continuous case. We start in Section 2 by reviewing the framework of Cohn and Varzi and showing that the discrete connection of Galton's work lies outside this framework. In Section 3 we review the semantics of a multi-modal logic where formulas are interpreted as subgraphs. This is used in Section 4 to express qualitative spatial relations within the logic. We provide conclusions in Section 5.

2 Connection in Continuous and Discrete Space

In this section we review the approach of Cohn and Varzi [3] and show that it needs to be generalized if it is to capture the notion of discrete connection defined by Galton [6].

2.1 Mereotopological Connection from Topological Closure

Giving a set A, a topology τ on A is usually given as a collection of subsets of A which is closed under finite intersections and arbitrary unions. The set A together with the topology τ on A is a topological space, and the elements of τ are the open sets of the space. A set is closed if and only if it is the complement of an open set. An alternative formulation of topology, dual to the one in terms of open sets, can be given. Closed sets are the fundamental elements, and a topology on a set A is a collection of subsets that is closed under finite union and arbitrary intersection [14].

Cohn and Varzi [3] give three definitions of connection which depend on the notion of topological closure.

▶ **Definition 1.** A closure operator on a set A is a function c associating with each $x \subseteq A$ a set $c(x) \subseteq A$, which satisfies the following axioms (Kuratowski axioms) for all $x, y \subseteq A$.
K1. $c(\varnothing) = \varnothing$.
K2. $x \subseteq c(x)$.
K3. $c(c(x)) \subseteq c(x)$.
K4. $c(x \cup y) = c(x) \cup c(y)$.

Given a set A together with an operator c satisfying $K1$-$K4$ axioms is equivalent to specifying a topological space in terms of open sets or in terms of closed sets. The closed sets of a topological space correspond to the sets $x \subseteq A$ for which $c(x) = x$.

▶ **Definition 2.** Let c be a topological closure on A. Three binary relations of connection between subsets $x, y \subseteq A$ are defined as follows.
1. $C_1(x, y) \Leftrightarrow x \cap y \neq \varnothing$.
2. $C_2(x, y) \Leftrightarrow c(x) \cap y \neq \varnothing$ or $x \cap c(y) \neq \varnothing$.
3. $C_3(x, y) \Leftrightarrow c(x) \cap c(y) \neq \varnothing$.

Cohn and Varzi use a setting where a mereotopological theory might allow only certain subsets of the topological space as its entities. For example, they draw a line between theories which allow as elements in the domain of quantification *boundary* elements, which, intuitively are elements with an empty interior, such as points, lines and surfaces, and theories which exclude such boundary elements. However, the spatial relationship of connection is defined in a way that is applicable to arbitrary subsets of the topological space.

The Region-Connection Calculus (RCC) is one of the most popular theories in qualitative spatial reasoning. An important models of the RCC is to take regions to be non-empty regular closed subsets of $\mathbb{R}^2$, with the usual topology. A subset is called regular closed when it is equal to the closure of its interior. In particular, this means that although a single point, or a line including its endpoints, is a closed set in $\mathbb{R}^2$, it is not regular closed, as its interior is empty. Therefore, such elements are not considered as regions in this context, and the RCC belongs to those theories which do not allow boundary elements in their domain. In the regular-closed model of RCC, all three connections above yield the same relation between regions, and connection means sharing at least one point. External connection, or abutting in the language of Cohn and Varzi, is distinguished from connection as regions that abut share points in this model but do not share regions. However, in other models of RCC and in other mereotopological systems the three notions of connection can have substantially different properties.

The contribution of Cohn and Varzi is to have provided a framework within which numerous mereotopological notions are expressible by varying the notion of connection used as well as the two key derived notions of part and fusion. In the case of part, the three connections yield three parthoods as follows.

$$P_i(x, y) \Leftrightarrow \forall z(C_i(z, x) \to C_i(z, y))\,.$$

Although Cohn and Varzi [3, p359] aim for neutrality with respect to density of space, that is whether space can be repeatedly sub-divided *ad infinitum*, we shall see next that the use of topological closure prevents the framework including one of the most straightforward examples of connection in a discrete space.

2.2 Galton's Discrete Connection

Galton [6] studied a notion of connection between subsets of a particular kind of discrete space. The spatial setting is a set N together with a relation of adjacency $\alpha \subseteq N \times N$. The relation α is symmetric and reflexive, but not transitive. Connection, C_α, is defined for subsets $x, y \subseteq N$ by $C_\alpha(x, y)$ if there are $a \in x$ and $b \in y$ such that $(a, b) \in \alpha$. We shall show next that there are spaces N, α where this connection is not expressible as any C_i, in the sense of Cohn and Varzi, for any topological closure on N. A specific example appears in Figure 1 where the links indicate adjacencies between distinct elements of the five element set $N = \{m, n, p, q, r\}$.

First, C_α cannot be C_1 as two adjacent nodes give disjoint singleton subsets which are C_α connected. So suppose that $C_\alpha = C_2$ for some topological closure c. If k is any node in N then $\{k\}$ is C_α connected to no singletons except those $\{k'\}$ such that $\alpha(k, k')$. Thus $c(\{k\})$ contains only nodes which are adjacent to k. Hence for the specific nodes m and n we have $c(\{m\}) \subseteq \{r, m, n\}$ and $c(\{n\}) \subseteq \{m, n, p\}$. Now $\{m\}$ and $\{n\}$ are connected in the connection C_α so if they are C_2 connected we must have $n \in c(\{m\})$ or $m \in c(\{n\})$. Consider first the case that $n \in c(\{m\})$. This implies $\{n\} \subseteq c(\{m\})$ so that $c(\{n\}) \subseteq c(c(\{m\})) \subseteq c(\{m\})$. Thus $p \notin c(\{n\})$ and $c(\{n\}) \subseteq \{m, n\}$. But $\{n\}$ and $\{p\}$ are connected in C_α, so $n \in c(\{p\})$, and hence $c(\{n\}) \subseteq c(\{p\})$. As $m \notin c(\{p\})$ we conclude $c(\{n\}) = \{n\}$ in the case that $n \in c(\{m\})$. In the case that $m \in c(\{n\})$ we conclude that $c(\{m\}) = \{m\}$. Thus in either case one of the sets $\{m\}$ and $\{n\}$ is a closed set, and they cannot both be closed since they need to be C_2 connected.

This applies to each pair of adjacent nodes in N; one of them is a closed set and the other is not. With an odd number of nodes in total this is a contradiction. Hence no such topological closure, c, can generate a C_2 connection equal to C_α. There remains the possibility that C_α is of the form C_3. Suppose then that some topological closure on N generates C_α as C_3. We must have $c(\{m\}) \cap c(\{n\}) \neq \varnothing$. For similar reasons to the C_2

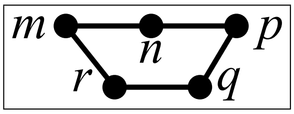


Figure 1 A discrete space where connection cannot be defined in terms of topological closure.

case we must have $c(\{m\}) \subseteq \{r, m, n\}$ and $c(\{n\}) \subseteq \{m, n, p\}$, so to obtain the non-empty intersection either $m \in c(\{n\})$ or $n \in c(\{m\})$. In the case that $n \in c(\{m\})$ we get $p \notin c(\{n\})$ as $c(\{n\}) \subseteq c(\{m\})$. But n and p are adjacent so are connected singletons and $c(\{n\})$ and $c(\{p\})$ must intersect and the only possibility for this intersection is n. Now if $m \in c(\{n\})$ we get $m \in c(c(\{p\}))$ which would make m and p adjacent. Hence $c(\{n\})$ can only be $\{n\}$. It is straightforward to continue to a contradiction as in the C_2 case.

3 Modal Logic with Graph Morphology Semantics

3.1 Classical Modal Logic

The syntax of classical propositional modal logic provides propositional variables $p, q, r, \ldots$, the usual logical connectives $\vee$, $\wedge$, $\rightarrow$, $\neg$, and the modalities $\Diamond$, and $\Box$. Formulae are defined by stipulating that propositional variable are formulae, and if φ, ψ are formulae then so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \rightarrow \psi$, $\neg\varphi$, $\Diamond\varphi$, and $\Box\varphi$. The semantics for this logic allows an interpretation of atomic propositions as subsets of a set of 'worlds' and formulae correspond to subsets constructed out of these. While an abstract set has no spatial structure by itself, we shall see that a more elaborate logic has a natural semantics in which formulae correspond to subgraphs of a graph. This means that spatial relations between subgraphs can be expressed in the logic. Before introducing this logic we need to review the connection between classical propositional modal logic and the morphological operations of dilation and erosion.

Kripke semantics for propositional modal logic is based on a binary relation on a set of worlds, W (see [1] for an introduction to Kripke semantics). Propositional variables are then interpreted as subsets of W, and truth and falsity in the language, often denoted $\top$ and $\bot$ are interpreted respectively as W and $\varnothing$. In this setting the logical connectives $\vee$, $\wedge$, $\neg$ are interpreted as the set-theoretic operations of union, intersection and complement. Once we are given a subset $[\![p]\!] \subseteq W$ for each propositional variable p, we can assign to each non-modal formula φ its interpretation as a subset $[\![\varphi]\!] \subseteq W$. Implication $\rightarrow$ is handled by defining $[\![\varphi \rightarrow \psi]\!] = -[\![\varphi]\!] \cup [\![\psi]\!]$ where $-$ is set-theoretic complement. This means $[\![\varphi \rightarrow \psi]\!]$ holds in a given interpretation if and only if $[\![\varphi]\!] \subseteq [\![\psi]\!]$ making a connection between the logic and the mereological assertion that one set is a part of another.

The semantics of the modalities $\Diamond$ and $\Box$ can easily be expressed in terms of the morphological operations of dilation and erosion which are defined as follows.

▶ **Definition 3.** For any subset $X \subseteq W$, and any relation $R \subseteq W \times W$, we define:

- **Dilation:** $X \oplus R = \{w \in W \mid \exists x(x\ R\ w \text{ and } x \in X)\}$,
- **Erosion:** $R \ominus X = \{w \in W \mid \forall x(w\ R\ x \text{ implies } x \in X)\}$.

In order to understand how dilation and erosion work, we do an example

▶ **Example 4.** Let $W = \{a, b, c, d, e\}$ and $X = \{a, b, c\}$ and $R = \{(a, a), (a, b), (c, d), (e, c)\}$. Then $X \oplus R = \{a, b, d\}$ and $R \ominus X = \{a, e\}$.

Using these operations we then define $[\![\Diamond\varphi]\!] = [\![\varphi]\!] \oplus \breve{R}$ and $[\![\Box\varphi]\!] = R \ominus [\![\varphi]\!]$. Note the use of the converse $\breve{R}$; that is $\Diamond\varphi$ holds at worlds accessible from worlds in $[\![\varphi]\!]$ via the converse of the accessibility relation.

3.2 Graphs and Relations on Graphs

We move now to consider not merely subsets of a set but subgraphs of a graph. This builds on the use of algebraic operations on subgraphs as introduced to the COSIT community by Stell and Worboys [16], but now in the context of a modal logic which allows the expression of mereotopological relations between subgraphs. The logic itself appears in [15] in a more general context, but the applications to discrete spatial representation have not been investigated before. We need first to explain what we mean by a graph.

A graph in which there are potentially multiple edges between nodes, and potentially (multiple) loops on the nodes can be defined a set W, thought of as consisting of all the nodes and edges together, with a relation $Q \subseteq W \times W$. This relation relates every edge to its incident nodes and no other elements of W are related. Thus $w\ Q\ v$ holds iff w is an edge incident with node v. From Q we derive its reflexive closure, which we denote by H. Given just W and H we can distinguish nodes from edges as a node is an element of W related only to itself by H, whereas an edge must be related both to itself and at least one other element of W. The subgraphs of a graph (W, H) are the subsets which for each edge include all the incident nodes. A set $X \subseteq W$ will be a subgraph iff $X \oplus H \subseteq X$ or equivalently $X \oplus Q \subseteq X$.

The algebra of subgraphs, already noted by Lawvere as cited in [16], provides unions and intersections of subgraphs but most significantly two distinct types of complement. Given a subgraph $X \subseteq W$ we can obtain both a largest subgraph disjoint from X and also a smallest subgraph whose union with X gives all of W. These are denoted $\neg X$ and $\lrcorner X$ respectively and can be expressed as $H \ominus (-X)$ and $(-X) \oplus H$ respectively.

To give a semantics for a modal logic where formulae are interpreted as subgraphs we need a notion of a relation on a graph which extends the notion of a relation on a set as used in classical Kripke semantics.

Let W be a set, let $\mathcal{P}_W$ be its power set, and let S be a function such that $S : \mathcal{P}_W \mapsto \mathcal{P}_W$; then

▶ **Definition 5.** S is a union preserving function on $\mathcal{P}_W$ if and only if for any family of indexed set $Z_i \subseteq X$ we have $S(\cup_i(Z_i) = \cup_i S(Z_i)$.

Given a union preserving function S on $\mathcal{P}_W$, it is always possible to define a binary relation $R \subseteq W \times W$ as follows: for any w and $v \in W$, $w\ R\ v$ if and only if $v \in S(\{w\})$. All binary relations on sets come in this way. On the other hand, given a relation $R \subseteq W \times W$, a union preserving function on $\mathcal{P}_W$ can be defined as follows, given $V \subseteq W$

$$S(V) = \{w \in W \mid \exists v \in V \text{ and } v\ R\ w\}.$$

Therefore, relations on a set W can be modelled as union-preserving functions on the power set of W. When it comes to the case of a graph, so a set carrying a pre order H, the union preserving functions on the lattice of subgraphs correspond to relations on W that are stable. Given a graph (W, H) we say a relation $R \subseteq W \times W$ is stable with respect to H provided $H ; \breve{R} ; H \subseteq R$ where ; denotes the composition of relations. The stable relations include the universal relation $U = W \times W$, and the relations Q and H.

Stable relations are closed under composition, H being the identity element, but are not, in general, closed under converse. Denoting the standard converse of R by $\breve{R}$, is not always the case that $H; \breve{R}; H \subseteq \breve{R}$. However, for a stable R, it is possible to define a relation, called *left converse*, characterized as the smallest stable relation containing $\breve{R}$.

▶ **Definition 6.** The **left converse** of a stable relation R, is $\backsim R = H; \breve{R}; H$, where $\breve{R}$ is the (ordinary) converse of R.

3.3 Graph-based Modal Logic

Bi-intuitionistic stable tense logics are a group of logics, described in [15], with a Kripke semantics where worlds in a frame are equipped with a pre-order as well as with an accessibility relation which is 'stable' with respect to the pre-order. We do not need the full generality of this setting here, and will give a semantics in a graph $G = (W, H)$. The relation H is easily seen to be reflexive and transitive, so that it is a pre-order. The syntax of the multi-modal version of *BISKT*, called so from [15] because is the system K of this group of logics, is that of classical propositional logic extended with dual negation $\lrcorner$, dual implication $\succ$, and four indexed modalities: $[R]$, $\langle R \rangle$, $\rangle R \langle$, and $]R[$.

The semantics needs, besides a graph $G = (W, H)$, a stable relation for each index R. Such a structure will be called a *BISKT*-model, and often denoted $\mathcal{M}$. Given a valuation, assigning to each propositional variable p a subgraph $[\![p]\!]$, we extend the semantic function $[\![\_]\!]_\nu$ thus (we will omit the subscription 'ν' when no confusion arises):

$$\begin{array}{rclrcl}
[\![\bot]\!]_\nu & = & \varnothing & [\![\top]\!]_\nu & = & W \\
[\![\varphi \vee \psi]\!]_\nu & = & [\![\varphi]\!]_\nu \cup [\![\psi]\!]_\nu & [\![\varphi \wedge \psi]\!]_\nu & = & [\![\varphi]\!]_\nu \cap [\![\psi]\!]_\nu \\
[\![\neg \varphi]\!]_\nu & = & \neg [\![\varphi]\!]_\nu & [\![\lrcorner\, \varphi]\!]_\nu & = & \lrcorner [\![\varphi]\!]_\nu \\
[\![\varphi \to \psi]\!]_\nu & = & H \ominus ((-[\![\varphi]\!]_\nu) \cup [\![\psi]\!]_\nu) & [\![\varphi \succ \psi]\!]_\nu & = & ([\![\varphi]\!]_\nu \cap (-[\![\psi]\!]_\nu)) \oplus H \\
[\![[R]\, \varphi]\!]_\nu & = & R \ominus [\![\varphi]\!]_\nu & [\![\langle R \rangle\, \varphi]\!]_\nu & = & [\![\varphi]\!]_\nu \oplus (\backsim R) \\
[\![\rangle R \langle\, \varphi]\!]_\nu & = & [\![\varphi]\!]_\nu \oplus R & [\![]R[\, \varphi]\!]_\nu & = & (\backsim R) \ominus [\![\varphi]\!]_\nu
\end{array}$$

A graph will be indicated by $G = (W, H)$; a valuation function ν, is a function going from formulas in the logic to subgraphs, such that for a formula φ, $\nu(\varphi) = [\![\varphi]\!]_\nu$. The pair G, ν is a *BISKT*-model and we write $G, \nu \vDash \varphi$ when $[\![\varphi]\!]_\nu$ is the whole graph G.

The use of morphology in connection with modal logic for spatial reasoning by Bloch [2] is in a classical setting. In *BISKT*, unlike the classical case, box and diamond modalities are not mutually interdefinable. Working in discrete space the bi-intuitionistic logic is essential to express spatial relations as we will see in the next section.

4 Expressing qualitative relations

In this section we express in a direct way, in *BISKT*, some qualitative spatial relationships between graphs. Then, we compare these expressions with their correspondents found in [3] and [10]. The first relation we analyse is the one of connection[1].

4.1 A Cech Closure for Graphs

We have seen that connection in Galton's sense cannot always be expressed as one of the connections in the framework of Cohn and Varzi by using a topological closure. However, a

[1] The reader must be aware of the fact that the notion of connection *between* graphs, which we refer to, is not the same as the notion of connectivity *of* a graph. In the latter case a graph is connected if between any two nodes there is always an edge connecting them. In our context, two subgraphs, that are connected in the sense expressed by one of the forthcoming relations of connection, are not necessarily connected subgraphs themselves.

weaker notion of closure does have the right properties. An operator satisfying Kuratowski axioms $K1$, $K2$ and $K4$ but not necessarily $K3$ is known as a Cech closure (see definition from [14, p.657]).

In $BISKT$ the following formulas are tautologies:

- $\lrcorner\neg\bot \leftrightarrow \bot$,
- $\varphi \to \lrcorner\neg\varphi$,
- $\lrcorner\neg(\varphi \vee \psi) \leftrightarrow \lrcorner\neg\varphi \vee \lrcorner\neg\psi$.

However, the following formula is not a tautology:

- $\lrcorner\neg(\lrcorner\neg\varphi) \to \lrcorner\neg\varphi$.

Interpreting $\bot$ as the empty subgraph $\varnothing$, the formulas φ and ψ on the respective subgraphs $[\![\varphi]\!]$ and $[\![\psi]\!]$, we have that the operator on graphs '$\lrcorner\neg$' satisfies

K1. $\lrcorner\neg[\![\bot]\!] = \varnothing$,
K2. $[\![\varphi]\!] \subseteq \lrcorner\neg[\![\varphi]\!]$,
K4 $\lrcorner\neg([\![\varphi]\!] \cup [\![\psi]\!]) = \lrcorner\neg[\![\varphi]\!] \cup \lrcorner\neg[\![\psi]\!]$,
but not necessarily
K3. $\lrcorner\neg(\lrcorner\neg[\![\varphi]\!]) \subseteq \lrcorner\neg[\![\varphi]\!]$.

Given the above tautologies we can define a Cech closure c on a graph G by

$$c(K) = \lrcorner\neg K$$

for any $K \subseteq G$.

4.2 Connection expressed modally

In this section we show that Cech closure operator defined earlier, is expressible by a modality in *BISKT*, for a suitable choice of a stable accessibility relation R on W.

Consider the relation Q introduced in Section 3.2. When it is taken as the stable relation, then the Cech closure of a subgraph, can be expressed by a modality indexed by Q.

▶ **Theorem 7.** *Let Q be the stable relation introduced above. Consider the graph G. For any formula φ, the following holds:*

$$\lrcorner\neg[\![\varphi]\!] = [\![\langle Q\rangle\,\varphi]\!]\,.$$

Proof. We sketch the idea of the proof in Figure 2, showing how $\backsim\!\!\!\!\!\cdot\, Q$ works. When dilation by $\backsim\!\!\!\!\!\cdot\, Q$ is applied to a node, it takes the node itself and all the nodes one-edge-away from it. When dilation by $\backsim\!\!\!\!\!\cdot\, Q$ is applied to an edge, it produces the edge itself , the edges one-node-apart from it, and all the nodes incident with these edges. Since, in **BISKT**, the smallest subgraph including a node is the node itself, and the smallest subgraph including an edge is the edge plus the nodes incident to it, $\lrcorner\neg$ acts extending any subgraph with all the nodes one-edge away and all the edges incident to them, which means dilating the subgraph by $\backsim\!\!\!\!\!\cdot\, Q$. ◀

By taking R to be the universal relation U on W, that is $U = W \times W$ we can interpret "somewhere φ" by $\langle U\rangle\,\varphi$ and "everywhere φ" by $[U]\,\varphi$, or equivalently just φ. The three notions of connection expressed by Cohn and Varzi [3] depend on being able to express that a subset is non-empty. We can handle this within our modal logic as $\langle U\rangle(\varphi)$ holds if and only if $[\![\varphi]\!] \neq \varnothing$. Thus, for any graph G and any valuation ν, given two formula x and y, and given the Cech closure c introduced above we have that

- $G, \nu \vDash \langle U\rangle(x \wedge y)$ iff $([\![x]\!]_\nu \cap [\![y]\!]_\nu) \neq \varnothing$,
- $G, \nu \vDash \langle U\rangle(((\langle Q\rangle\, x \wedge y) \vee (x \wedge \langle Q\rangle\, y))$ iff $((c([\![x]\!]_\nu) \cap [\![y]\!]_\nu) \cup (x \cap c([\![y]\!]_\nu))) \neq \varnothing$,
- $G, \nu \vDash \langle U\rangle((\langle Q\rangle\, x \wedge \langle Q\rangle\, y)$ iff $c([\![x]\!]_\nu) \cap c([\![y]\!]_\nu) \neq \varnothing$.

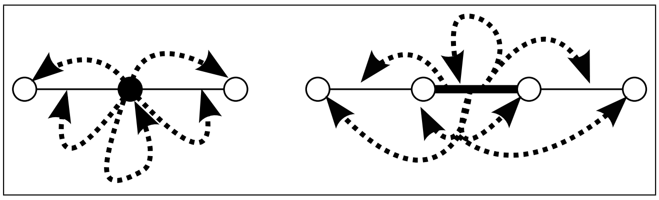


Figure 2 Effect of $\circlearrowleft Q$ on one node and on one edge.

This gives us three definitions of connection for regions in discrete space, analogous to C_1, C_2, C_3 defined by Cohn and Varzi for regions in continuous space:

1. $C_1([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle U \rangle(x \wedge y)$.
2. $C_2([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle U \rangle((\langle Q \rangle x \wedge y) \vee (x \wedge \langle Q \rangle y))$.
3. $C_3([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle U \rangle(\langle Q \rangle x \wedge \langle Q \rangle y)$.

4.3 Qualitative Relations Modally

Other qualitative spatial relations can be expressed in a direct way in this modal logic. We consider here the notions of non empty part, general part, proper part, tangential proper part, non-tangential proper part, and external connection. We index them notationally with $\Diamond$ in order to make the distinction with their mereotopological correspondents.

- $NEP_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle U \rangle\, x \wedge x \to y$
- $GP_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash x \to y$
- $PP_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash x \to y \wedge \langle U \rangle(\neg x \wedge y)$
- $O_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle U \rangle(x \wedge y)$
- $EC_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \neg(x \wedge y) \wedge \langle U \rangle(\langle Q \rangle x \wedge y) \vee \langle U \rangle(x \wedge \langle Q \rangle y)$
- $TP_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash x \to y \wedge \langle U \rangle(\langle Q \rangle x \wedge \neg y)$
- $NTP_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \langle Q \rangle x \to y$
- $EQ_\Diamond([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash x \leftrightarrow y$
- $DC([\![x]\!]_\nu, [\![y]\!]_\nu)$ iff $G, \nu \vDash \neg((\langle Q \rangle x \wedge y) \vee (x \wedge \langle Q \rangle y))$

The notion of parthood comprises three different relations: the general notion of part ($GP_\Diamond$), the one restricted to those parts x of y different from the empty graph ($NEP_\Diamond$), and the notion of proper part ($PP_\Diamond$). A separate section will be dedicated to the notion of boundary graph, since, as we shall see, a variety of possible definitions arises in *BISKT*.

In the next sections we will compare our definitions with those given in [3, 10]. In order to do so, we introduce the following lemmas

▶ **Lemma 8.** *Given a world $w \in W$, if for some $v \in W$ $v \circlearrowleft Q\, w$, then $w \circlearrowleft Q\, v$ or there exists a world u such that $w \circlearrowleft Q\, u$ and $u \in v \oplus H$.*

Proof. Four cases need to be addressed. (i) v and w are both nodes. The case $v = w$ is trivial. So suppose $v \neq w$. $v \circlearrowleft Q\, w$ holds. Then there is an edge u incident with both v and w such that $v\ H\ u\ \breve{Q}\ w\ H\ w$. Therefore $w\ H\ w\ \breve{Q}\ u\ H\ v$. (ii) v and w are both edges. The case $v = w$ is trivial. So suppose $v \neq w$. Then there exists a node u incident with both v and w such that $v\ H\ u\ \breve{Q}\ w\ H\ w$. So $w\ H\ u\ \breve{Q}\ v\ H\ v$. (iii) v is a node and w is an edge. If $v \circlearrowleft Q\, w$ then the only possibility is $v\ H\ v\ \breve{Q}\ w\ H\ w$. Therefore $w\ H\ v\ \breve{Q}\ w\ H\ v$. (iv) Suppose v is an edge and w is a node. Then (iv.i) w is a node incident to v: $v\ H\ w\ \breve{Q}\ v\ H\ w$. In this case $w\ H\ w\ \breve{Q}\ v\ H\ v$; or (iv.ii) there is a node u such that $u \in v \oplus H$ and an edge k, such that $v\ H\ u\ \breve{Q}\ k\ H\ w$. Therefore $w\ H\ w\ \breve{Q}\ k\ H\ u$, and $u \in v \oplus H$. ◀

▶ **Lemma 9.** *Given a BISKT-model G, ν, and two formulas x and y representing two subgraphs $[\![x]\!]$ and $[\![y]\!]$*

$$G, \nu \vDash \langle U \rangle (\langle Q \rangle x \wedge y) \iff G, \nu \vDash \langle U \rangle (x \wedge \langle Q \rangle y)$$

Proof. Given a graph G, a valuation ν and a world w, we can give the semantics clause for $\langle Q \rangle \varphi$ as follows

$$w \vDash \langle Q \rangle \varphi \text{ iff for some } v, (v \smile Q\ w) \text{ and } v \vDash \varphi \,.$$

For the left-to-right direction: assume that $G, \nu \vDash \langle U \rangle (\langle Q \rangle x \wedge y)$. Then, for all $w \in W$ exists a $v \in W$ such that $(w\ U\ v)$ and $v \vDash \langle Q \rangle x \wedge y$. This means that $v \vDash y$, and for some u such that $(u \smile Q\ v)$, $u \vDash x$. For lemma 1, or $v \smile Q\ u$, so that $u \vDash x \wedge \langle Q \rangle y$, or there is a $j \in v \oplus H$ such that $j \smile Q\ u$. So $j \vDash y$ because $j \in v \oplus H$, and $j \vDash \langle Q \rangle x$. So $j \vDash x \wedge \langle Q \rangle y$. Therefore, under the initial assumption, $G, \nu \vDash \langle U \rangle (x \wedge \langle Q \rangle y)$ holds. The right-to-left direction works in analogous way. ◀

Given this lemma, the notion of connection C_2 can be shortened to $\langle U \rangle (\langle Q \rangle x \wedge y)$. When this holds for x and y, the corresponding subgraphs are C_2-connected.

▶ **Lemma 10.** *Given a graph G and a valuation function, and x, y propositional variables, ν, if $G, \nu \vDash \langle U \rangle x$ and $G, \nu \vDash x \to y$ then $G, \nu \vDash \langle U \rangle y$.*

Proof. $G, \nu \vDash \langle U \rangle x$ iff for all $w \in v$, there is a $v \in W$ such that $v \vDash x$. $G, \nu \vDash x \to y$ iff for all $w \in W$, for all $u \in W$ such that $w\ H\ u$, if $u \vDash x$ then $u \vDash y$. Take v: $v \vDash \langle U \rangle x$ and $v \vDash x$. Also, $v\ H\ v$. Then $v \vDash y$. Therefore, under the initila assumptions, somewhere in the graph y must hold: $G, \nu \vDash \langle U \rangle y$. ◀

4.4 Overlapping

Overlapping regions are defined by Cohn and Varzi [3] as $O(X, Y) \equiv \exists Z (P(Z, X) \& P(Z, Y))$, where the predicate of parthood is restricted just to non-empty regions. We show that

▶ **Theorem 11.** *Let G be a graph and ν a valuation. Given x, y and z propositional variables, the following holds*

$$G, \nu \vDash \langle U \rangle (x \wedge y)$$

iff there is a subgraph K of G such that for any valuation ν' which agrees with ν on x and y, and where $[\![z]\!]_{\nu'} = K$

$$G, \nu' \vDash \langle U \rangle z \wedge z \to x \text{ and } G, \nu' \vDash \langle U \rangle z \wedge z \to y \,.$$

Proof. Suppose $G, \nu \vDash \langle U \rangle (x \wedge y)$. Then, for all $w \in W$, $w \vDash \langle U \rangle (x \wedge y)$. So for all $w \in W$ there is a world $v \in W$ such that $v \vDash x$ and $v \vDash y$. This means that $v \in [\![x]\!]_\nu$ and $v \in [\![y]\!]_\nu$. So $v \in [\![x]\!]_\nu \cap [\![y]\!]_\nu$. Then, for all the valuations ν' that agree with ν on x and y, there is a subgraph $[\![z]\!]_{\nu'} = K = v \oplus H$ such that $K, \nu' \vDash z$, because $K = [\![z]\!]_{\nu'}$, and $K, \nu' \vDash z \to x$, because $[\![z]\!]_{\nu'} \subseteq [\![x]\!]_{\nu'}$, and $K, \nu' \vDash z \to y$, because $[\![z]\!]_{\nu'} \subseteq [\![y]\!]_{\nu'}$. Therefore $G, \nu' \vDash \langle U \rangle z \wedge z \to x$ and $G, \nu' \vDash \langle U \rangle z \wedge z \to y$.

On the other hand, suppose $[\![z]\!]_{\nu'} = K$, and $G, \nu' \vDash \langle U \rangle z \wedge z \to x$ and $G, \nu' \vDash \langle U \rangle z \wedge z \to y$. But then $G, \nu' \vDash \langle U \rangle z \wedge z \to (x \wedge y)$, since, given p, q, r propositional variables, $(p \to q) \wedge (p \to r) \leftrightarrow p \to (q \wedge r)$ is a tautology in BISKT. But then for lemma 4, $G, \nu' \vDash \langle U \rangle (x \wedge y)$ and since ν' and ν agree on x and y, $G, \nu \vDash \langle U \rangle (x \wedge y)$. ◀

4.5 Tangential Part

Cohn and Varzi give the following definition of Tangential part: $TP(X,Y) \equiv P(X,Y)$ & $\exists Z(C(X,Z)\& - O(Z,Y))$, where '$-$', is the symbol for classical negation. We show that our relation $TP_{\Diamond}(\llbracket x \rrbracket, \llbracket y \rrbracket)$ gives the same entailments on subgraphs.

▶ **Theorem 12.** *Let G be a graph and ν a valuation. Given x, y and z propositional variables, the following holds*

$$G, \nu \vDash (x \to y) \wedge \langle U \rangle (\langle Q \rangle\, x \wedge \neg y)$$

iff there is a subgraph K of G such that for any valuation ν' which agrees with ν on x and y, and where $\llbracket z \rrbracket_{\nu'} = K$

$$G, \nu' \vDash x \to y \text{ and } G, \nu' \vDash \langle U \rangle (\langle Q \rangle\, x \wedge z) \text{ and } G, \nu' \nvDash z \wedge y\,.$$

Proof. Assume $G, \nu \vDash (x \to y) \wedge \langle U \rangle (\langle Q \rangle\, x \wedge \neg y)$. Then, for all $w \in W$ in the graph $w \in W$, $w \vDash (x \to y)$. Then, since ν and ν' agree on x and y, also $G, \nu' \vDash (x \to y)$. From the assumption it follows also that $G, \nu \vDash \langle U \rangle (\langle Q \rangle\, x \wedge \neg y)$. Again since ν and ν' agree on x and y, also $G, \nu' \vDash \langle U \rangle (\langle Q \rangle\, x \wedge \neg y)$. Take as $\llbracket z \rrbracket_{\nu'} = K$ the subgraph $\llbracket \neg y \rrbracket_{\nu'}$. There exists such a subgraph such that $\langle U \rangle (\langle Q \rangle\, x \wedge z)$. Suppose $G, \nu' \vDash z \wedge y$. That means that $G, \nu' \vDash \neg y \wedge y$. But this is impossible since in *BISKT*, for any propositional variable p, $(p \wedge \neg p) \to \bot$ is a tautology. Therefore, for the chosen subgraph $\llbracket z \rrbracket_{\nu'}$, necessarily $G, \nu' \nvDash z \wedge y$.

For the other direction, assume $G, \nu' \vDash x \to y$. Since ν' and ν agree on x and y, $G, \nu \vDash x \to y$. Suppose there exists some $\llbracket z \rrbracket_{\nu'} = K$, subgraph of G, and G, $\nu' \vDash \langle U \rangle (\langle Q \rangle\, x \wedge z)$ and $G, \nu' \nvDash z \wedge y$. That means that i) for all $w \in W$, there is the $v \in W$ such that $v \vDash \langle Q \rangle\, x$ and $v \vDash z$; and ii) it does not exist any $t \in W$ such that $t \vDash z \wedge y$. So, for all $t \in W$, if $t \vDash z$ then $t \nvDash y$. But, since $\llbracket z \rrbracket_{\nu'} = K$, for all $k \in W'$ such that $(W', H) = K$, $k \vDash z$. So, also, for all those k, $k \nvDash y$. That means that for all m such that vHm, $m \nvDash y$, and then, $v \vDash \neg y$. Therefore, $v \vDash \langle Q \rangle\, x$ and $v \vDash \neg y$. That means that, considering the valuation ν, $G, \nu \vDash \langle U \rangle (\langle Q \rangle\, x \wedge \neg y)$. ◀

4.6 Non-tangential part

The spatial relation of non-tangential parthood is defined by Cohn and Varzi as follows, for any two regions X and Y: $NTP(X,Y) \equiv P(X,Y)$ & $\forall Z(C(X,Y) \to O(Z,X))$. We show that just one direction of this entailment holds in *BISKT*.

▶ **Theorem 13.** *Let G be a graph and ν a valuation. Given x, y and z propositional variables, the following holds*

$$\text{if } G, \nu \vDash \langle Q \rangle (x) \to y\,,$$

then, for any valuation ν' which agrees with ν on x and y

$$G, \nu' \vDash x \to y$$

and for all the subgraph K of G such that $\llbracket z \rrbracket_{\nu'} = K$

$$\text{if } G, \nu' \vDash \langle U \rangle (\langle Q \rangle\, x \wedge z) \text{ then } G, \nu' \vDash \langle U \rangle (z \wedge y)\,.$$

Proof. Assume $G, \nu \vDash \langle Q \rangle\, x \to y$. Then, for all $w \in W$, if $w \vDash \langle Q \rangle\, x$ then $w \vDash y$, or, if $w \in \llbracket \langle Q \rangle\, x \rrbracket_{\nu}$ then $w \in \llbracket y \rrbracket_{\nu}$. But '$\lrcorner\neg$' is a Cech closure, and by theorem 1, $\lrcorner\neg\llbracket x \rrbracket_{\nu} = \llbracket \langle Q \rangle\, x \rrbracket_{\nu}$.

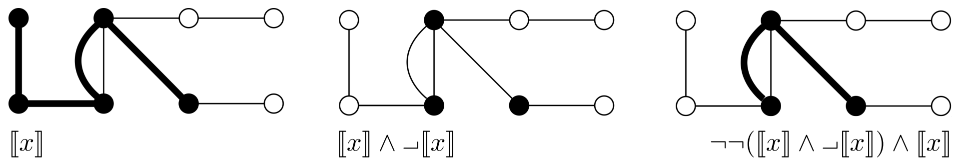


Figure 3 The Boundaries of $[\![x]\!]$.

So if $w \in [\![x]\!]_\nu$ then $w \in [\![\langle Q\rangle\, x]\!]_\nu$, and then $w \in [\![x]\!]_\nu$ implies that $w \in [\![y]\!]_\nu$. So, for all the valuations ν' agreeing with ν on x and y, we have that $G, \nu' \vDash x \to y$. Suppose that for some subgraph K such that $[\![z]\!]_{\nu'} = K$, i) $G, \nu' \vDash \langle U\rangle(\langle Q\rangle\, x \wedge z)$, and ii) $G, \nu' \nvDash \langle U\rangle(x \wedge y)$. So for i) there exists some $v \in W$ such that $v \vDash \langle Q\rangle\, x$ and $v \vDash z$. Since $[\![z]\!]_{\nu'} = K$, $v \in K$, and $K, \nu' \vDash \langle Q\rangle\, x$ and $K, \nu' \vDash z$. For ii) for all $w \in W$, if $w \vDash z$ then $w \nvDash y$. So for all the $k \in K$ $k \nvDash y$, that implies that for all the m such that vHm, $m \nvDash y$. That implies that $v \vDash \neg y$. Therefore $G, \nu' \vDash \langle U\rangle(\langle Q\rangle\, x \wedge \neg y)$. That means x is both non-tangential and tangential part of y. However, in *BISKT* this formula is a tautology: $(\langle Q\rangle\, x \to y) \to \neg(\langle Q\rangle\, x \wedge \neg y)$. So, under the assumption that $G, \nu \vDash \langle Q\rangle\, x \to y$, it cannot be the case that $G, \nu' \vDash \langle U\rangle(\langle Q\rangle\, x \wedge \neg y)$, because $\neg(\langle Q\rangle\, x \wedge \neg y)$ must hold everywhere in the graph. This also means that if a $[\![x]\!]$ is a non tangential part of $[\![y]\!]$, it cannot be also its tangential part. Therefore, under the initial assumption, if a subgraph $[\![z]\!]_{\nu'}$ is such that $G, \nu' \vDash \langle U\rangle(\langle Q\rangle\, x \wedge z$, we must have $G, \nu' \vDash \langle U\rangle(z \wedge y)$.

The other direction of the entailment does not hold. Consider the example of a graph G composed of a node n and an edge e going from n to n itself. Consider $[\![x]\!]_{\nu'} = [\![y]\!]_{\nu'} = n$. We have that $G, \nu' \vDash x \to y$. The possible subgraphs K such that $[\![z]\!]_{\nu'} = K$ are $[\![x]\!]_{\nu'}$, $[\![y]\!]_{\nu'}$ and the whole graph G. All of them are C_2-connected to $[\![x]\!]_{\nu'}$ and overlap $[\![y]\!]_{\nu'}$. Anyway, the closure of $[\![x]\!]_{\nu'}$ is $[\![\langle Q\rangle\, x]\!]_{\nu'} = G$, and G is not part of $[\![y]\!]_{\nu'}$. Therefore $[\![x]\!]_{\nu'}$ is not non-tangential part of $[\![y]\!]_{\nu'}$, and $G, \nu \vDash \langle Q\rangle\, x \to y$ does not hold. ◀

4.7 Boundary and Boundary part

Galton uses a notion of boundary graph already found in [8], that is, in our notation

$$B([\![x]\!]) = [\![x \wedge \lrcorner\, x]\!]\,.$$

However this is not the only notion of boundary possible in our setting, that is different from Galton's one because, in a *BISKT*-graph, multiple edges may occur between a pair of nodes. We examine, in this section the notion of boundary-graph and the spatial relation of boundary part that can be expressed in *BISKT*.

Consider the subgraph $[\![x]\!]$, in bold in Figure 3, with its underlying graph. The subgraph $([\![x \wedge \lrcorner\, x]\!])$ corresponds to the nodes incident with the edges which are not in $[\![x]\!]$. However, it is reasonable to ask that also the edges incident with these nodes, and which belong to $[\![x]\!]$, are considered as part of the boundary of $[\![x]\!]$. We can define another notion of graph boundary as

$$B_\Diamond([\![x]\!]) = [\![\neg\neg(x \wedge \lrcorner\, x) \wedge x]\!]$$

shown in Figure 3.

Cohn and Varzi's definition of the spatial relationship of Boundary part is

$$BP(X, Y) \equiv \forall Z(P(Z, X) \to (TP(Z, Y)))\,.$$

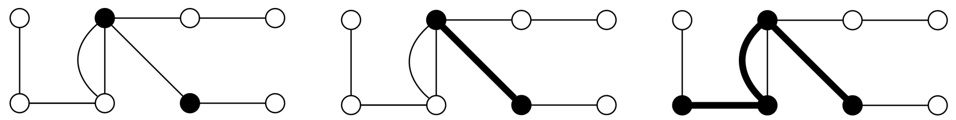


Figure 4 The Boundaries of $[\![x]\!]$.

A region X is part of the boundary of a region Y if for any of its parts Z, Z is a tangential part of Y. This spatial relationship is based on the intuition that the boundary of a region is connected with the *outside* of the region. Translating this definition in our language, the expected definition of Boundary part is the following (we assume any boundary part to be not empty)

$$BP_{\Diamond}([\![x]\!]_{\nu}, [\![y]\!]_{\nu}) \text{ iff for all } [\![z]\!]_{\nu}, \text{ if } NEP_{\Diamond}([\![z]\!]_{\nu}, [\![x]\!]_{\nu}) \text{ then } (TP_{\Diamond}([\![z]\!]_{\nu}, [\![y]\!]_{\nu}))\,.$$

We want to explore whether the definition of graph boundary $B([\![x]\!])$ given above is coherent with the spatial relation of Boundary part.

We notice the following:

(i) The definition of boundary part does not hold when $[\![\neg y]\!]$ is empty. Take the example of a graph composed of two nodes n_1, n_2 and two edges e_1, e_2 incident with them. Consider the subgraph $[\![x]\!] = \{n_1, e_1, n_2\}$. Here $B([\![x]\!]) = [\![x]\!]$ and $[\![\neg x]\!] = \varnothing$. For any $[\![z]\!]$ part of $[\![x]\!]$, $[\![z]\!] \subseteq [\![x]\!]$, the closure of $[\![z]\!]$ is the whole underlying graph, and the intersection of the whole graph with the empty set is empty. Therefore $G, \nu \nvDash \langle U\rangle(\langle Q\rangle z \wedge \neg x)$ is contradictory. This last consideration gives the hint that a better notion of boundary graph is "what leads *outside* of the graph" where the *outside* is $[\![\neg x]\!]$. Therefore, another sensible definition of boundary of x may be " everything which is connected to $[\![\neg x]\!]$".

(ii) If we adopt $B_{\Diamond}([\![x]\!]) = [\![\neg\neg(x \wedge \lrcorner\, x) \wedge x]\!]$ as definition of graph boundary, the notion of boundary part does not hold. Consider, again, the graph $[\![x]\!]$ in figure3. Its negation $\neg x$ is composed of all the nodes not in $[\![x]\!]$, plus the edges incident with them. According to the notion of Boundaty graph, every subgraph which is part of $[\![\neg\neg(x \wedge \lrcorner\, x) \wedge x]\!]$ is such that its closure overlaps $[\![\neg x]\!]$. It is easy to see that this is not true. Just for two of the three nodes are such that their closure overlap $[\![\neg]\!]$.

We put forward other two definitions of graph boundary:

$$B_{\Diamond}^{\neg,N}([\![x]\!]) = [\![x \wedge \langle Q\rangle\neg x]\!] \text{ and } B_{\Diamond}^{\neg}(x) = [\![\neg\neg(x \wedge \langle Q\rangle\neg x) \wedge x]\!]\,.$$

These new definitions single out the boundary subgraphs which are connected with $[\![\neg x]\!]$, and which support the definition of boundary part of [3]. The former considers just the nodes adjacent with edges adjacent with $[\![\neg x]\!]$, the latter adds also the edges between those nodes as shown in Figure 4. Eventually, another possible notions of boundary part is

$$B^{\lrcorner}([\![x]\!]) = [\![x \wedge \langle Q\rangle \lrcorner\, x]\!]$$

shown also in Figure 4.

5 Conclusions and Further Work

We have examined discrete space from the viewpoint of a modal logic based on relations on graphs, rather than on sets, as the accessibility relations. This has enabled us to bring together for the first time the approach to spatial reasoning using a modal logic based in mathematical morphology proposed by Bloch, with the mereotopological analysis of discrete space developed by Galton.

We have shown that the general framework of Cohn and Varzi can be generalized to accommodate discrete spatial relationships, but that closure operators which satisfy all of the Kuratowski axioms cannot be used to describe the notion of connection in some discrete spaces. By adopting the less restrictive version of closure due to Cech we have been able to realize the connection described by Galton as a C_2 connection in the framework of Cohn and Varzi.

The specific form of closure needed can be expressed as the negation, ¬, and dual negation, ⌐, in the logic BISKT. The combination of the semantic counterparts of these operations to express the idea of extending a subgraph by one step along the links is by no means new. This was already noted at COSIT 1997 by Stell and Worboys [16] citing the work of Reyes and Zolfaghari [11]. However, in the present work we have been able to express this closure as a modality using the stable relation Q which describes the incidence structure of the graph. Reyes and Zolfaghari [11] view this closure in a modal setting quite different from our use of stable relations on graphs. By working within the context of the BISKT logic we have been able to express not only connection itself, but other spatial relations including non-tangential parthood and a variety of notions of boundary.

In our formulation stating that two regions are connected is expressible through a formula in our logic. This depends on being able to express non-emptiness, which we achieve through the universal modality $\langle U \rangle$, "somewhere". In the setting of Bloch [2] a relation such as tangential part is expressed not just by a formula holding but by one formula holding and another being consistent. Expressing mereotopological relations entirely within our logic can be expected to facilitate the use of automated reasoning tools for modal logics, such as in [15], for spatial reasoning. We will explore this in further work, as well as extending our analysis to a wider range of relationships and examining these with notions of uncertainty and vagueness for discrete spatial regions.

References

1 P. Blackburn, J. F. van Benthem, and F. Wolter. *Handbook of modal logic*, volume 3. Elsevier, 2006.

2 I. Bloch. Modal logics based on mathematical morphology for qualitative spatial reasoning. *Journal of Applied Non-Classical Logics*, 12(3–4):399–423, 2002.

3 A. G. Cohn and A. C. Varzi. Mereotopological connection. *Journal of Philosophical Logic*, 32:357–390, 2003.

4 J. Cousty, L. Najman, F. Dias, and J. Serra. Morphological filtering on graphs. *Computer Vision and Image Understanding*, 117:370–385, 2013.

5 M. J. Egenhofer and J. Herring. Categorizing binary topological relations between regions, lines, and points in geographic databases. *Department of Surveying Engineering, , University of Maine, Orono, ME*, 9(94-1):76, 1991.

6 A. Galton. The mereotopology of discrete space. In C. Freksa and D. Mark, editors, *COSIT'99 proceedings*, volume 1661 of *LNCS*, pages 251–266. Springer, 1999.

7 Antony Galton. Discrete mereotopology. In *Mereology and the Sciences*, pages 293–321. Springer, 2014.

8 F. W. Lawvere. Intrinsic co-heyting boundaries and the leibniz rule in certain toposes. In *Category Theory*, pages 279–281. Springer, 1991.

9 L. Najman and H. Talbot. *Mathematical Morphology. From theory to applications.* Wiley, 2010.

10 D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In B. Nebel, C. Rich, and W. Swartout, editors, *Principles of Knowledge Representation and Reasoning. Proceedings of the Third International Conference (KR92)*, pages 165–176. Morgan Kaufmann, 1992.

11 G. E. Reyes and H. Zolfaghari. Bi-Heyting algebras, toposes and modalities. *Journal of Philosophical Logic*, 25:25–43, 1996.

12 Azriel Rosenfeld. Digital topology. *American Mathematical Monthly*, pages 621–630, 1979.

13 P. Simons. *Parts. A Study in Ontology.* Clarendon Press, Oxford, 1987.

14 M. B. Smyth. Topology. In *Handbook of Logic in Computer Science, Vol 1*, pages 641–761. Clarendon Press, Oxford, 1992.

15 J. G. Stell, R. A. Schmidt, and D. Rydeheard. A bi-intuitionistic modal logic: Foundations and automation. *J. Logical and Algebraic Methods in Programming*, 85(4):500–519, 2016.

16 J. G. Stell and M. F. Worboys. The algebraic structure of sets of regions. In S. C. Hirtle and A. U. Frank, editors, *Spatial Information Theory, International Conference COSIT'97, Proceedings*, volume 1329 of *LNCS*, pages 163–174. Springer, 1997.

17 D. Vakarelov, G. Dimov, I. Düntsch, and B. Bennett. A proximity approach to some region-based theories of space. *Journal of Applied Non-Classical Logics*, 12:527–559, 2002.

A New Perspective on the Mereotopology of RCC8

Michael Grüninger[1] and Bahar Aameri[2]

1 **University of Toronto, Toronto, ON, Canada**
gruninger@mie.utoronto.ca
2 **University of Toronto, Toronto, ON, Canada**
bahar@cs.toronto.edu

Abstract

RCC8 is a set of eight jointly exhaustive and pairwise disjoint binary relations representing mereotopological relationships between ordered pairs of individuals. Although the RCC8 relations were originally presented as defined relations of Region Connection Calculus (RCC), virtually all implementations use the RCC8 Composition Table (CT) rather than the axioms of RCC. This raises the question of which mereotopology actually underlies the RCC8 composition table. In this paper, we characterize the algebraic and mereotopological properties of the RCC8 CT based on the metalogical relationship between the first-order theory that captures the RCC8 CT and Ground Mereotopology (MT) of Casati and Varzi. In particular, we show that the RCC8 theory and MT are relatively interpretable in each other. We further show that a nonconservative extension of the RCC8 theory that captures the intended interpretation of the RCC8 relations is logically synonymous with MT, and that a conservative extension of MT is logically synonymous with the RCC8 theory. We also present a characterization of models of MT up to isomorphism, and explain how such a characterization provides insights for understanding models of the RCC8 theory.

1998 ACM Subject Classification F.4.1 Mathematical Logic, I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases RCC8, mereotopology, spatial reasoning, ontologies

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.2

1 Introduction

Representations of space, and their use in qualitative spatial reasoning, are widely recognized as key aspects in commonsense reasoning, with applications ranging from biology to geography. The predominant approach to spatial representation within the applied ontology community has used mereotopologies, which combine topological (expressing connectedness) with mereological (expressing parthood) relations. A variety of first-order mereotopological ontologies have been proposed, the most widespread being the Region Connection Calculus (RCC) [17], the ontology RT [1], and the ontologies introduced by Casati and Varzi [4]. Properties of RCC in particular have been studied extensively; [18, 5] present algebraic representations for the RCC theory, and [9] describes various mereotopological settings that satisfy axioms of RCC.

While theoretical work has focused on the first-order theories for mereotopologies, work within the qualitative spatial reasoning community has primarily used a formalism known as RCC8, which is a set of eight jointly exhaustive and pairwise disjoint binary relations representing mereotopological relationships between ordered pairs of individuals. Reasoning

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 2; pp. 2:1–2:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

is supported through the use of a composition table, which specifies all possible mereotopological relationships between pairs of elements; deduction is implemented through constraint propagation algorithms.

Although the RCC8 relations were originally presented as defined relations within RCC, the theoretical analyses of RCC have not been helpful in understanding properties of formalisms that use the RCC8 relations. The reason is that virtually all implementations use the RCC8 composition table rather than the axioms of RCC, and the composition table has very different mereotopological properties than RCC. Of particular importance is the widespread use of RCC8 in efforts such as GeoSPARQL, which is an international standard for the representation of geospatial linked data developed by the Open Geospatial Consortium. A characterization of all solutions for a set of RCC8 constraints presumes an understanding of the possible models of some first-order logical theory.

In this paper, we investigate algebraic and mereotopological properties of the RCC8 composition table based on the metalogical relationship between the first-order theory that captures the RCC8 composition table and Ground Mereotopology (MT) of Casati and Varzi. After reviewing the basic axiomatizations of the mereotopological theories in Section 2, we discuss the relationship between the RCC8 theory and MT in Section 3. Our key result is that a nonconservative extension of the RCC8 theory, we called RCC8*, is logically synonymous with the MT theory, meaning MT and RCC8* axiomatize the same class of structures. In other words, MT and RCC8* are semantically equivalent, and only differ in signature (i.e., the non-logical symbols). Further, we present a conservative extension of MT which is logically synonymous with the RCC8 theory. We also show that the RCC8 theory and MT are relatively interpretable in each other. Finally, in Section 4, we present a characterization of models of MT up to isomorphism, and explain how such a characterization can be used in characterizing algebraic properties of models of the RCC8 theory.

2 Preliminaries: Mereotopological Theories

2.1 Region Connection Calculus

The Region Connection Calculus (RCC) is a first-order theory whose signature contains the single primitive binary relation $C(x, y)$ denoting "x is *connected* to y". Parthood is defined in terms of connection alone, being equivalent to the topological notion of enclosure. Representation theorems [18] have shown that the models of RCC are equivalent to mathematical structures known as Boolean contact algebras which consist of a standard Boolean algebra together with a binary relation C that is reflexive, anti-symmetric, and extensional.

2.2 RCC8

RCC8 is a set of eight binary relations representing mereotopological relationships between (ordered) pairs of individuals. These relations and their intended interpretations are illustrated in Figure 1. The RCC8 relations have been proven to be jointly exhaustive and pairwise disjoint (JEPD), that is, every ordered pair of individuals are related by exactly one RCC8 relation.

Originally, RCC8 relations were presented as defined relations of RCC (throughout the paper, free variables in a displayed formula are assumed to be universally quantified):

$$DC(x, y) \equiv \neg C(x, y). \tag{1}$$

$$EC(x, y) \equiv C(x, y) \wedge \neg O(x, y). \tag{2}$$

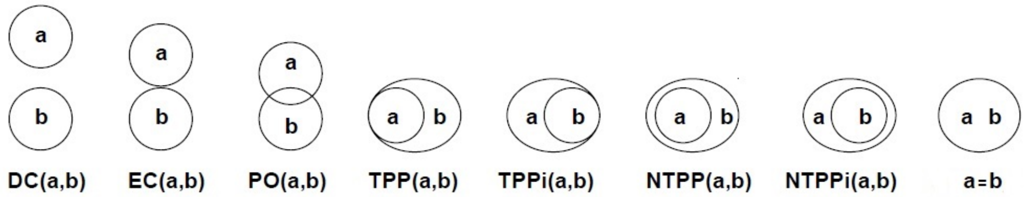


Figure 1 Illustration of RCC8 relations – $DC(a, b)$ (a is *disconnected* from b), $EC(a, b)$ (a is *externally connected* with b), $PO(a, b)$ (a *partially overlaps* b), $TPP(a, b)$ (a is a *tangential proper part* of b), $TPPi(a, b)$ (b is a *tangential proper part* of a), $NTPP(a, b)$ (a is a *nontangential proper part* of b), $NTPPi(a, b)$ (b is a *nontangential proper part* of a), $a = b$ (a is *identical* with b).

	DC	**EC**	**PO**	**TPP**	**NTPP**	**TPPi**	**NTPPi**	=
DC	*	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC	DC	DC
EC	DC, EC, PO, TPPi, NTPPi	DC, EC, PO, TPP, TPPi, =	DC, EC, PO, TPP, NTPP	EC, PO, TPP, NTPP	PO, TPP, NTPP	DC, EC	DC	EC
PO	DC, EC, PO, TPPi, NTPPi	DC, EC, PO, TPPi, NTPPi	*	PO, TPP, NTPP	PO, TPP, NTPP	DC, EC, PO, TPPi, NTPPi	DC, EC, PO, TPPi, NTPPi	PO
TPP	DC	DC, EC	DC, EC, PO, TPP, NTPP	TPP, NTPP	NTPP	DC, EC, PO, TPP, TPPi, =	DC, EC, PO, TPPi, NTPPi	TPP
NTPP	DC	DC	DC, EC, PO, TPP, NTPP	NTPP	NTPP	DC, EC, PO, TPP, NTPP	*	NTPP
TPPi	DC, EC, PO, TPPi, NTPPi	EC, PO, TPPi, NTPPi	PO,TPPi,NTPPi	PO, TPP, TPPi, =	PO, TPP, NTPP	TPPi, NTPPi	NTPPi	TPPi
NTPPi	DC, EC, PO, TPPi, NTPPi	PO,TPPi,NTPPi	PO,TPPi,NTPPi	PO,TPPi,NTPPi	PO, TPP, NTPP, TPPi, NTPPi, =	NTPPi	NTPPi	NTPPi
=	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	=

Figure 2 RCC8 Composition Table. "$*$" indicates that all RCC8 relations are possible.

$$PO(x, y) \equiv O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x). \tag{3}$$
$$(x = y) \equiv P(x, y) \wedge P(y, x). \tag{4}$$
$$TPPi(x, y) \equiv TPP(y, x). \tag{5}$$
$$NTPPi(x, y) \equiv NTPP(y, x). \tag{6}$$
$$TPP(x, y) \equiv PP(x, y) \wedge \neg NTPP(x, y). \tag{7}$$
$$NTPP(x, y) \equiv PP(x, y) \wedge \neg(\exists z)\, EC(z, x) \wedge EC(z, y). \tag{8}$$

In the axioms above, $C(x, y)$ denotes "x is *connected* to y," $P(x, y)$ denotes "x is a *part* of y," $O(x, y)$ denotes "x *overlaps* y," $PP(x, y)$ denotes "x is a *proper part* of y"):

$$O(x, y) \equiv (\exists z)\, P(z, x) \wedge P(z, y). \tag{9}$$
$$PP(x, y) \equiv P(x, y) \wedge \neg P(y, x). \tag{10}$$

Given its origin within RCC, it is interesting to note that RCC8 is typically used independently of the RCC theory – the RCC axioms are not considered to be part of the RCC8 formalism, and in most reasoning tasks even the axiomatic descriptions of RCC8 relations are not explicitly used. Instead, the RCC8 Composition Table (CT) is used. The RCC8 CT (illustrated in Figure 2) is an 8×8 matrix such that for each ordered pair of RCC8 relations R_i, R_j, the cell $CT(R_i, R_j)$ indicates possible mereotopological relationships between two individuals a and c assuming that $R_i(a, b)$ and $R_j(b, c)$ holds. For example, $CT(EC, NTPP) = \{PO, TPP, NTPP\}$, meaning that if $EC(a, b)$ and $NTPP(b, c)$, then a is related to c by either PO or TPP or $NTPP$.

2.3 Combined Mereotopology

Even though the RCC8 CT is entailed by the RCC theory, they have very different mereotopological properties. In fact, the RCC8 CT seems to be closely related to Ground Mereotopology (also called MT), which is the weakest theory among the mereotopological theories proposed in [4]. The signature of the MT theory (which we will denote by T_{mt}) consist of two primitive binary relations, parthood (P) and connection (C). The axioms of the theory (Axioms 11 to 16) state that connection is a reflexive and symmetric relation, while parthood is a reflexive, transitive, and anti-symmetric relation.[1] In addition, if one individual is connected to another, then the first one is also connected to any individual which the second is part of.

$$C(x,x). \tag{11}$$

$$C(x,y) \supset C(y,x). \tag{12}$$

$$P(x,x). \tag{13}$$

$$P(x,y) \wedge P(y,x) \supset (x = y). \tag{14}$$

$$P(x,y) \wedge P(y,z) \supset P(x,z). \tag{15}$$

$$P(y,z) \wedge C(x,y) \supset C(x,z). \tag{16}$$

3 Relationship between MT and RCC8

Even though the RCC8 CT has been derived based on the RCC theory, they show very different mereotopological properties. For instance, while in models of RCC every individual is atomless (i.e., has a proper part) and externally connected to another individual, individuals that satisfy the RCC8 CT may have no proper part, or may not be connect to any other individual. These differences raise the question of which mereotopology actually underlies the RCC8 composition table.

We begin this section by describing the logical theory that captures RCC8 CT. We then show that the closest mereotopology to this theory is MT.

3.1 The First-order Theory of RCC8

We denote the logical theory of RCC8 CT by T_{rcc8}. Following [2], we assume that for each cell in the RCC8 CT, T_{rcc8} contains an axiom of the following form

$$R_i(x,y) \wedge R_j(y,z) \supset T_1(x,z) \vee ... \vee T_n(x,z)$$

where $CT(R_i, R_j) = \{T_1, ..., T_n\}$. The following sentence, for example, is the axiom of T_{rcc8} which corresponds with $CT(TPP, EC)$:

$$TPP(x,y) \wedge EC(y,z) \supset DC(x,z) \vee EC(x,z).$$

Since RCC8 CT consists of 8×8 cells, T_{rcc8} must contain 64 axioms corresponding with the table. In addition to these axioms, we assume that T_{rcc8} contains an axiom that specifies RCC8 relations are jointly exhaustive:

$$DC(x,y) \vee EC(x,y) \vee PO(x,y) \vee NTPP(x,y) \vee TPP(x,y) \vee TPPi(x,y) \vee$$
$$NTPPi(x,y) \vee (x = y).$$

[1] In this paper, we consider a *theory* to be a set of first-order sentences closed under logical entailment. A collection of sentences of a theory which entail all other sentences in the theory are called *axioms* of the theory.

We also assume that for each RCC8 relation R_1, T_{rcc8} contains a sentence of the following form stating that RCC8 relations are pairwise disjoint (PD):

$$R_1(x, y) \supset \neg(R_2(x, y) \vee ... \vee R_7(x, y))$$

where $R_2, ..., R_7$ are RCC8 relations other than R_1. The following sentence, for example, is the PD axiom corresponding with DC:

$$DC(x, y) \supset \neg[EC(x, y) \vee PO(x, y) \vee TPP(x, y) \vee NTPP(x, y) \vee$$
$$TPPi(x, y) \vee NTPPi(x, y) \vee (x = y)].$$

As there are 8 RCC8 relations, T_{rcc8} contains 8 PD axioms. All other sentences in T_{rcc8} are those which are entailed by the $64 + 1 + 8$ above-mentioned axioms.

3.2 MT Theory vs. RCC8 Theory

The RCC8 CT is commonly considered to be related to RCC because RCC8 relations were originally defined as part of the RCC theory, and the RCC8 CT was proved using the RCC theory. It turns out, however, that the RCC8 CT can also be deduced from a definitional extension of MT:

▶ **Definition 1** (adopted from [11]). Let T be a first-order theory and Π be a set containing sentences of the following form[2]

$$R(x_1, ..., x_n) \equiv \Phi(x_1, ..., x_n)$$

where R is a predicate which is not in $\Sigma(T)$ and Φ is a formula in $\mathcal{L}(T)$ in which at most variables $x_1, ..., x_n$ occur free. $T \cup \Pi$ is called a *definitional extension* of T.

Notice that Definitions 1 to 10 are defined in terms of C and P, which are primitives of T_{mt}, and so if we extend T_{mt} by Definitions 1 to 10, we get a definitional extension of T_{mt}. This extension entails T_{rcc8}.

▶ **Theorem 2.** *T_{rcc8} is entailed by a definitional extension of T_{mt}.*

Proof. Suppose Π denotes the set containing Definitions (1) to (10). Using an automated theorem prover, Prover9 [13], we showed that $T_{mt} \cup \Pi$ entails axioms of T_{rcc8}. Hence, $T_{mt} \cup \Pi$ entails T_{rcc8}. ◀

Recall that DC, EC, PO, TPP, $NTPP$, $TPPi$, $NTPPi$, and $=$ are primitives of T_{rcc8}. Using these primitives, one can extend T_{rcc8} with the following definitions for parthood and connection:

$$P(x, y) \equiv TPP(x, y) \vee NTPP(x, y) \vee (x = y). \tag{17}$$

$$C(x, y) \equiv \neg DC(x, y). \tag{18}$$

A more interesting result is that this definitional extension of T_{rcc8} actually entails MT:

▶ **Theorem 3.** *T_{mt} is entailed by a definitional extension of T_{rcc8}.*

[2] For a theory T, $\Sigma(T)$ denotes the *signature* of T, i.e., the set of non-logical symbols used in sentences of T; $\mathcal{L}(T)$ denotes the *language* of T, i.e., the set of all first-order formulae generated by symbols in $\Sigma(T)$; $Mod(T)$ denotes the class of all models of T.

Proof. Suppose Δ denotes the set containing Definitions (17) and (18). Using Prover9, we showed that $T_{rcc8} \cup \Delta$ entails Axioms (11) to (16). Since Axioms (11) to (16) axiomatize T_{mt}, we can conclude that $T_{rcc8} \cup \Delta$ entails T_{mt}. ◀

Using Theorems 2 and 3, it can be shown that T_{mt} and T_{rcc8} are *relatively interpretable* [7] in each other. Informally, a theory T_1 has a relative interpretation in another theory T_2 if every sentence in T_1 can be translated into a sentence in T_2. In other words, for all sentences $\Phi \in \mathcal{L}(T_1)$, if T_1 entails Φ, then T_2 entails a translation of Φ into the language of T_2. [10] show that if a definitional extension of T_2 entails T_1, translations for sentences of T_1 is obtained based on the formulas which define predicates of T_1 in the definitional extension. For instance, a translation of Axiom (16) of T_{mt} into the language of T_{rcc8} can be obtained by replacing C and P with the formulas defining them in Definitions (17) and (18). The result is the following sentence, which provably is a sentence in T_{rcc8}:

$$\neg DC(x,y) \wedge (TPP(y,z) \vee NTPP(y,z) \vee (y=z)) \\ \supset \neg DC(x,z).$$

When T_1 is interpretable in T_2, every model of T_2 defines a model of T_1 using the translation definitions between T_1 and T_2 [7]. Consider a model $\mathcal{M}_1$ of T_{rcc8} with two elements $\mathbf{a}, \mathbf{b}$ that are externally connected:[3]

$\mathbf{EC}^{\mathcal{M}_1} = \{(\mathbf{a},\mathbf{b}),(\mathbf{b},\mathbf{a})\}$.

Now, consider a structure $\mathcal{N}_1$ with the same domain but in the signature of T_{mt} such that relations between elements are obtained based on $\mathcal{M}_1$ and Definitions (17) and (18). By Definitions (17) and (18), for any pair $\mathbf{x}, \mathbf{y}$:

$$\begin{aligned} (\mathbf{x},\mathbf{y}) \in \mathbf{C}^{\mathcal{N}_1} \quad &\text{iff} \quad (\mathbf{x},\mathbf{y}) \notin \mathbf{DC}^{\mathcal{M}_1}, \\ (\mathbf{x},\mathbf{y}) \in \mathbf{P}^{\mathcal{N}_1} \quad &\text{iff} \quad \Big[(\mathbf{x},\mathbf{y}) \in \mathbf{TPP}^{\mathcal{M}_1} \text{ or} \\ &\qquad (\mathbf{x},\mathbf{y}) \in \mathbf{NTPP}^{\mathcal{M}_1} \text{ or } \mathbf{x}=\mathbf{y}\Big]. \end{aligned}$$

So, $\mathbf{C}^{\mathcal{N}_1} = \{(\mathbf{a},\mathbf{a}),(\mathbf{b},\mathbf{b}),(\mathbf{a},\mathbf{b}),(\mathbf{b},\mathbf{a})\}$ and $\mathbf{P}^{\mathcal{N}_1} = \{(\mathbf{a},\mathbf{a}),(\mathbf{b},\mathbf{b})\}$.

To study models of T_{rcc8} based on models of T_{mt}, we need a notion stronger than relative interpretation:

▶ **Definition 4** ([11])**.** Two theories T_1 and T_2 are logically synonymous iff they have a common definitional extension.

Considering Definition 4, it is easy to see that T_1 and T_2 are synonymous iff there exist two sets of translation definitions, Δ and Π, such that $T_1 \cup \Pi$ is a definitional extension of T_1, $T_2 \cup \Delta$ is a definitional extension of T_2, and $T_1 \cup \Pi$ and $T_2 \cup \Delta$ are logically equivalent.

T_{mt} and T_{rcc8} are not synonymous. In the following part of this section we will explain why, and present an extension of T_{rcc8} which is synonymous with T_{mt}.

3.3 MT and RCC8*

When two theories are synonymous, there is a one-to-one correspondence between their models such that the corresponding models can be defined based on each other [15]. Such a

[3] We denote *structures* by calligraphic uppercase letters, e.g. $\mathcal{M}, \mathcal{N}$; elements of a structure by **boldface** font, e.g., $\mathbf{a}, \mathbf{b}$; and the *extension of predicate* R in a structure $\mathcal{M}$ by $\mathbf{R}^{\mathcal{M}}$.

correspondence does not exists between models of T_{mt} and T_{rcc8}. Consider two models $\mathcal{M}_2$ and $\mathcal{M}_3$ of T_{rcc8}, both with two elements $\mathbf{a}, \mathbf{b}$ such that

$$\mathbf{TPP}^{\mathcal{M}_2} = \{(\mathbf{a}, \mathbf{b})\}, \qquad \mathbf{NTPP}^{\mathcal{M}_3} = \{(\mathbf{a}, \mathbf{b})\}.$$

Both $\mathcal{M}_2$ and $\mathcal{M}_3$ define the same model $\mathcal{N}_2$ of T_{mt}:

$$\mathbf{C}^{\mathcal{N}_2} = \{(\mathbf{a}, \mathbf{a}), (\mathbf{b}, \mathbf{b}), (\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a})\}, \quad \mathbf{P}^{\mathcal{N}_2} = \{(\mathbf{a}, \mathbf{a}), (\mathbf{b}, \mathbf{b}), (\mathbf{a}, \mathbf{b})\}.$$

$\mathcal{M}_2$ and $\mathcal{M}_3$ correspond with the same model of T_{mt} because the only way for MT to distinguish TPP from $NTPP$ is the existence of a third element that is externally connected to the inner element (i.e., $\mathbf{a}$). However, such an element does not exists in either of $\mathcal{M}_2$ and $\mathcal{M}_3$.

A similar issue arises when two individuals overlap, but they do not have a common part. Consider a model $\mathcal{M}_4$ of T_{rcc8} with two elements $\mathbf{a}, \mathbf{b}$ and $\mathbf{PO}^{\mathcal{M}_4} = \{(\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a})\}$.

$\mathcal{M}_4$ defines the following model of T_{mt}, which is isomorphic to $\mathcal{N}_1$ in the previous subsection:

$$\mathbf{C}^{\mathcal{N}_4} = \{(\mathbf{a}, \mathbf{a}), (\mathbf{b}, \mathbf{b}), (\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a})\}, \quad \mathbf{P}^{\mathcal{N}_4} = \{(\mathbf{a}, \mathbf{a}), (\mathbf{b}, \mathbf{b})\}.$$

Thus $\mathcal{M}_1$ and $\mathcal{M}_4$ correspond with the same model of T_{mt}. This is because within MT, 'overlap' is defined based on a third element that is a common part of the overlapping individuals. If such an element does not exists (as is the case with $\mathcal{M}_4$), MT cannot distinguish PO from EC.

It is interesting to observe that although $\mathcal{M}_2$ and $\mathcal{M}_4$ are models of T_{rcc8}, they do not satisfy the original (axiomatic) definitions of PO, TPP or O (i.e., Definitions (3), (7), and (9)); that is definitions which are part of the RCC theory, and the RCC8 CT is derived based on them. Notice also that no model of T_{mt} defines $\mathcal{M}_2$ and $\mathcal{M}_4$ because without the existence of a third element TPP and O are not definable in MT.

Since a one-to-one correspondence between models of RCC8 and MT does not exists, they are not synonymous. To get synonymy, we need to extend T_{rcc8} by axioms that eliminate those models of T_{rcc8} which are not definable by any model of T_{mt}. Based on the examples we just discussed, undefinable models are those that do not satisfy the axiomatic definitions of TPP or O: That is, models (like $\mathcal{M}_2$) in which an element is related to another element by TPP, but there is no other element that externally connects with the inner element; or models (like $\mathcal{M}_4$) in which two elements are related by O, but they do not have a common part. To eliminate such models, we extend T_{rcc8} by the following axioms:

$$TPP(x, y) \supset (\exists z) EC(z, x) \wedge EC(z, y). \tag{19}$$

$$O(x, y) \supset (\exists z)\, P(z, x) \wedge P(z, y). \tag{20}$$

We call the resulting theory RCC8* and denote it by T_{rcc8^*}.

▶ **Theorem 5.** *T_{mt} is logically synonymous with T_{rcc8^*}.*

Proof. Suppose Π contains Definitions (1) to (10); and Δ contains Definitions (17), (18), (21), (22).

$$O(x, y) \equiv \neg DC(x, y) \wedge \neg EC(x, y). \tag{21}$$

$$PP(x, y) \equiv TPP(x, y) \vee NTPP(x, y). \tag{22}$$

Using Prover9, we showed

$$T_{mt} \cup \Pi \models T_{rcc8^*} \cup \Delta \qquad \text{and} \qquad T_{rcc8^*} \cup \Delta \models T_{mt} \cup \Pi\,.$$

Hence, $T_{mt} \cup \Pi$ and $T_{rcc8^*} \cup \Delta$ are logically equivalent. So, by definition, T_{mt} and T_{rcc8^*} are logically synonymous. ◀

According to [15], synonymous theories axiomatize the same class of structures. Thus, T_{mt} and T_{rcc8^*} are semantically equivalent and only differ in signature.

All relations in RCC8 CT can be deduced from T_{rcc8^*} as it is an extension of T_{rcc8}. In addition, for every entry $CT(R_i, R_j)$ of the RCC8 CT and every RCC8 relation $S \notin CT(R_i, R_j)$ we proved a sentence of the following form (proofs are done by Prover9):

$$R_i(x, y) \wedge R_j(y, z) \supset \neg S(x, z)\,.$$

Thus, the additional axioms of RCC8* does not change RCC8 CT, but only eliminate those models of T_{rcc8} that do not satisfy the axiomatic definitions of RCC8 relations.

3.4 Extending MT

As we explained in Section 3.3, logical synonymy between MT and RCC8 does not achieved because of the way $NTPP$, TPP, PO and EC are defined within the MT theory: The difference between $NTPP$ and TPP is defined with respect to a third element. Hence, only models of MT with more than two elements can distinguish between $NTPP$ and TPP. However, within the RCC8 theory $NTPP$ and TPP are distinguishable even by models of size two. A similar arguments applies to PO and EC. Thus, a one-to-one correspondence between models of MT and RCC8 does not exist.

In Section 3.3 we demonstrate how extending T_{rcc8} to T_{rcc8*} gives a one-to-one correspondence between models of T_{mt} and T_{rcc8*}, meaning that T_{mt} and T_{rcc8*} are logically synonymous. Another way of getting logical synonymy is to extend MT with Axioms (23) to (30), which specify properties of $NTPP$ and O (Axioms (23) to (26) are borrowed from [8]). We call the resulting theory MTNO and denote it by T_{mtno}.

$$NTPP(x, y) \wedge P(y, z) \supset NTPP(x, z). \tag{23}$$
$$P(x, y) \wedge NTPP(y, z) \supset NTPP(x, z). \tag{24}$$
$$NTPP(x, y) \supset PP(x, y). \tag{25}$$
$$C(x, y) \wedge NTPP(y, z) \supset O(x, z). \tag{26}$$
$$O(x, x). \tag{27}$$
$$O(x, y) \supset O(y, x). \tag{28}$$
$$O(x, y) \supset C(x, y). \tag{29}$$
$$O(x, y) \wedge P(y, z) \supset O(x, z). \tag{30}$$

Since O and $NTPP$ are primitive relations in T_{mtno}, PO and TPP can be defined based on them, without introducing a third element:

$$PO(x, y) \equiv O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x). \tag{31}$$
$$TPP(x, y) \equiv P(x, y) \wedge \neg P(y, x) \wedge \neg NTPP(x, y). \tag{32}$$

Therefore, a one-to-one correspondence between models of T_{mtno} and T_{rcc8} should exist, and it should be possible to show that the two theories are logically synonymous.

▶ **Theorem 6.** *T_{mtno} is logically synonymous with T_{rcc8}.*

Proof. To show that T_{mtno} and T_{rcc8} are synonymous we need to show that there exist conservative definitions Π and Δ for T_{mtno} and T_{rcc8} such that $T_{mtno} \cup \Pi$ and $T_{rcc8} \cup \Delta$ are logically equivalent.

Suppose Π contains Definitions (1) to (6) and (31) to (32); and Δ contains Definitions (17), (18), (21), (22). Using Prover9, we showed that

$$T_{mtno} \cup \Pi \models T_{rcc8} \cup \Delta \quad \text{and} \quad T_{rcc8} \cup \Delta \models T_{mtno} \cup \Pi\,.$$

Hence, $T_{mtno} \cup \Pi$ and $T_{rcc8} \cup \Delta$ are logically equivalent. ◀

4 Model-Theoretic Characterization of MT

Is the equivalence between RCC8* and MT simply an intellectual curiosity, or does it give us new insights into RCC8? If we consider that RCC8 is primarily used in constraint satisfaction problems, in which one constructs a satisfying interpretation of a set of expressions in the signature of RCC8, then the set of all possible solutions of RCC8 problems, excluding those eliminated by RCC8*, is equivalent to the set of all possible models of T_{mt}. In this section, we provide a characterization of the models of T_{mt} up to isomorphism, by first specifying a class of mathematical structures, and then showing that T_{mt} axiomatizes this class of structures.

4.1 Representation Theorem for Models of T_{mt}

We begin by introducing the two classes of mathematical structures that capture the intended interpretations of the the connection and parthood relations in MT. The connection relation in T_{mt} corresponds to a class of graphs:

▶ **Definition 7.** A graph with loops is a pair $\mathcal{G} = \langle V, E \rangle$ of sets such that:

1. $E \subseteq V \times V$.
2. For each $\mathbf{v} \in V$, $\mathbf{v} \in N(\mathbf{v})$, where $N(\mathbf{x})$, $\mathbf{x} \in V$, denotes the set of neighbors of $\mathbf{x}$ and is defined as

$$N(\mathbf{x}) = \{\mathbf{y} \,:\, (\mathbf{x}, \mathbf{y}) \in E\}\,.$$

$\mathfrak{M}^{graph\_loops}$ is the class of structures which are graphs with loops.

It is well-known that Ground Mereology, the subtheory of T_{mt} which describes the parthood relation, is synonymous with the theory of partial orderings [4]. That is, the parthood relation in models of T_{mt} forms a partial ordering:

▶ **Definition 8.** A partial ordering is a pair $\mathcal{Q} = \langle V, \preceq \rangle$ s.t. $\preceq$ is a reflexive, antisymmetric, and transitive binary relation. For each $\mathbf{x} \in V$ and each set $X \subseteq V$ the upper set, denoted by $U(\mathbf{x})$ and $U(X)$ respectively, is defined as

$$U(\mathbf{x}) = \{\mathbf{y} \,:\, \mathbf{x} \preceq \mathbf{y}\} \qquad U(X) = \bigcup_{\mathbf{x} \in X} U(\mathbf{x})\,.$$

$\mathfrak{M}^{par\_orders}$ denotes the class of partial orderings.

We pull all of these ideas together to define the class of mathematical structures which we will eventually show are equivalent to the models of T_{mt}:

▶ **Definition 9.** $\mathfrak{M}^{mt}$ is the following class of structures. $\mathcal{M} \in \mathfrak{M}^{mt}$ iff $\mathcal{M} = \langle V, E, \preceq \rangle$ such that
1. $\mathcal{G} = \langle V, E \rangle$ and $\mathcal{G} \in \mathfrak{M}^{graph\_loops}$;
2. $\mathcal{Q} = \langle V, \preceq \rangle$ and $\mathcal{Q} \in \mathfrak{M}^{par\_orders}$;
3. $U(N(\mathbf{x})) \subset N(\mathbf{x})$, for each $\mathbf{x} \in V$.

Condition (3) constrains how the two graph and partial ordering substructures are related to each other – the neighborhood of a vertex in the graph is closed under upper sets in the partial ordering. An example of a structure in $\mathfrak{M}^{mt}$ can be seen in the graph of Figure 3(i) and the corresponding partial ordering in Figure 3(ii); note that the vertices in the graph are the elements of the partial ordering.

The following theorem shows that there is a one-to-one correspondence between the models of T_{mt} and class of structures $\mathfrak{M}^{mt}$ that capture the intended semantics of the mereotopology of MT.

▶ **Theorem 10.** *There exists a bijection*

$$\varphi : Mod(T_{mt}) \rightarrow \mathfrak{M}^{mt}$$

such that
1. *the domain of $\mathcal{M}$ and $\varphi(\mathcal{M})$ are the same;*
2. $(\mathbf{x}, \mathbf{y}) \in \mathbf{C}^{\mathcal{M}}$ *iff* $(\mathbf{x}, \mathbf{y}) \in E^{\varphi(\mathcal{M})}$*;*
3. $(\mathbf{x}, \mathbf{y}) \in \mathbf{P}^{\mathcal{M}}$ *iff* $\mathbf{x} \preceq^{\varphi(\mathcal{M})} \mathbf{y}$.

Suppose $\mathcal{M} \in Mod(T_{mt})$ and $\mathcal{N} = \varphi(\mathcal{M})$. Then $\mathcal{N} \in \mathfrak{M}^{mt}$, and the domain of $\mathcal{M}$ and $\mathcal{N}$ are the same. For each element $\mathbf{x}$ in $\mathcal{M}$ and $\mathcal{N}$, the neighbors of $\mathbf{x}$ in the graph (i.e., $N(\mathbf{x})$) in $\mathcal{N}$ are those which are connected to $\mathbf{x}$ in $\mathcal{M}$. Also, $U(\mathbf{x})$ contains those elements which $\mathbf{x}$ is part of them in $\mathcal{M}$. Thus, Condition (3) in Definition 9 basically captures the monotonicity axiom in T_{mt} (Axiom 16) which says that every element that has a part which is connected to $\mathbf{x}$ is also connected to $\mathbf{x}$.

Theorem 10 gives a characterization of the models of T_{mt} up to isomorphism. Furthermore, since T_{mt} and T_{rcc8^*} are synonymous, this provides a characterization of the models of T_{rcc8^*}.

4.2 Characterization of $\mathfrak{M}^{mt}$

Although Definition 9 gives us a precise specification of the models of T_{mt}, it only provides an implicit characterization; we now outline an explicit characterization that gives us a complete understanding of the possible structures in $\mathfrak{M}^{mt}$, and so models of T_{mt}. The key to the characterization of $\mathfrak{M}^{mt}$ lies in understanding the graphs. In particular, we identify three distinct subgraphs in any structure in $\mathfrak{M}^{mt}$. The first graph is an instance of the following class:

▶ **Definition 11.** Suppose $\mathcal{Q} \in \mathfrak{M}^{par\_orders}$ and $\mathcal{Q} = \langle V, \preceq \rangle$. $\mathcal{U}_{\mathcal{Q}} = \langle V, E \rangle$ is the lower bound graph for $\mathcal{Q}$ iff : $(\mathbf{x}, \mathbf{y}) \in E$ iff exists $\mathbf{z} \in V$ s.t. $\mathbf{z} \preceq \mathbf{x}$, $\mathbf{z} \preceq \mathbf{y}$.

The lower bound graph for the partial ordering in Figure 3(ii) can be seen in Figure 3(iii). Note that the upper sets of elements form cliques in the graph. Lower bound graphs are well-understood within graph theory [14, 12, 3], with two different characterizations. The first is to consider them strictly from a graph-theoretic perspective: $\mathcal{G} = \langle V, E \rangle$ is a lower bound graph iff its vertex clique cover number is equal to its edge clique cover number, where the vertex clique cover number of $\mathcal{G}$ is the minimum number of cliques needed to cover V and the edge clique cover number of $\mathcal{G}$ is the minimum number of cliques needed to cover E.

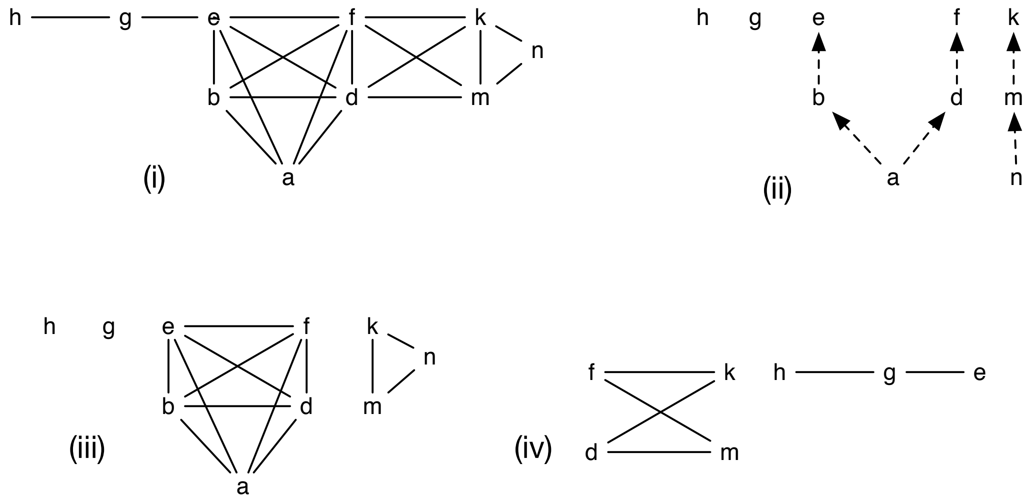


Figure 3 Examples structures and substructures of $\mathfrak{M}^{mt}$. The loops for each vertex in a graph are suppressed to enhance readability.

The second way to look at lower bound graphs is that they are isomorphic to the extension of the "overlaps" relation O. The other two subgraphs within a structure in $\mathfrak{M}^{mt}$ will be used to characterize the relationship between nonoverlapping (externally connected) elements.

We first need to define a few other classes of graphs before we get to the characterization theorem.

▶ **Definition 12.** Let $P = \langle V, \leq \rangle$ be a poset. The graph $G_{\mathbb{P}} = (V, E_P)$ is the comparability graph for P iff $(x, y) \in E_P$ whenever $x < y$ or $y < x$. $G = (V, E)$ is a comparability graph iff there is a poset P such that $G \cong G_P$.

▶ **Definition 13.** A graph G is a permutation graph iff G and $\overline{G}$ are comparability graphs.

This definition is actually the statement of a characterization theorem from [6]; the original definition of permutation graphs with respect to the representation of the elements of a permutation can be found in [16].

▶ **Definition 14.** Suppose $\mathcal{Q} \in \mathfrak{M}^{par\_orders}$. A graph $\mathcal{H}$ is an upper bipartite permutation graph for $\mathcal{Q}$ iff $\mathcal{H} = (V_1 \cup V_2, E)$ is a bipartite permutation graph such that V_1, V_2 are upper sets in $\mathcal{Q}$.

The first subgraph in Figure 3(iv) is an upper bipartite permutation graph for the partial ordering in Figure 3(ii), in which the upper sets are $V_1 = \{\mathbf{d}, \mathbf{f}\}$ and $V_2 = \{\mathbf{m}, \mathbf{k}\}$.

The third subgraph is not an instance of any special class of graphs, but rather can be an arbitrary graph, the only condition being that the vertices are all maximal elements in $\mathcal{Q}$.

▶ **Definition 15.** Suppose $\mathcal{Q} \in \mathfrak{M}^{par\_orders}$. A crown for $\mathcal{Q}$ is a graph $G = (V, E)$ such that all vertices in V are maximal elements of $\mathcal{Q}$ and which are not externally connected to proper parts of any other element.

The second subgraph in Figure 3(iv) is a crown, since its vertices consist entirely of elements $\{\mathbf{e}, \mathbf{g}, \mathbf{h}\}$ that are maximal in $\mathbb{Q}$.

The lower bound graph, the upper bipartite permutation graphs, and the crowns must be combined to form a graph that satisfies the conditions in Definition 9.

▶ **Definition 16.** A graph $\mathcal{G} = \langle V, E\rangle$ is edge-decomposable into a set of graphs H iff

1. $\mathcal{H}_i \subset \mathcal{G}$, for each $\mathcal{H}_i \in H$;
2. $E_i \cap E_j = \emptyset$, for each $\mathcal{H}_i = \langle V_i, E_i\rangle$ and $\mathcal{H}_j = \langle V_j, E_j\rangle$;
3. $E = \bigcup_i E_i$.

Thus, a graph $\mathcal{G}$ is edge-decomposable into a set of subgraphs iff the set of edges in $\mathcal{G}$ can be partitioned. We will use the notation $\mathcal{G} = \mathcal{H}_1 \uplus ... \uplus \mathcal{H}_n$ to indicate that $\mathcal{G}$ is edge-decomposable into $\mathcal{H}_1, ..., \mathcal{H}_n$.

▶ **Theorem 17.** *$\mathcal{M} \in \mathfrak{M}^{mt}$ iff $\mathcal{M} = \langle V, E, \preceq\rangle$ such that*

1. *$\mathcal{Q} = \langle V, \preceq\rangle$ and $\mathcal{Q} \in \mathfrak{M}^{par\_orders}$;*
2. *$\mathcal{G} = \langle V, E\rangle$ and $\mathcal{G} \in \mathfrak{M}^{graph\_loops}$;*
3. *$\mathcal{G} = \mathcal{U}_{\mathcal{Q}} \uplus \mathcal{G}_u \uplus \mathcal{G}_m$ such that*
   - **(a)** *$\mathcal{U}_{\mathcal{Q}}$ is the lower bound graph for $\mathcal{Q}$;*
   - **(b)** *$\mathcal{G}_u$ is decomposable into a set of upper bipartite permutation graphs for $\mathcal{Q}$;*
   - **(c)** *$\mathcal{G}_m$ is a crown for $\mathcal{Q}$.*

Suppose a structure $\mathcal{M} \in \mathfrak{M}^{mt}$ is composed of the graph $\mathcal{G}$ depicted in Figure 3(i) and the corresponding partial ordering $\mathcal{Q}$ depicted in Figure 3(ii). The graph $\mathcal{G}$ is edge-decomposable into $\mathcal{U}_{\mathcal{Q}}$, $\mathcal{G}_u$, and $\mathcal{G}_m$, where $\mathcal{U}_{\mathcal{Q}}$ is the lower bound graph depicted in Figure 3(iii), while $\mathcal{G}_u$ and $\mathcal{G}_m$ are depicted in Figure 3(iv). Suppose $\mathcal{N} \in Mod(T_{mt})$ is the model corresponding with $\mathcal{M}$. Intuitively speaking, $\mathcal{U}_{\mathcal{Q}}$ is the subgraph of $\mathcal{G}$ in which two vertices $\mathbf{x}, \mathbf{y}$ are neighbors whenever $\mathbf{x}$ and $\mathbf{y}$ overlap in $\mathcal{N}$. That is, $\mathcal{U}_{\mathcal{Q}}$ captures the connection relation between overlapping elements of $\mathcal{N}$. $\mathcal{G}_u \cup \mathcal{G}_m$ represents (externally) connected non-overlapping elements of $\mathcal{N}$; that is, $\mathbf{x}$ and $\mathbf{y}$ are neighbours in $\mathcal{G}_u$ whenever in $\mathcal{N}$ they are connected but do not overlap.

Theorem 17 is a characterization theorem for $Mod(T_{mt})$ because it tells us how to construct all possible models of T_{mt} up to isomorphism. We can take an arbitrary lower bound graph, together with a set of upper bipartite permutation graphs, and an arbitrary graph, and combine these graphs together to yield a model of T_{mt}. Given the synonymy of T_{mt} and T_{rcc8*}, this Theorem also characterizes all possible solutions of a set of RCC8 constraints; by synonymy, any solution is isomorphic to a mereology together with a graph that is decomposable into the three subgraphs specified in Theorem 17.

5 Summary

Constraint satisfaction with spatial calculi such as RCC8 has been the predominant application of mereotopology within commonsense reasoning. Yet in some way, this has diminished the role played by the different mereotopology ontologies that were the original sources. It has long been known that the first-order theory of RCC8 is interpretable by the mereotopology ontologies, not only RCC, but also including the rather weak ontology T_{mt}. This perspective has been considered sufficient for showing that RCC8 was in some sense sound with respect to its mereotopological foundations. On the other hand, it has been thought that the first-order theory of RCC8 was too weak to be considered to be a mereotopological ontology in its own right. In this paper, we have shown that indeed the RCC8 theory is mutually interpretable with T_{mt}. Furthermore, by extending the RCC8 theory with sentences that fully capture the intended interpretations of the RCC8 relations, we obtain a theory that is logically synonymous with T_{mt}. Finally, we have provided a characterization of the models of T_{mt} up to isomorphism, by first specifying a class of mathematical structures, and then showing that T_{mt} axiomatizes this class of structures. This characterization gives us insights into the set of all possible solutions for a set of RCC8 constraints. The characterization also lays the groundwork for a new approach to location ontologies, in which we embed the models of a mereotopology of physical objects in a mereotopology of abstract spatial regions.

References

1 N. Asher and L. Vieu. Toward a geometry of common sense: A semantics and a complete axiomatization of mereotopology. In *IJCAI*, pages 846–852, 1995.

2 B. Bennett, A. Isli, and A. G. Cohn. When does a composition table provide a complete and tractable proof procedure for a relational constraint language? In *Proceedings of the IJCAI-97 workshop on Spatial and Temporal Reasoning*, 1997.

3 D. J. Bergstrand and K. F. Jones. On upper bound graphs of partially ordered sets. *Congression Numerica*, 66:185–193, 1988.

4 R. Casati and A. C. Varzi. *Parts and places: The structures of spatial representation*. MIT Press, 1999.

5 I. Düntsch and M. Winter. A representation theorem for boolean contact algebras. *Theoretical Computer Science*, 347(3):498–512, 2005.

6 B. Dushnik and E. W. Miller. Partially ordered sets. *American Journal of Mathematics*, 63:600–610, 1941.

7 H. Enderton. *Mathematical Introduction to Logic*. Academic Press, 1972.

8 C. Eschenbach. Viewing composition tables as axiomatic systems. In *Proc. of the Conference on Formal Ontology in Information Systems*, pages 93–104. ACM, 2001.

9 N. M. Gotts. An axiomatic approach to topology for spatial information systems. *Research Report Series-University of Leeds School of Computer Studies*, 1996.

10 M. Gruninger, T. Hahmann, A. Hashemi, D. Ong, and A. Ozgovde. Modular first-order ontologies via repositories. *Applied Ontology*, 7(2):169–209, 2012.

11 W. Hodges. *Model theory*. Cambridge University Press Cambridge, 1993.

12 J. R. Lundgren and J. S. Maybee. A characterization of upper bound graphs. *Congression Numerica*, 40:189–193, 1983.

13 W. McCune. Prover9 and mace4. `http://www.cs.unm.edu/~mccune/prover9/`, 2010.

14 F. R. McMorris and T. Zaslavsky. Bound graphs of a partially ordered set. *Journal of Combinatorics and Information Systems Science*, 7:134–138, 1982.

15 Charles Pinter. Properties preserved under definitional equivalence and interpretations. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 24:481–488, 1978.

16 A. Pnueli, A. Lempel, and Even S. Transitive orientations of graphs and identification of permutation graphs. *Canadian Journal of Mathematics*, 23:160–175, 1971.

17 D. A. Randell, Z. Cui, and A. Cohn. A spatial logic based on regions and connection. In *Proceedings of the KR92*, pages 165–176. Morgan Kaufmann, 1992.

18 J. G. Stell. Boolean Connection Algebras: A New Approach to the Region-Connection Calculus. *Artificial Intelligence*, 122:111–136, 1999.

A Qualitative Spatial Descriptor of Group-Robot Interactions

Zoe Falomir[1] and Cecilio Angulo[2]

1 **University of Bremen, Bremen, Germany**
`zfalomir@uni-bremen.de`
2 **Universitat Politècnica Catalunya, Barcelona, Spain**
`cecilio.angulo@upc.edu`

Abstract

The problem of finding a suitable qualitative representation for robots to reason about activity spaces where they carry out tasks such as leading or interacting with a group of people is tackled in this paper. For that, a Qualitative Spatial model for Group Robot Interaction (QS-GRI) is proposed to define Kendon's F-formations [16] depending on: (i) the relative location of the robot with respect to other individuals involved in that interaction; (ii) the individuals' orientation; (iii) the shared peri-personal distance; and (iv) the role of the individuals (observer, main character or interactive). An iconic representation is provided and Kendon's formations are defined logically. The conceptual neighborhood of the evolution of Kendon formations is studied, that is, how one formation is transformed into another. These transformations can depend on the role that the robot have, and on the amount of people involved.

1998 ACM Subject Classification I.2 Artificial Intelligence, I.2.3 Deduction and Problem Solving, I.2.9 Robotics, I.6 Simulation and Modeling

Keywords and phrases qualitative modeling, spatial reasoning, location, distance, orientation, cognitive robotics, human-robot interaction, group-robot interaction, logics

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.3

1 Introduction

Robot tour guides appeared in the late 90s: Rhino [4] was located at the Deustche Museum in Bonn, Germany; Minerva [25] at the Smithsonian's National Museum of American History in Washington, etc. Nowadays flying quadcopters are used at MIT for personal guiding to labs (Skycall[1] project). Robots and other automats are getting gradually involved in human daily living activities, and in human environments, social robots must have the ability to communicate with people closely and fluidly both in a verbal and in a non-verbal way.

Spatial relationships are involved in human-robot interaction (HRI), e.g. combinations of distance, relative position and spatial arrangements that occur naturally when two or more people engage in an interaction [15, 20]. Empirical studies in robotics [17] identified spatial relations between people and a robot as a key issue to improve the quality of interaction noticing that interpersonal distances convey significant and relevant social information. Social interaction when navigating, specifically when robots pass people [22, 1] was also studied.

Qualitative descriptors for reasoning about moving objects appeared in the literature to represent HRIs in navigation situations where one robot and one human (or a group of

[1] `http://senseable.mit.edu/skycall/`

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 3; pp. 3:1–3:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

humans as a whole) are involved [11]. Qualitative spatial representations for activity spaces where a robot carry out a task or collaborate with more that one person are not available in the literature, as far as we are concerned. This paper refers to social interactions among humans and HRI in social environments, which may involve several individuals (sometimes arranged as a group) and one robot –from now on named as Group-Robot Interactions, GRI.

Few approaches in the literature have dealt with the challenge of formalizing social conventions for robots to behave more cognitively in human populated scenarios. The Qualitative Trajectory Calculus (QTC) was used to model HRI [8, 9, 2, 14]. QTC uses points as primitives to represent both the human and the robot, and their relative motion is expressed in a set of tuples of qualitative relationships. Qualitative social rules for robots to have a polite pedestrian behavior while navigating were proposed [10] using $OPRA_4$ calculus to formalize polite navigation rules (in situations as crossing, narrow passages, passing groups from the outside, etc.) and motion planning and pedestrian behavior were simulated using JWalkerS and SparQ toolbox[2] to investigate how traveling time is influenced by being polite. These pedestrian rules were also modeled in QLTL (Linear Temporal Logic with Qualitative Spatial Primitives) [11] and tested in a case study using a Kinect camera and a laser range scanner on a mobile robot. However, spatial arrangements of a robot interacting with a group of people (i.e. carrying a joint action) has not been studied yet.

The *Groups in Human-Robot Interaction* community discussed at IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN 2016)[3] that inter-group interaction differs from inter-individual (dyadic) interaction. Ideally a robot should have different models of behavior depending on the number of people around it [18]. Thus, the first step is identifying the interactive situation a robot is facing.

This paper is organised as follows. Section 2 presents Kendon's [16] F-formations for group behavior. As these F-formations are described in a linguistic manner, next sections formalize them using qualitative representations and first order logics. Final sections provide an experiment for testing the logics defined, a discussion, conclusions and future work.

2 F-formations by Kendon

The *F-formation* system proposed by Kendon [16] studied spatial structures, both in position and orientation, generated when two or more people interact, and affirmed that *"behaviour of any sort occurs in a three dimensional world and any activity whatever requires space of some sort "* [ibid, p. 1.] This space allows an individual to perform any activity and it is differentiated from other spaces [20]. According to Kendon, in any scenario it is common that several individuals are co-present, but the way they are positioned and oriented in relation to the others reflects directly how they can be involved together. Based on his observations, Kendon defines a transactional space, *o-space*, as the space where people can interact and manipulate shared objects. In dyadic interactions, Kendon observed two types of formations: *vis-a-vis* (individuals are facing to each other) and *L-shape* (individuals are standing perpendicularly to each other facing an object). When the interaction occurs between two or more people, Kendon observed three types of formations: *circular form* (all people are looking at each other), *side-by-side* (people stand closely together and facing the same segment of the environment), and *horseshoe shape* (a kind of compromise between *side-by-side* and *circular form*). Typical spatial arrangements also happen in occasions where

[2] SparQ toolbox: http://www.sfbtr8.uni-bremen.de/project/r3/sparq/
[3] https://grouprobot.wordpress.com/home/

there is an unequal distribution of rights to start a conversation or action, for example, in the *performer-audience* interaction. In contrast, if a group of people does not follow any spatial arrangement between them is known as *cluster*.

3 A Qualitative Spatial Descriptor of Group-Robot Interactions

This section presents a Qualitative Spatial descriptor for representing Group-Robot Interactions (QS-GRI). First, the representation for an individual is provided: an iconic representation is given, the location, orientation and distance reference systems used are defined and the first order logic statements generated are described (Section 3.1). Then the relations which can be obtained by QS-GRI between two individuals are described (Section 3.2).

3.1 QS-GRI Iconic Representation for an Individual

QS-GRI defines interactions between robots and people depending on: location, orientation and distance. Robots must be aware that people's personal space usually is not interfered by other people unless they are family, and this space is not allowed to be interfered by robots. So, an interactive distance for a robot is that distance which is not too close to any person but not too far away for them. Kendon [16] defined the *o-space* as the space where people can interact and manipulate shared objects. Similarly, in psychology, peri-personal space is defined as the space wherein individuals manipulate objects, whereas extra-personal space –which extends beyond the peri-personal space– is defined as the portion of space relevant for locomotion and orienting [12, 6]. Therefore, let us determine that two individuals that share their peri-personal space can be considered to have an interaction.

Moreover, any person distinguishes spatial orientations inside his/her personal and peri-personal space. These areas are usually named as: *front, back, right* and *left*. A person is also an oriented entity in space, defined by his/her *front*, indicated by their eyes.

The iconic representation of an individual (robot or person) used in QS-GRI is shown in Figure 1. That is, any individual fills an area in space (in blue), and (s)he has a personal space (in red) which is private, and a peri-personal space (in green) which is that space that (s)he can reach using their body or a tool. The white space is the extra-personal space.

These locations are defined using a Location interval Reference System, that is, Lo_{RS}= $\{\alpha$, Lo_n, $\mathrm{Lo}_{int(\alpha)}\}$ where α is the angular amplitude starting from 0 –located following the unit circle convention in trigonometry, that is, on the right-hand of an individual– to a range of $[0, 2\pi]$ measured in radians; Lo_n refers to the set of names given as locations; and $\mathrm{Lo}_{int(\alpha)}$ refers to a function which returns the corresponding Lo_n depending on α. In general:

$$\mathrm{Lo}_n = \{\mathrm{Lo}_1, \mathrm{Lo}_2, \dots, \mathrm{Lo}_K\}\,,$$
$$\mathrm{Lo}_{int(\alpha)} = \{[\mathrm{lo}_0, \mathrm{lo}_1], (\mathrm{lo}_1, \mathrm{lo}_2], \dots, (\mathrm{lo}_{K-1}, \mathrm{lo}_K]\}\,,$$

where K is the number of concepts used for defining orientations. The Lo_n and $\mathrm{Lo}_{int(\alpha)}$ can be defined for the QS-GRI adapting to the case of study. Therefore, for modeling Kendon F-formations, the following Lo_{RS} can be selected:

$$\mathrm{Lo}_n = \{\mathit{right}, \mathit{front}, \mathit{left}, \mathit{back}\}\,,$$
$$\mathrm{Lo}_{int(\alpha)} = \{(-\pi/4, \pi/4], (\pi/4, 3/4\pi], (-\pi/4, -3/4\pi](3/4\pi, -3/4\pi]\}\,.$$

An individual can rotate its *front* towards any direction in the space. Thus, the orientation of an individual is also taken into account by QS-GRI, which is calculated with respect to its *front* defined in the Lo_{RS}, and can be determined by the following RS: $\text{O}_{RS}= \{\sigma,\ \text{O}_i,\ \text{O}_{g(\sigma)}\}$ where σ is the angle of rotation measured from the *front* with a range of $[0, 2\pi]$ in radians; O_i refers to the set of names (n) given to the orientations; and $\text{O}_{g(\sigma)}$ refers to the function which relates the σ with a given name. In general:

$$\begin{aligned} \text{O}_i &= \{\text{O}_1, \text{O}_2, \ldots, \text{O}_M\}\,, \\ \text{O}_{g(\sigma)} &= \{\text{o}_1(\sigma), \text{o}_2(\sigma), \ldots, \text{o}_M(\sigma)\}\,, \end{aligned}$$

where M is the number of concepts used for defining orientations. The O_i and O_g can be adapted to the case of study. Therefore, for QS-GRI, the following O_{RS} is selected:

$$\begin{aligned} \text{O}_i = \{&\textit{towards-front}(tf), \textit{towards-front-right}(tfr), \textit{towards-right}(tr), \\ &\textit{towards-back-right}(tbr), \textit{towards-back}(tb), \textit{towards-back-left}(tbl), \\ &\textit{towards-left}(tl), \textit{towards-front-left}(tfl)\}\,, \\ \text{O}_{g(\sigma)} = \{&0, (0, \pi/2), \pi/2, (\pi/2, \pi), \pi, (\pi, 3/2\pi), 3/2\pi, (3/2\pi, 2\pi)\}\,. \end{aligned}$$

In order to define the spaces surrounding an individual, QS-GRI uses a Distance Reference System or $\text{D}_{RS} = \{d,\ \text{D}_n,\ \text{D}_f\}$, where d refers to a distance measured in meters (m), D_n refers to the set of names corresponding to the spaces defined; and D_f refers to the values of d related to each label. In general,

$$\begin{aligned} \text{D}_n &= \{\text{d}_1, \text{d}_2, \ldots, \text{d}_Q\}\,, \\ \text{D}_f &= \{[0, \text{d}_1], (\text{d}_1, \text{d}_2], \ldots, (\text{d}_{Q-1}, \text{d}_Q]\}\,, \end{aligned}$$

where Q is the quantity of concepts defined. Both D_n and D_f can be parameterized depending on the case of study. For QS-GRI:

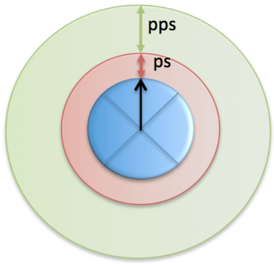
$$\begin{aligned} \text{D}_n &= \{\textit{ps}, \textit{pp}, \textit{eps}\}\,, \\ \text{D}_f &= \{[0, 0.46], (0.46, 0.46 + ToolLength], (0.46 + ToolLength, \infty)\}\,, \end{aligned}$$

where *ps* is the personal space, *pps* is the peripersonal space, and *eps* is the extra-personal space. The width of the *ps* depends on the person, their social abilities and culture. Some people would need a wider personal space than other people. The *pps* is dynamic and adaptable, depending on the tool used by the person/robot and their abilities (i.e. flexibility of legs/arms for a person, actuator possibilities in a robot, etc). Thus, these areas can be customized for an individual but also parameterized based on psychological experimental studies [3]. For example, Hall [13] defined 4 kinds of interpersonal distances, each with its own significance in a social context: intimate ($0 - 0.46$ m), personal ($0.46 - 1.22$ m), social ($1.22 - 3.66$ m) and public (> 3.66 m). In QS-GRI, the *ps* may correspond to Hall's intimate distance, and the peripersonal space may involve the personal and social distance.

The QS-GRI can represent any individual using Horn clause logic [19] and Prolog programming language [23]. A possible description for an individual is given in Figure 1.

3.2 Relations between Individuals Inferred by QS-GRI

According to the previous definitions given for QS-GRI, relations of location, topology and distance can be inferred with respect to (wrt.) each individual. In this section, the logical rules for these inferences are provided.




```
has_location_xy(ind, 10,10).
has_orientation(ind, pi/2, towards-front).
has_width(ind,1).
has_ps(ind, 0.46).
has_tool_length(ind, stick, 0.2).
has_area(ind, right, -pi/4, pi/4).
has_area(ind, front, pi/4, 3/4*pi).
has_area(ind, left, 3/4*pi, 5/4*pi).
has_area(ind, back, 5/4*pi, 7/4*pi).
```

Figure 1 Iconic and logic representation of an individual.

Topological relation A wrt. B. An individual B, is inside the peripersonal space of another individual A, if the distance between the location of B and the location of A is smaller than their peri-personal limits. Moreover, if A is in the peri-personal space of B, B is also in the peri-personal space of A.

```
in_pps(A,B):-
  has_location_xy(A,X,Y), has_location_xy(B,X2,Y2),
  has_pps(A,LimitA), has_pps(B,LimitB),
  distance(X,Y,X2,Y2,D), D < LimitA+LimitB.
```

```
in_pps(A,B):-
  in_pps(B,A).
```

Relative Location of A wrt. B. The area around any individual is divided in locations according to the Lo_{RS}. So, the location of an individual A wrt. another individual B, is computed. For that, the rL_{RS} is built: rL_{RS}= $\{\alpha, \mathrm{rLo}_j, \mathrm{rL}_{f(\alpha)}\}$ where α is the angle of location of A wrt. B in radians; rLo_j refers to the set of names (n) defined as locations in Lo_n and its combinations; and $\mathrm{Lo}_{f(\alpha)}$ takes the values in radians as parameters in a belonging function ($h(\sigma)$) which returns the corresponding location in rLo_j and a value of certainty ($Grade$). This $Grade$ is needed to evaluate how to the $front$, for example, is an individual located. It depends on the relative angle between the individuals as indicated below.

$$\mathrm{rLo}_j = \mathrm{Lo}_n \cup \{\textit{front-right}, \textit{front-left}, \textit{back-right}, \textit{back-left}\}\,,$$
$$\mathrm{rLo}_{h(\sigma,grade)} = \mathrm{rLo}_h(\mathrm{Lo}_{int(\sigma)}, \mathrm{grade})\,.$$

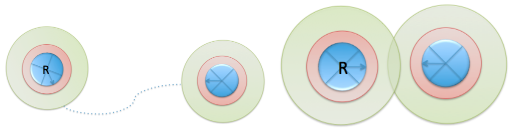
```
located(Lon,A,B,Grade):-
  relative_coordinates_to_A(A,B,Xr,Yr),
  location(Lon,Xr,Yr,Grade).

relative_coordinates_to_A(A,B,Xr2,Yr2):-
  has_location_xy(A,X,Y),
  has_location_xy(B,X2,Y2),
  has_orientation(A,RAngle, _),
  Xr is X2-X, Yr is Y2-Y,
  Xr2 is round((Xr*cos(RAngle))-(Yr*sin(RAngle))),
  Yr2 is round((Xr*sin(RAngle))+(Yr*cos(RAngle))).
```

```
location(front,0,Yr,Grade):-
  Yr >= 0, Grade is 1.
location(front,Xr,Yr,GradeS2):-
  Xr <> 0,
  Yr > 0,
  GradeS is sin(Yr/Xr),
  GradeC is cos(Yr/Xr),
  GradeC2 is abs(GradeC),
  GradeS2 is abs(GradeS),
  GradeS2 > GradeC2.
```

Similarly, the rest of the locations (*right, left, back*) of an individual B wrt. another individual A are obtained. And the combined location relations (*front-right, front-left, back-right, back-left*) are inferred. Note that, as individuals have a *ps* area, then the points on the boundary of this *ps* must be used to obtain correct locations.

```
located(front-right,A,B,Grade):-
  boundary_point_loc(front,A,B,GradeF),
  boundary_point_loc(right,A,B,GradeR),
  Grade is GradeF * GradeR.
```

```
located(back-left,A,B,Grade):-
  boundary_point_loc(back,A,B,GradeB),
  boundary_point_loc(left,A,B,GradeL),
  Grade is GradeB * GradeL.
```




```
vis_a_vis(A, B):-
  facing_each_other(A,B, _),
  in_pps(A, B).

facing_each_other(A,B, GradeResult):-
  located(front,A,B, GradeA),
  located(front,B,A, GradeB),
  GradeResult is (GradeA+GradeB)/2.
```

Figure 2 Vis-a-vis formation. Note that A and B are variables which can refer to any individual.

Location-neighbourhood relations that can help us to define the F-formations (i.e. *next to*, *in middle*, *neighbour*) and are inferred using QS-GRI as follows. Note that other relations (i.e. *behind*, *in front*) are also possible to define.

```
next_to(A,B):-
  in_pps(A,B),
  located(right,A,B,GradeR),
  located(left,B,A,GradeL).
in_middle(A,B,C):-
  next_to(A,C), next_to(C,B).
```

```
neighbour(A,B):-
  in_pps(A,B),
  located(front-right,A,B,GradeR),
  located(front-left,B,A,GradeL).
```

Orientation relation A wrt. B. By expressing the orientation of an object A wrt. another object B, relations of opposition (*towards-right* vs. *towards-left*, *towards-front* vs. *towards-back*, *towards-front-left* vs. *towards-back-right*, and *towards-front-right* vs. *towards-back-left*) and relations of perpendicularity (*towards-right* vs. *towards-front*, *towards-left* vs. *towards-down*, *towards-left* vs. *towards-front*, and *towards-right* vs. *towards-down*) can be extracted, which are useful to identify individual group formations. Logically, these relations can be written as the following examples:

```
opposed_orientation(A,B):-
  has_orientation(A,_, towards-right),
  has_orientation(B,_, towards-left).
```

```
perpendicular_orientation(A,B):-
  has_orientation(A,_, towards-down),
  has_orientation(B,_, towards-right).
```

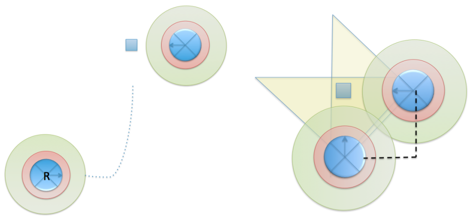
4 Recognizing Social Formations in Groups of Individuals

In this Section the F-formations defined by Kendon [16] are described logically using the predicates defined by QS-GRI: vis-a-vis, L-shape, circular, horse-shoe, side-by-side, performer-audience or cluster formation.

Vis-a-vis Formation: Individuals are facing each other and their *pps* intersect in the *front* area of both individuals, as Figure 2 shows. Note that the *front* of each individual must be oriented to each other relative *front* and that their orientations are opposite.

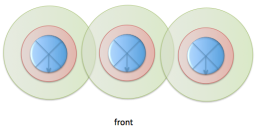
L-shape Formation: Two individuals are facing an object (Figure 3). This object is located in the front area of both individuals. The object observed is not animated, so it has no personal and no peri-personal space. These two individuals must share some peri-personal space. The intersection of this peri-personal space intersects at their *front-left* area of one individual and at the *front-right* area of the other individual.

The individuals are observers, they are not carrying out any physical activity together, otherwise they would face each other (e.g., they may be talking about the object). The roles of speaker and listener can be taken in turns. Note that the orientation of each individual is perpendicular to each other.




```
l-shape(A, B):-
  in_pps(A, B),
  located(front,A,object,_),
  located(front,B,object,_),
  dif(A,B).
```

Figure 3 L-shape formation.




```
side_by_side(A,B):-
  next_to(A,B),
  has_orientation(A,_,O),
  has_orientation(B,_,O).
```

```
side_by_side_group(A,[]).
side_by_side_group(A,[B|L]):-
  side_by_side(A,B),
  side_by_side_group(B,L).
```

Figure 4 Side-by-side formation.

Side-by-side formation: Individuals have the same orientation. They share their peri-personal space with the individuals *next to* them on their *left* and on their *right*.

In the queuing variation, individuals have also the same orientation, but they share their peri-personal space with their neighbors at their *front* and at their *back*. In both cases, individuals' role is passive. They are listeners-observers. Usually, they do not take the speaker role unless they are given permission for (i.e. for the queuing variation, until they reach the head of the queue). Note that, in both *side-by-side* and *queuing* formations, individuals only must change their orientation to establish a *facing each other* relation.

Horse-shoe formation: Individuals share their peri-personal space with their neighbors, in the *right* and *left* area. They all share their *front* area. All the individuals are observers: they are displaced to listen to somebody or to see some object (Figure 6).

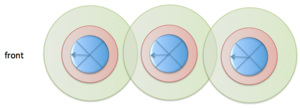
Hence, they hold the role of listeners. This is a passive role which can be changed with permission of the speaker, which is usually located at the shared *front*. Note that, the first and last individuals in the group-chain are *facing each other*.

Circular formation: Individuals are displaced in a triangular spatial formation sharing a common peri-personal space (Figure 8) on their right, and on their left. They are oriented towards a shared front.

In the general circular formation, each member of the group shares her personal space on her $right$ and also on her $left$, so each of the members in the group have two neighbors at the mentioned locations.

The individuals in the group are not only observers, they can interact with each other. The roles of speaker and listener can be exchanged constantly. Therefore, in order to maintain a circular shape, each member of the group has at least one other member located at its $front$ or facing each other in the distance (not sharing peri-personal space).

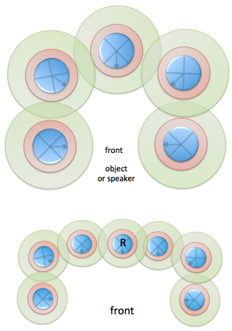
Performer-audience or cluster formation: All the individuals have the same perspective and they share their *pps* with their neighbors at their *front*, *right*, *left* and at their *back*, that is, they are *next to* someone and also *cueing* with someone (Figure 9). Their role is passive since they are listeners-observers. They do not take the speaker roll unless they are given permission.




```
cueing(A,B):-
  in_pps(A,B),
  located(back,A,B,GradeF),
  located(front,B,A,GradeB).
```

```
cueing_group(A,[]).
cueing_group(A,[B|L]):-
  cueing(A,B),
  cueing_group(B,L).
```

Figure 5 Cueing formation.




```
horse_shoe_form([Head|L]):-
  length(L,NumInd),
  NumInd >= 3,
  last(L,Tail),
  facing_each_other(Head,Tail,_),
  horse_shoe(Head,L).

horse_shoe(_,[Tail]):-
  located(front,Tail,obj2,_).
```

```
horse_shoe(H,[X|L]):-
  located(front,H,obj2,_),
  (next_to(X,H);
  neighbour(X,H)),
  horse_shoe(X,L).
```

Figure 6 Horse-shoe formation: individuals observe someone/ something while sharing its left/right peripersonal space and its front.

5 Dynamics of Social Formations: Exploring QS-GRI Neighbourhoods

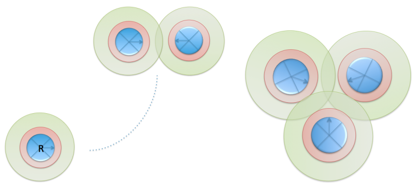
This section deals with the following challenge: *where the robot should locate itself to be included in a group? and towards which direction should it be oriented?* The first step towards the solution is to identify which kind of F-formation is the group taking. Then, for some F-formations, the role of the robot is relevant because it determines the location where the robot should place itself. For example, in the *horse-shoe* formation, most of the individuals have an *observer* role, while the individual at the front has a *leading* role. If the situation evolves so that the one leading allow others to lead and their roles are exchanged, then an interactive situation is happening and the *horse-shoe* formation evolves to a *circular* formation. For this reason, how a F-formation can evolve by including individuals is studied depending on the roles involved: leading, observer or interactive.

If the robot has an interactive goal, and detects:

- a person, it can select the vis-a-vis formation to locate itself and start this interaction. For that, it must be located in front of the person, oriented towards the person, and it must share that person's *pps* but their *ps* must not intersect (Figure 2).
- two people in a vis-a-vis formation, then the robot can select the triangular formation to locate itself to try to start an interaction.
- a group of more than 3 people who interact among themselves, then the robot can select a circular formation. The evolving formations are those where the circle is getting bigger: 4-circular formation, 5-circular formation, n-circular-formation (Figure 8).

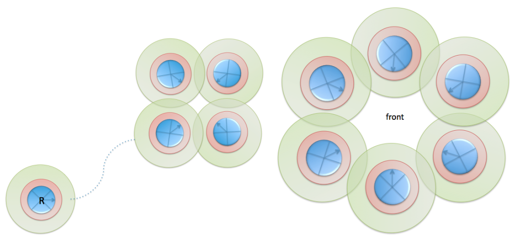
Let us explain how the rest of F-formations are useful for the robot to place itself depending on its goal, which may be:

- interacting with one person while observing an object, then the robot selects the L-shape formation to start this interaction (Figure 3).
- leading, i.e. performing a speech to a group of people located in a horse-shoe formation (Figure 6). The robot must locate itself at the front. While if the robot takes an observer




```
triangular_form(A,B,C):-
  both_sides_neighbours(A,B,C),
  both_sides_neighbours(B,A,C),
  both_sides_neighbours(C,A,B).

both_sides_neighbours(A,B,C):-
  in_pps(A,B),
  in_pps(A,C),
  (located(front-left,A,B,_);
   located(front-right,A,B,_)),
  (located(front-right,A,C,_);
  located(front-left,A,C,_)).
```

Figure 7 Minimal circular or triangular formation.




```
circular(Group):-
  length(Group,NumInd),
  NumInd > 3,
  some_members_loc(front,Group,Group),
  two_neighbours_for_each(Group,Group).

two_neighbours_for_each([Head|L],Group):-
  last(Group,Tail),
  nextto(Head,Next,Group),
  ( neighbour(Head,Tail);
    next_to(Head,Tail)),
  ( neighbour(Head,Next);
    next_to(Head,Next)),
  two_neighbours_middle(L,Group).
```

Figure 8 General circular formation. The complete definition is available[5].

role, then the robot chooses to locate itself among the people. The robot shares its left and right peri-personal space with its neighbors.

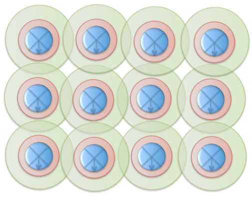
- leading, i.e. performing some speech to a group of people who are located in a side-by-side formation or in a cluster formation (i.e. performance), then the robot chooses to locate itself at the front, not in the crowd.
- observing, i.e. observing a performance with a group of people. These people are located in a side-by-side formation, and the robot incorporates itself in this side-by-side or cluster formation (Figure 4). In the cluster formation, the robot can have more than 2 left-right-neighbours and up to 4. In the situation depicted, the robot must also share its front *pps* with the person in front of it while they are sitting.

This relations among the F-formations have been summarized in Table 1. Note that a change of the robot activity/role involves a change in its location in the corresponding formation (see lines in Table 1), while adding a new person in the group also makes the formation to evolve to a different one (change in columns in Table 1).

6 Experimentation

In order to test the QS-GRI, we selected Prolog programming language [23], which is based on Horn clause logic [19]. Swi-Prolog[4] was the testing platform. Figure 10 presents the experimental world used to test QS-GRI logic algorithms[5] in an envisioned museum scenario

[4] SWI-Prolog: `http://www.swi-prolog.org/`
[5] Download from CogQDA project website: `https://sites.google.com/site/cogqda/publications`




```
cluster(L):-
  all_in_cluster(L,L).

all_in_cluster([],_).
all_in_cluster([X|L],Cluster)
  in_cluster(X,Cluster).
```

```
in_cluster(A,L):-
  side_by_side_with_sb(A,L,LNext),
  cueing_with_sb(A,L,LCue),
  length(LNext,NumNext),
  length(LCue,NumInCue),
  NumNext > 0, NumInCue > 0.
```

Figure 9 Performer-audience formation or cluster formation[5].

Table 1 Table of conceptual neighborhood situations.

	Leading	**Observer**	**Interactive**
1 person	vis-a-vis	L-shape	vis-a-vis
2 people	at front in: side-by-side or minimal circular	L-shape	minimal circular
3 people	at front in: side-by-side or horse-shoe	observer in: side-by-side	circular
4 people	at front in: side-by-side or horse-shoe	observer in: side-by-side horse-shoe	circular
5 people	at front in: side-by-side or horse-shoe	observer in: side-by-side horse-shoe	circular
N people	at front in: side-by-side, horse-shoe or performance	observer in: side-by-side horse-shoe or performance	circular

where the surveillance camera helps the robot to take a general perspective to identify human formations and to identify where should it stand to start the interaction.

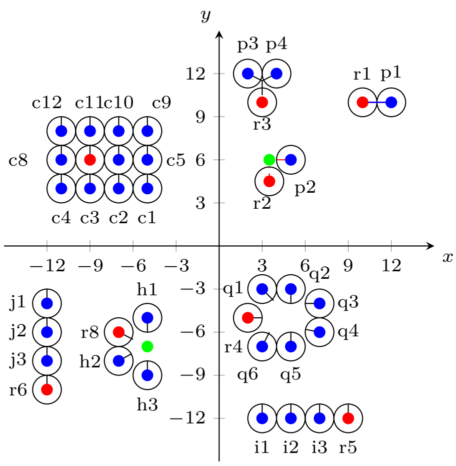
The following simulated environment has been implemented as facts in a close world. The elements showed are:

```
?- facing_each_other(R, P, G).
  R = r1,
  P = p1,
  G = 1 .

?- vis_a_vis(R,Ind).
  R = r1,
  Ind = p1 .

?- l-shape(A,B).
  A = r2,
  B = p2 ;
  A = p2,
  B = r2 ;
```

```
?- side_by_side_group(A,L).
L = [] ;
A = i1, L = [i2] ;
A = i1, L = [i2, i3] ;
A = i1, L = [i2, i3, r5] ;

?- cueing_group(A,L).
L = [] ;
A = j1, L = [j2] ;
A = j1, L = [j2, j3] ;
A = j1, L = [j2, j3, r6] ;
```

```
?- triangular_form(p3,p4,r3).
  true .

?- circular([r4,q1,q2,q3,q4,q5,q6]).
  true .

?- horse_shoe_form([h1,r8,h2,h3]).
  true .

?- cluster([c1,c2,c3,c4,c5,c6,
   r7,c8,c9,c10,c11,c12]).
  true .
```

- robot *r1*, located on the coordinates (10,10) in the simulated world and which is oriented *towards-right* and a person *p1*, who is located in the coordinates (12,10) and who is facing *towards-left*. According to these facts and the Lo_{RS}, it is inferred that *p1* is *in front of r1* and viceversa, and therefore, they are located in a *vis-a-vis* F-formation.
- robot *r2*, located on the coordinates (3.5,4.5) and which is oriented *towards-front* and a person *p2*, who is located in the coordinates (5,6) and who is facing towards-left. There is also a non-oriented *object* located on (3.5, 6). According to these facts and the Lo_{RS}, it is inferred that the *object* is *in front of r2* and *in front of p2*, sharing some peri-personal space, thus it is inferred that they are in a *L-shape* formation.
- individual *i1*, which is oriented *towards-front* has another individual *i2 next to*, which has another individual, *i3* also *next to*, which also has *r5 next to*, thus it is inferred that they are in a *side-by-side* formation.




```
has_loc_xy(r1, 10,10).                has_orient(r1,pi/2,tr).
has_loc_xy(p1, 12,10).                has_orient(p1,-pi/2,tl).
has_loc_xy(object, 3.5,6).            has_orient(p2,-pi/2,tl).
has_loc_xy(object2,-5,-7).            has_orient(r2,0,tf).
has_loc_xy(r2, 3.5,4.5).              has_orient(object,0,none).
has_loc_xy(p2, 5,6).                  has_orient(object2,0,none).
has_loc_xy(p3, 2,12).                 has_orient(p3,3/4*pi,tbr).
has_loc_xy(r3,3,10).                  has_orient(p4,5/4*pi,tbl).
has_loc_xy(p4,4,12).                  has_orientation(r3, 0, tf).
has_loc_xy(q1,3,-3).                  has_orient(q1,3/4*pi,tbr).
has_loc_xy(q2,5,-3).                  has_orient(q2,pi,tb).
has_loc_xy(q3,7,-4).                  has_orient(q3,3/2*pi,tl).
has_loc_xy(q4,7,-6).                  has_orient(q4,7/4*pi,tfr).
has_loc_xy(q5,5,-7).                  has_orient(q5,0,tf).
has_loc_xy(q6,3,-7).                  has_orient(q6,pi/4,tfr).
has_loc_xy(i1,3,-12).                 has_orient(r5,0,tf).
has_loc_xy(i2,5,-12).                 has_orient(i1,0,tf).
has_loc_xy(i3,7,-12).                 has_orient(i2,0,tf).
has_loc_xy(r5,9,-12).                 has_orient(i3,0,tf).
has_loc_xy(j1,-12,-4).                has_orient(j1,0,tf).
has_loc_xy(j2,-12,-6).                has_orient(j2,0,tf).
has_loc_xy(j3,-12,-8).                has_orient(j3,0,tf).
has_loc_xy(h1,-5,-5).                 has_orient(h1, pi, tb).
has_loc_xy(r8,-7,-6).                 has_orient(r8, 7/12*pi, tbr).
has_loc_xy(h2,-7,-8).                 has_orient(h2, 5/12*pi, tfr).
has_loc_xy(h3,-5,-9). (...)           has_orient(h3, 0, tf). (...)
```

Figure 10 Virtual world created for testing QS-GRI logic algorithms in Prolog. These predicates and orientations (see O_{RS}) have been abbreviated for saving space.

- robot *r3* is sharing its peri-personal space and its *front-right* and *front-left* areas with two individuals *p3* and *p4*, which also have the same relation between them. Thus, they are located in a *triangular* formation.
- robot *r4* and individuals *q1-q6* are located in a *circular* formation, whereas robot *r8* and individuals *h1-h3* is located forming a *horse shoe*.
- finally, robot *r7* and individuals *c1-c12* are located in a cluster formation.

7 Discussion

In robotics, research works usually analyze spatial interactions from a quantitative point of view, expressing spatial relationships in terms of numerical distances and absolute orientations. Since distances and directions are constantly changing, the representation of the interaction based on these primitives is complex. A qualitative descriptor such as QS-GRI can abstract the necessary information, while dealing with incomplete or uncertain data to define HRI in a more cognitive way.

In the literature, $\mathcal{EPRA}_m$ calculus [21] integrated cardinal absolute direction information and local distances. Other works focused on HRI [22, 10] divided the robot space following proxemics using: intimate, personal, social and public. This paper proposes a more psychological point of view by dividing space in personal and peri-personal, which is more related to Kendon definition of o-space [16], where people can interact and manipulate shared objects.

Exploratory studies in robotics [7] for evaluating HRI in terms of spatial relationships observed that it is possible to distinguish different types of spatial arrangements and group sizes, and to chose a discretization of group individuals to points/regions in space (see Figure 11).

Other studies in psychology and linguistics [24] observed that, in a communicative process, the capabilities assumed for the addresse depend if they are a human or a robot since speakers usually conceptualized a robot as *"a communication partner who needs comparably simple instructions"* (p.22), e.g. humans usually took the robot's perspective when giving instructions

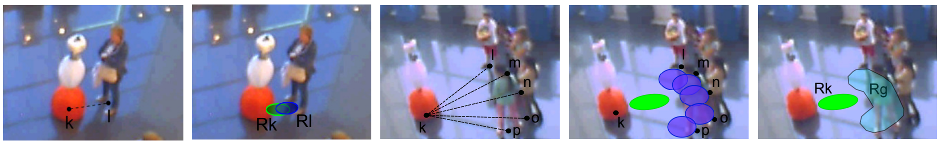


Figure 11 Real scenario: a *vis-a-vis* formation representing individuals as points/regions, a *horse-shoe* formation representing individuals as points/regions.

to it. The capacity of adaptation in humans in interactive situations facilitates HRI, which does not need to be so sophisticated as interaction among humans. However, the more the robot can reproduce human-similar utterances and behaviors, the more natural the interaction will get.

As far as we are concerned, there are not previous works in the literature that define Kendon's F-formations logically using qualitative descriptors and study their change/evolution as conceptual neighborhood. This evolution of F-formations may help robots to locate themselves following a social convention depending on the role they are assigned (main character/guide, observer/listener, or interactive). Further tests are intended for the QS-GRI logics I are aimed to be tested in a real scenario, where the robot perspective will substitute the general surveillance camera perspective with real human test subjects as future work.

8 Conclusions and Future Work

This paper presents a Qualitative Spatial model for Group Robot Interaction (QS-GRI) based on a location, orientation and a distance descriptor for representing individuals interacting in space. These descriptors are defined as first order logic statements and are used to infer relations of location, orientation and topology between individuals.

The QS-GRI identifies also Kendon's F-formations depending on: (i) the relative location of the robot with respect to other individuals involved in the interaction; (ii) the orientation of the individuals (shared front) or not; (iii) the shared peri-personal distance; and (iv) the role of the individuals (observer, main character or interactive). The recognition of these situations has been tested in a simulated world using Swi-Prolog.

Moreover, the evolution of Kendon-formations between them has also been studied to extract conceptual neighbourhood relations. That is, how one formation is transformed into another. These transformations depend on the robot role (i.e. interactive or observer), and on the number of people in the group.

As future work we intend to validate QS-GRI using the data available from the exploratory study carried out in a cultural centre where a robot guide is interacting to people [7]. QS-GRI is also envisioned to be applied in other human-robot collaboration (HRC) scenarios [5].

Acknowledgements. The support by the project *Cognitive Qualitative Descriptions and Applications* (CogQDA) funded by the Universität Bremen is gratefully acknowledged. We also thank the reviewers for their detailed comments and suggestions for this manuscript.

References

1 P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H.I. Christensen. Navigation for human-robot interaction tasks. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE Int. Conf. on*, volume 2, pages 1894–1900 Vol.2, April 2004. doi:10.1109/ROBOT.2004.1308100.

2 N. Bellotto, M. Hanheide, and N. Van de Weghe. Qualitative design and implementation of human-robot spatial interactions. In G. Herrmann et al., editor, *Social Robotics*, volume 8239 of *LNCS*, pages 331–340. Springer International Publishing, 2013.

3 R. J. Bufacchi, M. Liang, L. D. Griffin, and G. D. Iannetti. A geometric model of defensive peripersonal space. *Journal of Neurophysiology*, 115:218–225, 2016. `doi:10.1152/jn.00691.2015`.

4 W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114(1):3–55, 1999. `doi:10.1016/S0004-3702(99)00070-3`.

5 P. Bustos, L. J. Manso, J. P. Bandera Rubio, A. Romero-Garcés, L. V. Calderita, R. Marfil, and A. Bandera. A unified internal representation of the outer world for social robotics. In L. P. Reis et al., editor, *Robot 2015: Second Iberian Robotics Conference – Advances in Robotics, Lisbon, Portugal*, volume 418 of *Advances in Intelligent Systems and Computing*, pages 733–744. Springer, 2015. `doi:10.1007/978-3-319-27149-1_57`.

6 A. Couyoumdjian, F. Di Nocera, and F. Ferlazzo. Functional representation of 3D space in endogenous attention shifts. *The Quarterly Journal of Experimental Psychology Section A*, 56(1):155–183, 2003. `doi:10.1080/02724980244000215`.

7 M. Díaz-Boladeras, D. Paillacho, C. Angulo, O. Torres, J. Gonzalez, and J. Albo-Canals. Evaluating group-robot interaction in crowded public spaces: A week-long exploratory study in the wild with a humanoid robot guiding visitors through a science museum. *Int. Journal of Humanoid Robotics*, 12(04):1550022, 2015. `doi:10.1142/S021984361550022X`.

8 C. Dondrup, N. Bellotto, and M. Hanheide. A probabilistic model of human-robot spatial interaction using a qualitative trajectory calculus. In *AAAI Spring Symposium*, 2014.

9 C. Dondrup, N. Bellotto, and M. Hanheide. Social distance augmented qualitative trajectory calculus for human-robot spatial interaction. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pages 519–524, Aug 2014. `doi:10.1109/ROMAN.2014.6926305`.

10 F. Dylla, M. Coors, and M. Bhatt. Socially compliant navigation in crowded environments. In *Spatial and Temporal Dynamics (STeDy) Workshop at ECAI 2012*, pages 9–17, 2012.

11 F. Dylla, A. Kreutzmann, and D. Wolter. A qualitative representation of social conventions for application in robotics. In *2014 AAAI Spring Symposium – Qualitative Representations for Robots*, number SS-14-06 in Papers from the 2014 AAAI Spring Symposium, 2014.

12 C. Goerick. Towards cognitive robotics. In B. et al. Sendhoff, editor, *Creating Brain-Like Intelligence: From Basic Principles to Complex Intelligent Systems*, pages 192–214. Springer, Berlin, Heidelberg, 2009. `doi:10.1007/978-3-642-00616-6_10`.

13 E. Hall. *The hidden dimension: Man's Use of Space in Public and Private.* The Bodley Head Ltd, London, UK, 1966.

14 M. Hanheide, A. Peters, and N. Bellotto. Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus. In *RO-MAN, 2012 IEEE*, pages 689–694, 2012.

15 H. Huettenrauch, K. S. Eklundh, A. Green, and A. T. Elin. Investigating spatial relationships in human-robot interaction. In *Int. Conf. on Intelligent Robots and Systems*, pages 5052–5059, 2006.

16 A. Kendon. Spacing and orientation in co-present interaction. In A. Esposito, N. Campbell, C. Vogel, A. Hussain, and A. Nijholt, editors, *Development of Multimodal Interfaces: Active Listening and Synchrony*, volume 5967 of *LNCS*, pages 1–15. Springer, Berlin, 2010.

17 A. Kristoffersson, K. Severinson-Eklundh, and A. Loutfi. Measuring the quality of interaction in mobile robotic telepresence: A pilots perspective. *Int. Journal of Social Robotics*, pages 1–13, 2012.

18 I. Leite, M. McCoy, D. Ullman, N. Salomons, and B. Scassellati. Comparing models of disengagement in individual and group interactions. In *Proc. of 10th Annual ACM/IEEE*

Int. Conf. on Human-Robot Interaction, HRI'15, pages 99–105, New York, NY, USA, 2015. ACM. doi:10.1145/2696454.2696466.

19 J. W. Lloyd. *Foundations of logic programming. Symbolic computation: Artificial intelligence*. Springer-Verlag, 2nd, extended edition edition, 1987.

20 P. Marshall, Y. Rogers, and N. Pantidi. Using F-formations to analyse spatial patterns of interaction in physical environments. In *Proc. of the ACM 2011 Conf. on Computer Supported Cooperative Work*, CSCW'11, pages 445–454, New York, NY, USA, 2011. ACM. doi:10.1145/1958824.1958893.

21 R. Moratz and J. O. Wallgrün. Spatial reasoning with augmented points: Extending cardinal directions with local distances. *J. Spatial Information Science*, 5(1):1–30, 2012. doi:10.5311/JOSIS.2012.5.84.

22 E. Pacchierotti, H. I. Christensen, and P. Jensfelt. Embodied social interaction for service robots in hallway environments. In *Proc. of the International Conference on Field and Service Robotics (FSR'05)*, July 2005.

23 L. Sterling and E. Shapiro. *The Art of Prolog (2nd Ed.): Advanced Programming Techniques*. MIT Press, Cambridge, MA, USA, 1994.

24 T. Tenbrink, K. Fischer, and R. Moratz. Spatial strategies in linguistic human-robot communication. In Christian Freksa, editor, *KI-Themenheft 4/02 Spatial Cognition*, pages 19–23. arenDTaP Verlag, 2002.

25 S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. R. Rosenberg, N. Roy, J. Schulte, and D. Schulz. MINERVA: A second-generation museum tour-guide robot. In *1999 IEEE Int. Conf. on Robotics and Automation, Proc.*, pages 1999–2005, 1999. doi:10.1109/ROBOT.1999.770401.

An Efficient Representation of General Qualitative Spatial Information Using Bintrees

Leif Harald Karlsen[1] and Martin Giese[2]

1 **Department for Informatics, University of Oslo, Oslo, Norway**
leifhka@ifi.uio.no

2 **Department for Informatics, University of Oslo, Oslo, Norway**
martingi@ifi.uio.no

Abstract

In this paper we extend previous work on using bintrees as an efficient representation for qualitative information about spatial objects. Our approach represents each spatial object as a bintree satisfying the exact same qualitative relationships to other bintree representations as the corresponding spatial objects. We prove that such correct bintrees always exists and that they can be constructed as a sum of local representations, allowing a practically efficient construction. Our representation is both efficient, w.r.t. storage space and query time, and can represent many well-known qualitative relations, such as the relations in the Region Connection Calculus and Allen's Interval Algebra.

1998 ACM Subject Classification I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Qualitative spatial data, Bintree, Data structure

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.4

1 Introduction

Spatial and temporal data types are ubiquitous in today's software, with a growing number of spatially aware devices gathering and publishing data. Spatial and temporal data are used in a great number of highly valuable applications, such as route planning, automatic navigation, and modeling of physical processes. However, temporal and especially spatial data are normally represented as complex numerical objects, where relationships between objects are implicit, and advanced algorithms (e.g. from computational geometry) are needed to determine them. Indexing these objects for efficient query answering is also complex. During the last decades, several spatial and temporal database systems have been developed, featuring advanced indexing mechanisms and efficient numerical algorithms for answering queries over these data types (see e.g. [12, 20, 13]). Despite these advances, spatial and temporal data are still significantly more difficult to handle than more traditional types of data, often lag behind when new knowledge representations are introduced and in many cases need special treatment. The present work stems from the following observations:

1. Many applications of spatial data are mostly concerned with *qualitative* relations such as overlaps or containment of spatial objects, rather than *quantitative* properties like distance, area, etc.
2. For such qualitative applications, resorting to expensive computations on the numerical representations for each query seems wasteful. It would be sufficient to store a (precomputed) database table for overlap, containment, or any other relations of interest, treating these relations like any other in a relational database. But this, also seems wasteful, in terms of space, since such tables could be quadratic in the number of geometries

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 4; pp. 4:1–4:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

(for binary relations, cubic for ternary and so on), despite obvious redundancies, like e.g. the transitivity of the containment relation or symmetry of the overlaps relation.

3. Numerical representations of spatial objects are often subject to precision errors. E.g. even though two objects are touching in the real world, their numerical representations might not, due to insufficient precision in either their numerical representation or the measuring device reporting the objects' spatial extent. These errors are difficult or sometimes even impossible to fix numerically without introducing other errors [3]. However, if we construct a new qualitative representation of the objects, we can fix such errors during the translation by using domain knowledge about the objects. For instance, we might know that every country c touches all countries c' whenever their numerical representations have a smallest distance of 1 kilometer, or that the spatial extent of any capital of a country is contained in the extent of that country.
4. Most approaches to qualitative spatial representation can be divided into two types: they are either targeted at complex reasoning tasks (consistency checking, entailment, etc.) and is therefore not suitable as an efficient *data structure* for qualitative information extracted from a set of concrete spatial objects; or they focus on a particular set of relations for a particular type of spatial data. (See Section 7 for more details.)

Our approach is to construct a linear bintree-representation for each spatial object that are correct w.r.t. any given set of qualitative relations definable from a given first-order language. This representation scales to real-world datasets without limiting the approach to any fixed set of relations. The linear bintree [24, 25] consists of a set of bit-strings, each representing a small chunk of space obtained by recursively dividing space. Thus, bintrees represent a union of chunks of space, and two bintrees can therefore e.g. spatially overlap or one can contain the other. We can therefore make one bintree per spatial object that have the same relationships to each other as the spatial objects have, thereby becoming a *representation* of the qualitative relationships between the spatial objects.

Bintrees have the convenient property that they can be stored as a regular relation in a relational database. Furthermore, the bintrees can themselves be indexed by normal database index structures, like B-trees, since they only consist of sets of bit-strings where each bit-string can be represented by one integer. Another desirable feature of bintrees is that they allow variable resolution, so we can have low resolution (few and short bit-strings) for homogeneous areas and high resolution (many and long bit-strings) for heterogeneous areas where more detail is necessary. The bintree has previously been used as an indexing structure for geometries and as an efficient representation for images. Bintrees are now considered obsolete as index structures for geometries, as R-trees [8] and their variants (see e.g. [18] for an overview) have better performance. However, for our purpose of representing qualitative information, we will see that the bintree is a good fit.

The concrete problem this paper addresses is the following: Given a set of objects with a spatial interpretation and a set of qualitative relations, construct a bintree representation that returns the same answers to queries with the given relations over the spatial objects. We have previously constructed both theory [10] and an implementation [11] for constructing such qualitatively correct bintrees, with promising results. However, our previous work has been restricted to the construction of bintree-representations that are correct only w.r.t. part-of and overlaps relationships (as presented in Section 2). In this paper we will extend the theoretical foundation to allow for representations that are correct with respect to a more expressive set of relations.

The paper is outlined as follows: In Section 2 we introduce the spatial objects we work with and the key notions and results needed for expressing and constructing correct bintrees;

in Section 3 we explain how to construct correct bintrees and why this is a good representation for qualitative information; in the Sections 4 and 5 we extend the expressiveness of the relational language our bintrees are correct w.r.t. in two directions; in Section 6 we show several examples of common qualitative relations that our bintrees can represent; Section 7 discusses related work and Section 8 concludes the paper.

2 Spatial Objects and Correct Bintrees

We will start by introducing the central concept of spaces, the elements which we aim to represent correctly.

▶ **Definition 1.** A *space lattice* $\mathcal{S} = (S, \prec_{\mathcal{S}}, \top_{\mathcal{S}}, \bot_{\mathcal{S}})$ is a bounded, distributive lattice with top element $\top_{\mathcal{S}}$ and bottom element $\bot_{\mathcal{S}}$. We will let $\otimes_{\mathcal{S}}$ and $\oplus_{\mathcal{S}}$ be the induced meet and join operators respectively, and call the elements of S *spaces*. We will let $S^+ := S \setminus \{\bot_{\mathcal{S}}\}$.

Typical examples of such space lattices is the lattice of geometrical objects (polygons, lines points) where $\prec_{\mathcal{S}}$ is geometric containment, the lattice of temporal intervals where $\prec_{\mathcal{S}}$ is temporal containment, the lattice of sets where $\prec_{\mathcal{S}}$ is the subset-relation, and so on. Thus, the goal of this paper is to construct efficient representations of the qualitative relationships between such structures. In order to do this, we need to compute these relationships between the spatial objects, however checking all possible relationships between all possible spaces would be very complex, as this has a complexity of $\mathcal{O}(n^k)$ (for n elements with relations of arity k). A property of qualitative relationships like *overlaps* and *contains* is that they are local, that is, they depend only on the spatial parts *inside* the elements, and nothing more. Thus, we want to exploit this locality in a similar fashion as the *bucket sort*-algorithm does, where the elements to sort are first distributed into a set of of buckets/intervals partitioning the universe. The buckets are sorted individually, before being gathered into a sorted list.

In a similar fashion, we will construct a set of chunks of space partitioning the space-lattice's universe $\top_{\mathcal{S}}$, and construct locally correct representations in each chunk. We will call such a chunk a *block*. The blocks are most naturally construed in a recursive fashion where we start with $\top_{\mathcal{S}}$ and recursively split blocks into two smaller blocks, until we reach some desired property (e.g. the desired resolution or the desired number of spatial objects overlapping each block.) This splitting forms a binary tree, so each block can be represented as the path from the root ($\top_{\mathcal{S}}$) down to that block. Furthermore, such a path can be compactly represented as a bit-string (a 0-bit and 1-bit denotes a left-edge and a right-edge resp.) Note that every bit-string denotes a chunk of space, and that if s is a bit-string which is a prefix of s', then the block denoted by s spatially contains the block denoted by s'. If we let a set of bit-strings denote their union, we can represent more complex spaces that can spatially overlap and contain other spaces.

Our representation should allow efficient updates, and since relationships are locally determined, inserting a new object into our representation should only affect the representation of the blocks overlapping the object to insert. However, for such a local insert to be possible we need to know which block each representation was constructed in.

Therefore, it would seem natural to let each element's local block-representation be a set of bit-strings, each contained in that block, which satisfies the same qualitative relationships as the spatial objects they represent. A set of such bit-strings is in fact a linear bintree. The bintree is thus a binary trie data structure, similar to the quadtree and octree. For a discussion and comparison of these three structures, see e.g. [24]. Below follows the formal definition of both bit-strings and bintrees.

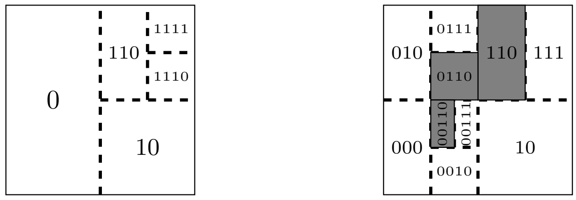


Figure 1 The left figure shows the bit-string representation of some blocks and the right figure the spatial extent (in gray) of the bintree $\{110, 0110, 00110\}$ in 2D.

▶ **Definition 2.** Let $\mathbb{B}$ to be the *set of bit-strings* with ε being the empty bit-string and $b \circ b'$ to be the *concatenation* of the bit-strings b and b'. Let the *prefix-relation* on blocks, $\lhd$, be defined as $b_1 \lhd b_2 \Leftrightarrow \exists b \in \mathbb{B}(b_2 \circ b = b_1)$ and the *neighbor-relation* on blocks, $\sim$, be defined as $b_1 \sim b_2 \Leftrightarrow \exists b \in \mathbb{B}(b_1 = b \circ 0 \wedge b_2 = b \circ 1)$. Define a *block-set* B to be a non-empty, finite set of bit-strings such that if $b \in B$ then B also contains all $b' \in \mathbb{B}$ such that either $b \sim b'$ or $b \lhd b'$.

▶ **Definition 3.** Define the *$\mathcal{T}$-lattice* $\mathcal{T} := (T, \prec_{\mathcal{T}}, \top_{\mathcal{T}}, \bot_{\mathcal{T}})$ where $T = \mathcal{P}_{fin}(\mathbb{B})$ is the set of bintrees, (where $\mathcal{P}_{fin}$ is the finite powerset) such that $t \in T$ contains no two distinct elements b_1, b_2 where either $b_1 \lhd b_2$ or $b_1 \sim b_2$. Furthermore, let $\top_{\mathcal{T}} = \{\epsilon\}$, $\bot_{\mathcal{T}} = \emptyset$, and $t \prec_{\mathcal{T}} t' \Leftrightarrow \forall b \in t \exists b' \in t(b \lhd b')$.

It should be easy to see that the $\mathcal{T}$-lattice is a space lattice. Thus, bintrees behave similarly to spaces, which allows them to be used as representations for spaces.

In Figure 1 we can see an example of both blocks and a bintree, and their spatial extent (assuming regular splitting in each space division). Note that we put no restriction on the number of dimensions our spaces has, and the same holds for our bintrees. In the same way we alternate between splitting along the x- and y-axis in the 2D case, we would cycle through all k dimensions in a k-dimensional space. We will now introduce our models, which will allow us to precisely define correctness of bintree-representations.

▶ **Definition 4.** Given a space lattice $\mathcal{S} = (S, \prec_{\mathcal{S}}, \top_{\mathcal{S}}, \bot_{\mathcal{S}})$, a finite set of constants C and a block-set B, an *S-model* $\mathcal{M}$ is a first order model over the similarity type $\langle \prec; C \cup B \rangle$ with universe S, but where $(\exists^+ z.\varphi)^{\mathcal{M}} \Leftrightarrow \varphi[s/z]^{\mathcal{M}}$ for some $s \in S^+$, and where $\epsilon^{\mathcal{M}} = \top_{\mathcal{S}}$, $\prec^{\mathcal{M}} = \prec_{\mathcal{S}}$, and $b^{\mathcal{M}} \neq \bot_{\mathcal{S}}$, $(b \circ 0)^{\mathcal{M}} \otimes_{\mathcal{S}} (b \circ 1)^{\mathcal{M}} = \bot_{\mathcal{S}}$, and $(b \circ 0)^{\mathcal{M}} \oplus_{\mathcal{S}} (b \circ 1)^{\mathcal{M}} = b^{\mathcal{M}}$ for any $b \in B$.

Note the interpretation of the new existential quantifier, and that if e.g. $\exists^+ z(z \prec c_1 \wedge z \prec c_2)$ holds in some model, then there is a non-empty intersection between c_1 and c_2 in that model. Observe also that given a space-lattice $\mathcal{S}$, the only difference between two $\mathcal{S}$-models is their interpretation of the constants $C \cup B$. The constants C will be the elements which have a spatial interpretation that we wish to correctly represent as bintrees. The constants of B will function as the *buckets* as described above. However, before we can talk about correct representations, we need to define the scope of this correctness. Our notion of correctness will be restricted to a language of first order sentences that nicely captures a core of qualitative relations, namely overlaps and containment relationships. We will in the later sections of this paper extend the expressiveness of the language.

▶ **Definition 5.** Let an *atomic spatial formula* be a first order formula on one of the two forms: $x_1 \prec x_2$ or $\exists^+ z \left(\bigwedge_{i \in I} z \prec x_i \right)$. A *spatial formula* is a first order formula $\varphi(\vec{x})$ defined by the BNF $\varphi := \psi \mid \neg\psi \mid \varphi_1 \wedge \varphi_2$, where ψ is an atomic spatial formula.

Given a set of constants C and a block-set B, an *(atomic) spatial sentence* $\varphi(\vec{c})$ is a first order sentence such that $\varphi(\vec{x})$ is an (atomic) spatial formula and $\vec{c} \in (B \cup C)^{|\vec{c}|}$.

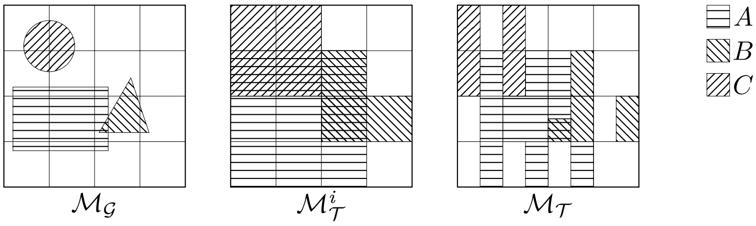


Figure 2 A figure with three polygons to the left, an example of a Γ-incorrect bintree-representation in the middle, and an example of a Γ-correct bintree-representation to the right.

▶ **Definition 6.** Given a set of spatial formulae Γ, we will say that a $\mathcal{T}$-model $\mathcal{M}_\mathcal{T}$ is Γ-*correct* w.r.t. an $\mathcal{S}$-model $\mathcal{M}_\mathcal{S}$ if $\mathcal{M}_\mathcal{T} \vDash \varphi(\vec{c}) \Leftrightarrow \mathcal{M}_\mathcal{S} \vDash \varphi(\vec{c})$ where $\vec{c} \in (B \cup C)^{|\vec{c}|}$ and $\varphi(\vec{x}) \in \Gamma$.

▶ **Example 7.** Let $\Gamma := \{x \prec y, \exists^+ z\,(z \prec x_1 \wedge z \prec x_2)\ \}$. In Figure 2 we can see an example of a Γ-incorrect and a Γ-correct bintree model for the constants $\{A, B, C\} \cup \{b \in \mathbb{B} \mid |b| \leq 4\}$ (where $|b|$ is the length of the bit-string b), w.r.t. a geometric model $\mathcal{M}_\mathcal{G}$. $\mathcal{M}^i_\mathcal{T}$ is just an approximation from above, which is how bintrees are normally used as index structures. We can see that such a representation is complete w.r.t. Γ, i.e. $\mathcal{M}_\mathcal{G} \vDash \varphi(\vec{c}) \Rightarrow \mathcal{M}^i_\mathcal{T} \vDash \varphi(\vec{c})$ for any spatial formula $\varphi(\vec{x}) \in \Gamma$ and any $\vec{c} \in (C \cup B)^{|\vec{c}|}$, but it is not sound, i.e. the converse implication does not necessarily hold. For instance, $\mathcal{M}^i_\mathcal{T} \vDash \mathit{ov}(A, C)$ but $\mathcal{M}_\mathcal{G} \nvDash \mathit{ov}(A, C)$, and $\mathcal{M}^i_\mathcal{T} \vDash \mathit{1011} \prec B$ but $\mathcal{M}_\mathcal{G} \nvDash \mathit{1011} \prec B$. However, $\mathcal{M}_\mathcal{T}$ is an example of a correct model, and it is easy to check that *any* spatial sentence is true in $\mathcal{M}_\mathcal{T}$ if and only if it is true in $\mathcal{M}_\mathcal{G}$.

As stated above, for efficiency reasons we will construct our bintrees locally. Thus, we need a notion of local correctness, that is, what a locally correct bintree-model is.

▶ **Definition 8.** Let $\vDash_b$ for a bit-string b, be equivalent to $\vDash$, but where $\mathcal{M}_\mathcal{S} \vDash_b c \prec d \Leftrightarrow \left(b^{\mathcal{M}_\mathcal{S}} \otimes_\mathcal{S} c^{\mathcal{M}_\mathcal{S}}\ \right) \prec_\mathcal{S} \left(b^{\mathcal{M}_\mathcal{S}} \otimes_\mathcal{S} d^{\mathcal{M}_\mathcal{S}}\ \right)$ and $\mathcal{M}_\mathcal{S} \vDash_b \exists^+ z.\varphi \Leftrightarrow \mathcal{M}_\mathcal{S} \vDash_b \varphi[s/z]$ for some $s \in S^+$ and $s \prec_\mathcal{S} b^{\mathcal{M}_\mathcal{S}}$. Given a block-set B, we will call $\mathcal{M}_\mathcal{T}$ *locally* Γ-*correct* if $\mathcal{M}_\mathcal{T} \vDash_b \varphi(\vec{c}) \Leftrightarrow \mathcal{M}_\mathcal{S} \vDash_b \varphi(\vec{c})$ for all spatial sentences $\varphi(\vec{c})$ where $\varphi(\vec{x}) \in \Gamma$ and all $\lhd$-smallest elements b of B.

So a locally correct model is a model that is correct if we limit out vision to one block at the time. We will now show that our qualitative relations are locally determined, that is, locally correct models are also globally correct.

▶ **Theorem 9.** *Given a set of constants C and a block-set B, any locally Γ-correct $\mathcal{T}$-model $\mathcal{M}_\mathcal{T}$ is Γ-correct, w.r.t. an $\mathcal{S}$-model $\mathcal{M}_\mathcal{S}$.*

Proof. Let β be set of $\lhd$-smallest elements of B. It is sufficient to prove that for any $\mathcal{S}$-model $\mathcal{M}_\mathcal{S}$ we have $\mathcal{M}_\mathcal{S} \vDash c_1 \prec c_2 \Leftrightarrow \forall b \in \beta\,(\mathcal{M}_\mathcal{S} \vDash_b c_1 \prec c_2\ )$ and $\mathcal{M}_\mathcal{S} \vDash \exists^+ z \left(\bigwedge_{i \leq k} z \prec c_i\ \right) \Leftrightarrow \exists b \in \beta \left(\mathcal{M}_\mathcal{S} \vDash_b \exists^+ z \left(\bigwedge_{i \leq k} z \prec c_i\ \right)\ \right)$ for any $c_1, \ldots, c_k \in C \cup B$. By definition the $\mathcal{M}_\mathcal{S}$-interpretation of the elements of β forms a partition on $\top_\mathcal{S}$, so $\top_\mathcal{S} = \bigoplus_{b \in \beta} b^{\mathcal{M}_\mathcal{S}}$ and $b_1^\mathcal{M} \otimes_\mathcal{S} b_2^\mathcal{M} = \bot_\mathcal{S}$. This, together with distributivity, we know that $c_1^{\mathcal{M}_\mathcal{S}} \prec_\mathcal{S} c_2^{\mathcal{M}_\mathcal{S}}$ is equivalent to $\forall b \in \beta \left(\left(b^{\mathcal{M}_\mathcal{S}} \otimes_\mathcal{S} c_1^{\mathcal{M}_\mathcal{S}}\right) \prec_\mathcal{S} \left(b^{\mathcal{M}_\mathcal{S}} \otimes_\mathcal{S} c_2^{\mathcal{M}_\mathcal{S}}\right)\ \right)$ for any $c_1, c_2 \in C \cup B$. By similar arguments, we have that $\exists z \in S^+ \left(\bigwedge_{i \leq k} z \prec_\mathcal{S} c_i^{\mathcal{M}_\mathcal{S}}\right)$ is equivalent to the local $\exists b \in \beta \exists z \in S^+ \left(\bigwedge_{i \leq k} z \prec \left(b^{\mathcal{M}_\mathcal{S}} \otimes_\mathcal{S}\ c_i^{\mathcal{M}_\mathcal{S}}\right)\right)$ for any $c_1, \ldots, c_k \in C \cup B$. ◀

▶ **Theorem 10.** *For any $\mathcal{S}$-model $\mathcal{M}_\mathcal{S}$ there exists a locally Γ-correct bintree-model $\mathcal{M}_\mathcal{T}$.*

Proof. Proof done by model construction: For each $b \in \beta$, construct the set of all locally true atomic spatial sentences occurring (either positively or negatively) in some $\varphi \in \Gamma$:

$$T_b := \left\{ \psi(\vec{c}) \mid \mathcal{M}_{\mathcal{S}} \vDash_b \psi(\vec{c}), \bigwedge_{1 \leq i \leq |\vec{c}|} \left(c_i \in C \cup B \wedge c_i^{\mathcal{M}_{\mathcal{S}}} \otimes b^{\mathcal{M}_{\mathcal{S}}} \neq \bot_{\mathcal{S}} \right) \right\}.$$

Any $\mathcal{T}$-model of all T_bs is locally Γ-correct. Then, let T_b' be the skolemization of T_b, and T_b'' be the set of atoms occurring in any sentence in T_b'. Define K_b^+ to be the set of all constants occurring in T_b'', and $K_b^\perp := (C \cup B) \setminus K_b^+$. So $\{K_b^+, K_b^\perp\}$ partitions $C \cup B$ and K_b^+ is the set of constants that should have a non-empty interpretation, locally in b.

We will now construct the $\mathcal{T}$-model. First, for each $b \in \beta$, generate a set $W_b \subseteq \mathbb{B}$ of size $|K_b^+|$ of pairwise $\lhd$-unrelated bit-strings b' such that $b' \lhd b$. Then, let $w_b : \mathcal{K}_b^+ \to W_b \cup \{b\}$ be a bijective function on $\mathcal{K}_b^+ \setminus B$ and $w_b(b') = b$ for $b' \in \mathcal{K}_b^+ \cap B$. Then, define $I_b(c) := \bigoplus_{\mathcal{T}} \{\{w_b(c')\} \mid (c' \prec c) \in T_b'', c' \in K_b^+\}$ for each $c \in \mathcal{K}_b^+$, and $I_b(c) := \bot_{\mathcal{T}}$ for $c \in \mathcal{K}_b^\perp$. So I_b is the locally correct interpretation of the constants in $C \cup B$, and it should be clear that $\left(\left(c_1^{\mathcal{M}_{\mathcal{S}}} \otimes_{\mathcal{S}} b^{\mathcal{M}_{\mathcal{S}}}\right) \prec_{\mathcal{S}} \left(c_2^{\mathcal{M}_{\mathcal{S}}} \otimes_{\mathcal{S}} b^{\mathcal{M}_{\mathcal{S}}}\right)\right)$ if and only if $I_b(c_1) \prec_{\mathcal{T}} I_b(c_2)$ and $\exists z \in S^+ \left(\bigwedge_{i \leq k} z \prec_{\mathcal{S}} \left(b^{\mathcal{M}_{\mathcal{S}}} \otimes_{\mathcal{S}} c_i^{\mathcal{M}_{\mathcal{S}}}\right)\right)$ if and only if $\exists z \in T^+ \left(\bigwedge_{i \leq k} z \prec_{\mathcal{T}} I_b(c_i)\right)$ for any $c_1, \ldots, c_k \in C \cup B$ and any $b \in \beta$. Finally, let $c^{\mathcal{M}_{\mathcal{T}}} := \bigoplus_{\mathcal{T}} \{I_b(c) \mid b \in \beta\}$ for each $c \in C \cup B$. $\mathcal{M}_{\mathcal{T}}$ is now a Γ-correct $\mathcal{T}$-model w.r.t $\mathcal{M}_{\mathcal{S}}$. ◀

3 How To and Why Construct Correct Bintrees

The proof of Theorem 10 illustrates how one could design an algorithm for construction of correct bintree-models. We can write an almost direct translation of the steps in the proof to an algorithm. That is, for each $b \in \beta$ do the following: Find all $c \in C \cup B$ overlapping b and compute their Γ-relationships, T_b; skolemize and extract the atomic sentences, T_b''; generate a set of blocks and assign each non-empty element a block, and propagate according to the $\prec$-relationships in T_b''; finally, sum up the local representations to form the model. The algorithmic complexity of such a model construction is $|\beta|$ times the complexities of first constructing T_b'' and then generating and distributing the elements of W_b. It should be easy to see that the latter has complexity $\mathcal{O}(|T_b''|)$. Note that constructing T_b' from T_b and T_b'' from T_b' are both linear in the size of T_b. Lastly, we have that constructing T_b requires computing whether $\mathcal{M}_{\mathcal{S}} \vDash \varphi(\vec{c})$ holds for each atomic spatial sentence generated from the atomic spatial formulas of Γ and the constants $B \cup C$. This gives us a total complexity of $\mathcal{O}(|\beta| \cdot o^k)$, where $o = \max_{b \in beta} |\{c \in C \cup B \mid \mathcal{M}_{\mathcal{S}} \vDash ov(c, b)\}|$, that is, the largest number of elements from $C \cup B$ that overlaps any $b \in \beta$, and k is the largest number of free variables occurring in any atomic spatial formula occurring in any $\varphi \in \Gamma$. This means that we in practice *can* construct correct bintree-models for any $\mathcal{S}$-model, however, *why* still remains to be answered. Below we discuss the main properties of the representation making it suitable for representing qualitative information.

The bintrees can be stored and queried in a relational database as a binary relation $(id, block)$, where we encode the bit-strings as integers and where both the IDs and the bit-string integers can be indexed by a normal B-tree. This allows for highly efficient query answering, in the complexity class AC_0 [1], of queries of the form "given $a \in C \cup B$ and $R \in \Gamma$, find all x such that $R(a, x)$ holds" and "given $a, b \in C \cup B$ and $R \in \Gamma$, check whether $R(a, b)$ holds". In [11] we discuss this representation in more detail and present a benchmark that shows that overlaps and containment queries are on average 2.7 times faster over our correct bintrees than over the corresponding geometries. The comparison was done with

PostGIS [20], a state of the art geospatial database, over real-world datasets where the largest sets has over a million geometries.

Note that our constructions also allow a more efficient insertion than reconstructing the entire model upon each insert: Assume we already have constructed a Γ-correct model $\mathcal{M}_\mathcal{T}$ for the $\mathcal{S}$-model $\mathcal{M}_\mathcal{S}$ and constants $C \cup B$, but now want to construct a Γ-correct model for the extended model $\mathcal{M}'_\mathcal{S}$ for $C \cup C' \cup B$. Since we only need local Γ-correctness, we only need to update I_b for each $b \in \beta$ where $\mathcal{M}'_\mathcal{S} \vDash ov(c, b)$ for any $c \in C'$. Thus, a larger B gives a more efficient insert-operation as we have a higher resolution. Observe also that the only requirements we put on the interpretations of the elements of B, is that $\{b \circ 0, b \circ 1\}$ partitions b. Thus, we are free to interpret $b \circ 0$ and $b \circ 1$ in such a way that there is approximately the same number of elements from C that overlap each. This will evenly spread the elements of C over the elements of β, thus making each T_b about equally complex to compute. This is important, as it can greatly reduce the value of o in the complexity measure. We present an algorithm for construction and update of Γ-correct bintrees with such balanced splitting of B in [11], with $\Gamma = \{x \prec y, \exists^+ z(\bigwedge_{i \leq k} z \prec x_i)\}$ for arbitrary k.

Our representation is also compact, as it does not need to explicitly store reflexive, symmetric or transitive closures of the containment and overlaps relationships. There are also many optimizations one can do to get an even more compact and efficient representation: E.g. we can remove all sentences φ from T_b if there is some sentence $\varphi' \in T_b$ such that $\varphi' \rightarrow \varphi$. This will remove all redundant overlaps-witnesses (either implied by a containment-relationship or another overlaps-relationship of higher arity) and reduce the overall size of the bintrees. In the benchmark in [11] we show that our bintree-representation uses only 62% of the space of the corresponding geometry-datasets, and only 22% of the explicit representations, for the largest datasets.

4 Extension: Roles

We have now seen that we can construct a correct bintree-representation for any space lattice, but the correctness is only for spatial sentences of containment and overlaps relationships. We will now see that a small extension to our bintree representations allows us to accommodate a much more interesting set of relationships. First observe that we, e.g., can express the well known RCC8-relations (see e.g. [22, 5]) with only containment and overlaps relations, if we can relate the different *types of parts*:

$$\begin{aligned}
DJ(x, y) &:= \neg ov(x, y) & EC(x, y) &:= ov(x, y) \wedge \neg ov(x^\circ, y^\circ) \\
PO(x, y) &:= ov(x^\circ, y^\circ) \wedge (x \not\prec y) \wedge (y \not\prec x) & EQ(x, y) &:= x \prec y \wedge y \prec x \\
TPP(x, y) &:= x \prec y \wedge ov(\partial x, \partial y) \wedge (y \not\prec x) & NTPP(x, y) &:= x \prec y \wedge \neg ov(\partial x, \partial y)
\end{aligned}$$

where $ov(x, y) := \exists^+ z(z \prec x \wedge z \prec y)$, ∂x is the boundary of x, x° is the interior of x and $x \not\prec y$ is short for $\neg(x \prec y)$. We will therefore extend our definitions above with the notion of *roles*, which allows us to talk about the different parts of a space, e.g. interior and boundary.

▶ **Definition 11.** A *role* is a set of names. A *role-set* is a set of roles containing $\emptyset$.

As we will see shortly, we only need roles that consist of a single name to express the relations of RCC8, namely i for interior and b for boundary. However, we will also see examples where using multiple names to denote a part is useful.

▶ **Definition 12.** Given a role-set R, an *R-roled space lattice* $\mathcal{S}$ is a tuple $(S, \prec_\mathcal{S}, \top_\mathcal{S}, \bot_\mathcal{S}, \pi_\mathcal{S})$ where $(S, \prec_\mathcal{S}, \top_\mathcal{S}, \bot_\mathcal{S})$ is a space lattice and $\pi_\mathcal{S} : R \times S \to S$ is a function where $\pi_\mathcal{S}(\emptyset, s) = s$ and $\pi_\mathcal{S}(r \cup u, s) = \pi_\mathcal{S}(r, s) \otimes_\mathcal{S} \pi_\mathcal{S}(u, s)$ for any $(r, u) \in R^2$ such that $r \cup u \in R$ and any $s \in S$.

The reader can read $\pi_{\mathcal{S}}(r, a)$ as "a's r-part". Intuitively one can think of a $\{n_1, \ldots, n_k\}$-part as an intersection of all the $\{n_i\}$-parts. For instance the role $\{i, h\}$, where h is short for hole, denotes holes in an interior, whereas $\{i\}$ denotes all of the interior, both with and without holes. Observe also that we always have $\pi_{\mathcal{S}}(r, s) \prec_{\mathcal{S}} s$. We will now introduce the corresponding bintrees.

▶ **Definition 13.** Given a role-set R, an *R-roled block* is a pair (r, b) such that $b \in \mathbb{B}$ and $r \in R$. Let $\mathbb{B}_R$ be the set of R-roled blocks. Also let $\delta(r, t) := \{b \mid (r, b) \in t\}$ for any $r \in R$ and $t \subseteq \mathbb{B}_R$ and let $\Sigma_R(t) := \bigoplus_{r \in R} \delta(r, t)$.

An *R-roled bintree* t is an element of $\mathcal{P}_{fin}(\mathbb{B}_R)$, such that for any role $r \in R$ we have that $\delta(r, t)$ is a bintree, and $\delta(r, t) \otimes_{\mathcal{T}} \delta(u, t) = \bot_{\mathcal{T}}$ for any $r \neq u$. Let T_R be the set of R-roled bintrees. Furthermore, let $\top_{\mathcal{T}_R} := \{(\emptyset, \varepsilon)\}$, $\bot_{\mathcal{T}_R} := \emptyset$, $\pi_{\mathcal{T}_R}(r, t) := \{(u, b) \in t \mid r \subseteq u\}$, $t \prec_{\mathcal{T}_R} t' \Leftrightarrow \Sigma_R(t) \prec_{\mathcal{T}} \Sigma_R(t')$ and $\mathcal{T}_R := (T_R, \prec_{\mathcal{T}_R}, \top_{\mathcal{T}_R}, \bot_{\mathcal{T}_R}, \pi_{\mathcal{T}_R})$.

While the different roles for the parts are implicitly defined for spaces like geometries, (such as being the interior of a polygon), we explicitly state the roles each block should have in the bintree. So the boundary of a bintree t, $\pi_{\mathcal{T}}(\{b\}, t)$, is the set of blocks having a role r such that $b \in r$. We can then define the touching relation as $ov(\pi_{\{b\}}(x), \pi_{\{b\}}(y))$. So even though two bintrees seem to touch geometrically (e.g. if one has a block b and the other a block b' and $b \sim b'$) they will not necessarily touch according to our definition. This makes it easier for us to construct correct bintree-models, as we still only have to care about overlaps and part-of relationships. Note also that it is possible to construct bintree-models that satisfy sentences that are unsatisfiable by any $\mathcal{S}$-model for a particular space lattice $\mathcal{S}$. For instance, it is easy to make a bintree model with two objects that have a partially overlapping interior, but that have disjoint boundaries, which is impossible for any geometrical model. Thus, we cannot use our representation for reasoning (that is, make a representation for a set of sentences and then query for all entailments). However, as our bintrees only function as a representation of the relationships of a given $\mathcal{S}$-model and is constructed to satisfy exactly these, this is not a problem.

▶ **Definition 14.** Given an R-roled space lattice $\mathcal{S}$, a set of constants C, and a block-set B, an *R-roled $\mathcal{S}$-model* $\mathcal{M}$ is a first order model over the similarity type $\langle \prec; \pi; C \cup B \rangle$, where π is a family of unary function symbols π_r for each $r \in R$, that is an $\mathcal{S}$-model over $\langle \prec; C \cup B \rangle$ and where $\pi_r(c)^{\mathcal{M}_{\mathcal{S}}} = \pi_{\mathcal{S}}(r, c^{\mathcal{M}_{\mathcal{S}}})$ for any $r \in R$ and $c \in C \cup B$.

▶ **Definition 15.** Given a role-set R, an *R-roled atomic spatial formula* is a first order formula on one of the forms $\pi_{r_1}(x) \prec \pi_{r_2}(y)$ or $\exists^+ z \left(\bigwedge_{i \leq k} z \prec \pi_{r_i}(x_i) \right)$ for some $r_1, \ldots, r_k \in R$. Let R-roled formulae and R-roled (atomic) spatial sentences be defined analogously as in Definition 5, but where ψ is an R-roled (atomic) spatial formula.

Note that $(\pi_\emptyset(s))^{\mathcal{M}_{\mathcal{S}}} = s$ for any R-roled $\mathcal{S}$-model $\mathcal{M}_{\mathcal{S}}$, so we sometimes write x instead of $\pi_\emptyset(x)$ in the definitions of spatial formulae. To save ink, let $\bar{r} = \{r\}$ for any role-name r.

▶ **Example 16** (RCC8)**.** Assume we have the names b for "boundary", and i for "interior", where $\pi_{\bar{b}}(x)$ denotes x's boundary and $\pi_{\bar{i}}(x)$ denotes x's interior, we can now express the RCC8-relations with Γ equal to the set of formulae:

$$DJ(x, y) := \neg ov(x, y) \qquad EC(x, y) := ov(x, y) \wedge \neg ov(\pi_{\bar{i}}(x), \pi_{\bar{i}}(y))$$
$$PO(x, y) := ov(\pi_{\bar{i}}(x), \pi_{\bar{i}}(y)) \wedge (x \not\prec y) \wedge (y \not\prec x) \qquad EQ(x, y) := x \prec y \wedge y \prec x$$
$$TPP(x, y) := x \prec y \wedge ov(\pi_{\bar{b}}(x), \pi_{\bar{b}}(y)) \wedge (y \not\prec x) \qquad NTPP(x, y) := x \prec y \wedge \neg ov(\pi_{\bar{b}}(x), \pi_{\bar{b}}(y))$$

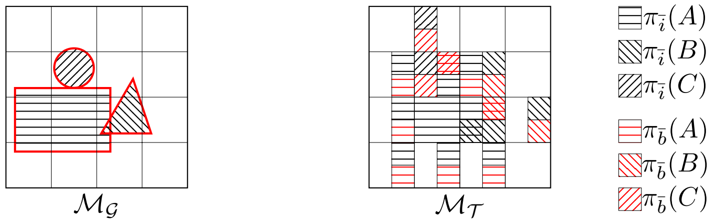


Figure 3 Polygons and an RCC8-correct bintree-model.

Let G be set of two-dimensional geometries (i.e. polygons, line-strings, points) contained in some universe $\top_{\mathcal{G}}$ and with $\prec_{\mathcal{G}}$ being geometric containment, then $\mathcal{G}$ is a space lattice. So a Γ-correct $\mathcal{T}_R$-model w.r.t. $\mathcal{G}$-models $\mathcal{M}$ will correctly represent all RCC8-relations between the elements of C as interpreted by $\mathcal{M}$. In Figure 3 we see an example of a correct bintree-model with respect to the RCC8-relations.

▶ **Theorem 17.** *Any R-roled $\mathcal{T}_R$-model $\mathcal{M}_{\mathcal{T}}$ is Γ-correct if and only if it is locally Γ-correct, w.r.t. an R-roled $\mathcal{S}$-model $\mathcal{M}_{\mathcal{S}}$. Furthermore, for any R-roled $\mathcal{S}$-model $\mathcal{M}_{\mathcal{S}}$ there exists a locally Γ-correct R-roled $\mathcal{T}_R$-model $\mathcal{M}_{\mathcal{T}}$.*

Proof. The arguments for the first part are analogous to the proof of Theorem 9, just substitute c_i with $\pi_{r_i}(c_i)$.

The second part is done by a similar model construction as for the proof of Theorem 10. So, construct T''_b in the same way, but note that now the elements of T''_b are on the forms $\pi_r(c_1) \prec \pi_u(c_2)$ and $v \prec \pi_r(c)$. Now, define K^+_b to be the set of expressions (v and $\pi_r(c)$ where v is a skolem-constant and $c \in C \cup B$) occurring in T''_b, and $K^\perp_b$ as before. Let W_b be a set of size $|K^+_b|$ of pairwise $\lhd$-unrelated bit-strings b' where $b' \lhd b$. Then, let $w_b : K^+_b \to W_b \cup \{b\}$ be a function such that assigns a unique element from W_b to each $e \in K^+_b \setminus \{\pi_\emptyset(b') \mid b' \in B\}$, and $w_b(e) := b$ for each $e \in K^+_b \cap \{\pi_\emptyset(b') \mid b' \in B\}$. We then define $I_b(\pi_r(c)) := \bigoplus_{\mathcal{T}_R} \{\{(r, w_b(e))\} \mid (e \prec \pi_r(c)) \in T''_b, e \in K^+_b\}$ for each $\pi_r(c) \in \mathcal{K}^+_b$ and $I_b(e) := \bot_{\mathcal{T}_R}$ for $e \in \mathcal{K}^\perp_b$. It should now be clear that $\left(b^{\mathcal{M}_S} \otimes_{\mathcal{S}} (\pi_r(c_1))^{\mathcal{M}_S}\right) \prec_{\mathcal{S}} \left(b^{\mathcal{M}_S} \otimes_{\mathcal{S}} (\pi_u(c_2))^{\mathcal{M}_S}\right)$ if and only if $I_b(\pi_r(c_1)) \prec_{\mathcal{T}_R} I_b(\pi_u(c_2))$ and $\exists z \in S^+ \left(\bigwedge_{i \leq k} z \prec_{\mathcal{S}} \left(b^{\mathcal{M}_S} \otimes_{\mathcal{S}} (\pi_{r_i}(c_i))^{\mathcal{M}_S}\right)\right)$ if and only if $\exists z \in T^+_R \left(\bigwedge_{i \leq k} z \prec_{\mathcal{T}_R} I_b(\pi_{r_i}(c_i))\right)$ for any $c_1, \ldots, c_k \in C \cup B$ and any $b \in \beta$. Finally, we let $c^{\mathcal{M}_{\mathcal{T}}} := \bigoplus_{b \in \beta} \bigoplus_{r \in R} I_b(\pi_r(c))$ for each $c \in C \cup B$. $\mathcal{M}_{\mathcal{T}}$ is now an R-roled $\mathcal{T}_R$-model that satisfies exactly the true spatial sentences of $\mathcal{M}_{\mathcal{S}}$ generated from Γ and $C \cup B$. ◀

From the above proof, we can see that the construction of correct roled bintrees is done in a similar fashion as the normal bintrees, and we only need a minor update of any algorithm used for constructing normal correct bintrees.

Observe also that we can compress our roled bintrees in the following manner: Assume that in the set of skolemized atoms T''_b we have a $\pi_r(c)$ such that $(\pi_r(c) \prec e_1) \in T''_b \Leftrightarrow (\pi_u(c) \prec e_1) \in T''_b$ and $(e_2 \prec \pi_r(c)) \in T''_b \Leftrightarrow (e_2 \prec \pi_u(c)) \in T''_b$ for any expressions $e_1, e_2 \in K^+_b$. If then there is no formula $\varphi(\vec{x}) \in \Gamma$ such that $eq(\pi_r(c), \pi_u(c)) \Leftrightarrow \varphi(\vec{c})$, we can let $w_b(r, c) = w_b(u, c)$, thus reducing the size of our bintree-representation. This can for instance be done for the RCC8-relations (letting $eq(\pi_{\bar{b}}(c), \pi_{\bar{i}}(c))$ cannot introduce a new relationship, if $\pi_{\bar{b}}(c)$ and $\pi_{\bar{i}}(c)$ has the exact same relationships to other elements).

Note also that any role can be represented as a fixed length bit-string by enumerating all role-names occurring in Γ and represent each role r as the bit-string having 1s at the bit-positions corresponding to the numbers given to the role-names in r, and 0 everywhere else. We can then represent our roled bintrees as ternary relations $(id, block, role)$ where each column can be index by a normal B-tree. Thus, querying roled bintrees is almost as efficient as querying our normal bintrees, as we only need to consult one additional index-structure (the B-tree over the *role*-column) during query execution.

5 Extension: Order

Introducing roles allows us to construct much richer bintree-models. However, having only the part-of relations allows only relations based on sharing of different types of parts, we are still unable to describe many interesting qualitative relationships, such as temporal relationships, relative size and relative direction. In this section we will extend our language to also include a different type of partial order which will enable us to express these relationships.

▶ **Definition 18.** An *ordered R-roled space lattice* $\mathcal{S}$ is a tuple $(S, \prec_{\mathcal{S}}, <_{\mathcal{S}}, \top_{\mathcal{S}}, \bot_{\mathcal{S}}, \pi_{\mathcal{S}})$ where $(S, \prec_{\mathcal{S}}, \top_{\mathcal{S}}, \bot_{\mathcal{S}}, \pi_{\mathcal{S}})$ is an R-roled space lattice and $<_{\mathcal{S}}$ is a strict partial order such that if $a <_{\mathcal{S}} b$ then $a \otimes_{\mathcal{S}} b = \bot_{\mathcal{S}}$ and for any pair $c, d \in S$ we have $c \prec_{\mathcal{S}} a \wedge d \prec_{\mathcal{S}} b \ \rightarrow c <_{\mathcal{S}} d$.

The reader can read the statement $x < y$ as "x is *before* y". The rest of the definitions are analogous to before:

▶ **Definition 19.** Let $t <_{\mathcal{T}} t' \Leftrightarrow \forall b \in t \forall b' \in t' (b <_{\mathbb{B}} b')$ where $b <_{\mathbb{B}} b'$ for bit-strings b, b' iff there exists some b'' such that $b \lhd b'' \circ 0$ and $b' \lhd b'' \circ 1$. Then let $t <_{\mathcal{T}_R} t' \Leftrightarrow \Sigma_R(t) <_{\mathcal{T}} \Sigma_R(t')$ and $\mathcal{T}_R^< := (T_R, \prec_{\mathcal{T}_R}, <_{\mathcal{T}_R}, \top_{\mathcal{T}_R}, \bot_{\mathcal{T}_R}, \pi_{\mathcal{T}_R})$.

▶ **Definition 20.** Given an ordered R-roled space lattice $\mathcal{S}$, a set of constants C, and a block-set B, an *ordered R-roled $\mathcal{S}$-model* $\mathcal{M}$ is a first order model over the similarity type $\langle \prec, <; \pi; C \cup B \rangle$ that is an R-roled $\mathcal{S}$-model over $\langle \prec; \pi; C \cup B \rangle$, and where $(<)^{\mathcal{M}} = <_{\mathcal{S}}$ and $(b \circ 0)^{\mathcal{M}_{\mathcal{S}}} <_{\mathcal{S}} (b \circ 1)^{\mathcal{M}_{\mathcal{S}}}$ for any $(b \circ 0), (b \circ 1) \in B$.

▶ **Definition 21.** Let an *atomic ordered R-roled spatial formula* be a first order formula that is either an atomic R-roled spatial formula or a formula on the form $x < y$. Let *ordered R-roled spatial formulae* and *(atomic) ordered R-roled spatial sentences* be defined analogously as R-roled spatial formulae and (atomic) R-roled spatial sentences, but where each ψ is an atomic ordered R-roled spatial formula.

▶ **Example 22** (Allen's Interval Algebra)**.** Assume we have the role-names i for *interior*, f for *first*, l for *last*, and the role-set $R := \{\emptyset, \bar{i}, \bar{f}, \bar{l}\}$. Let $\pi_{\bar{f}}(x)$ denote the interval consisting of only x's first point, and $\pi_{\bar{l}}(x)$ denote the interval consisting of only x's last point, and $\pi_{\bar{i}}(x)$ is the interior of x's interval. We can then express the relations of Allen's Interval Algebra [2]:

$$\begin{aligned} &\mathit{before}(x,y) := \pi_{\bar{l}}(x) < \pi_{\bar{f}}(y) && \mathit{meets}(x,y) := \mathit{eq}(\pi_{\bar{l}}(x), \pi_{\bar{f}}(y)) \\ &\mathit{overlaps}(x,y) := \mathit{ov}(\pi_{\bar{i}}(x), \pi_{\bar{i}}(y)) \wedge (x \not\prec y) \wedge (y \not\prec x) && \mathit{equal}(x,y) := \mathit{eq}(x,y) \\ &\mathit{starts}(x,y) := \mathit{eq}(\pi_{\bar{f}}(x), \pi_{\bar{i}}(y)) \wedge \pi_{\bar{l}}(x) \prec \pi_{\bar{i}}(y) && \mathit{during}(x,y) := \pi_{\bar{f}}(x) \prec \pi_{\bar{i}}(y) \wedge \pi_{\bar{l}}(x) \prec \pi_{\bar{i}}(y) \\ &\mathit{ends}(x,y) := \mathit{eq}(\pi_{\bar{l}}(x), \pi_{\bar{l}}(y)) \wedge \pi_{\bar{f}}(x) \prec \pi_{\bar{i}}(y) && \mathit{after}(x,y) := \pi_{\bar{l}}(y) < \pi_{\bar{f}}(x) \end{aligned}$$

Given the set I of time intervals contained in some universe $\top_{\mathcal{I}}$, with $\prec_{\mathcal{I}}$ being temporal containment, and $x <_{\mathcal{I}} y$ is the temporal *before*, it should be obvious that this forms an ordered R-roled space lattice. Thus, any correct $\mathcal{T}_R^<$-model w.r.t. such an $\mathcal{I}$-model $\mathcal{M}$ will correctly represent all Allen's Interval-relations between the elements of $C \cup B$ as $\mathcal{M}$.

▶ **Theorem 23.** *Any ordered R-roled $\mathcal{T}_R^<$-model $\mathcal{M}_{\mathcal{T}}$ is Γ-correct if and only if it is locally Γ-correct, w.r.t. an ordered R-roled $\mathcal{S}$-model $\mathcal{M}_{\mathcal{S}}$. Furthermore, for any ordered R-roled $\mathcal{S}$-model $\mathcal{M}_{\mathcal{S}}$ there exists a locally Γ-correct ordered R-roled $\mathcal{T}_R^<$-model $\mathcal{M}_{\mathcal{T}}$.*

Proof. For the first part, note that, since β is a partition of $\top_{\mathcal{S}}$ and $(b \circ 0) <_{\mathcal{S}} (b \circ 1)$ for any $(b \circ 0), (b \circ 1) \in B$, we have that $<_{\mathcal{S}}$ is a total order on β. This implies that $e_1^{\mathcal{M}_{\mathcal{S}}} <_{\mathcal{S}} e_2^{\mathcal{M}_{\mathcal{S}}}$ if and only if $\forall b \in \beta \left(\left(e_1^{\mathcal{M}_{\mathcal{S}}} \otimes_{\mathcal{S}} b^{\mathcal{M}_{\mathcal{S}}} \right) <_{\mathcal{S}} \left(e_2^{\mathcal{M}_{\mathcal{S}}} \otimes_{\mathcal{S}} b^{\mathcal{M}_{\mathcal{S}}} \right) \right)$. The rest of the proof is analogous to the proof of Theorem 9.

For the second part, we again have to construct a locally correct model. So, construct T''_b in the same way as before for each $b \in \beta$, but this time the elements of T''_b can also be on the form $e_1 < e_2$. Let W_b be as before but now with size $2|K_b^+|$. We then let $c <_b d \Leftrightarrow (c < d) \in T''_b$, and $<_b^t$ be some strict total ordering on K_b^+ containing $<_b$. Now, define $w_b(c) := \{b_c^f\} \oplus \{b_c^l\}$ (intuitively, one can think of b_c^f and b_c^l as representing the $<$-first and last part of c, respectively) for some $b_c^f, b_c^l \in W_b$ such that $b_c^f <_{\mathbb{B}} b_c^l$ and $c <_b d \Rightarrow b_c^l <_{\mathbb{B}} b_d^f$ and $c \not<_b d \wedge c <_b^t d \Rightarrow (b_c^f <_{\mathbb{B}} b_d^f <_{\mathbb{B}} b_c^l <_{\mathbb{B}} b_d^l)$. Now $w_b(c)$ and $w_b(d)$ are disjoint and $c <_b d \Leftrightarrow w_b(c) <_{\mathcal{T}_R} w_b(d)$ for any pair of distinct $c, d \in K_b^+$. We then define $I_b(e)$ and $\mathcal{M}_{\mathcal{T}}$ in the same way as before. Now, $\mathcal{M}_{\mathcal{T}}$ is an ordered R-roled model satisfying exactly the same ordered R-roled sentences generated from Γ as $\mathcal{M}_{\mathcal{S}}$. ◀

Again we see that the construction of correct bintrees with order requires only a small extension to the previous algorithm. Furthermore, a nice feature of encoding bit-strings as integers as described in Section 3 is that the $<$-ordering of the blocks corresponds to the normal $<$-ordering on their integer representations, thus we can reuse the B-tree index over the blocks to efficiently answer $<$-queries as well.

6 Expressiveness and More Examples

▶ **Example 24** (Holes). To both $\mathcal{G}$ and $\mathcal{I}$ we can add an additional role-name, h, for "hole", that can be combined with e.g. i to represent holes in the interior of a polygon or interval, or with b to represent geometries that have an open boundary. We can now express:

$$\mathit{surroundedBy}(x,y) := x \prec \pi_{\{h,i\}}(y) \qquad \mathit{hasHoles}(x) := \exists^+ z(z \prec \pi_{\bar{h}}(x))$$
$$\mathit{hasOpenBoundary}(x) := \pi_{\bar{b}}(x) \prec \pi_{\{b,h\}}(x) \quad \mathit{hasHole}(x,y) := \mathit{eq}(\pi_{\{h,i\}}(x), y)$$

▶ **Example 25** (Relative size and direction). One dimensional attributes like size, length, projection down to the north-south and east-west axis can easily be represented by introducing an appropriate role-name, e.g. d, and let $\pi_{\mathcal{S}}(\bar{d}, x) <_{\mathcal{S}} \pi_{\mathcal{S}}(\bar{d}, y)$ hold if x has a smaller value than y on the d-axis. If we then also let for each $b \in \beta$, $\pi_{\mathcal{S}}(\bar{d}, b)$ be an interval along this axis such that β contains both the smallest and largest values, our constructing algorithm will be a normal bucket-sort with β being the set of buckets.

If we introduce the role-names n for the projection along the north-south and e for the projection down to the east-west, we can express the following relations from the Cardinal Direction Calculus[14], e.g.:

$$\mathit{northOf}(x,y) := \pi_{\bar{n}}(y) < \pi_{\bar{n}}(x) \wedge \mathit{ov}(\pi_{\bar{e}}(x), \pi_{\bar{e}}(y))$$
$$\mathit{northEastOf}(x,y) := \pi_{\bar{n}}(y) < \pi_{\bar{n}}(x) \wedge \pi_{\bar{e}}(y) < \pi_{\bar{e}}(x)$$

and the rest of the directional-relations are defined similarly. Note that $\pi_{\mathcal{S}}(\bar{n}, x)$ and $\pi_{\mathcal{S}}(\bar{e}, x)$ is the projection of a two-dimensional object down to the each dimension. We can of course also do this for three-dimensional (or higher) objects and introduce a role-name, u for the up-down axis, and relations such as $\mathit{above}(x,y)$ and $\mathit{between}(x,y,z)$. If we combine the directional roles with the interior-role, e.g. $\{i,u\}$, we can express

$$\mathit{onTopOf}(x,y) := \mathit{ov}(\pi_{\{i,n\}}(x), \pi_{\{i,n\}}(y)) \wedge \mathit{ov}(\pi_{\{i,e\}}(x), \pi_{\{i,e\}}(y)) \wedge \mathit{ov}(\pi_{\{b,u\}}(x), \pi_{\{b,u\}}(y)) \wedge \pi_{\{i,u\}}(y) < \pi_{\{i,u\}}(x)$$

that is, x and y overlap in the two-dimensional plane, but x and y are touching along the up-down axis, yet x's interior is above y's.

▶ **Example 26** (Orientation). If we have the directional roles $\{n, e\}$ as described above, we can introduce two more role-names f for *front* and b for *back*, and then introduce orientational relations, e.g $\mathit{northOriented}(x) := \pi_{\{n,b\}}(x) < \pi_{\{n,f\}}(x) \wedge ov(\pi_{\{e,b\}}(x), \pi_{\{e,f\}}(x))$ and similarly for the rest of the directions. If we allow unions of relations in our query language (this is trivial in SQL), we can express relative orientation, that is, $\mathit{orientedTowards}(x, y)$ as the union of the 8 relations on the form $\mathit{northOf}(x, y) \wedge \mathit{southOriented}(x)$.

▶ **Example 27** (Egg-Yolk). If we have a space-lattice $\mathcal{S}$ with indeterminate boundaries (that is, an inner and outer boundary where the real boundary is somewhere in between) we can introduce two new role-names y, for *yolk*, and w, for *white*, and let $\pi_{\mathcal{S}}(\bar{y}, s)$ be the region within the inner boundary and $\pi_{\mathcal{S}}(\bar{w}, s)$ be the region within the outer boundary. We can then introduce all the 46 relations from the Egg-Yolk RCC5 calculus [6], e.g.:

$$R_2(x, y) := PO'(\pi_{\bar{w}}(x), \pi_{\bar{w}}(y)) \wedge \neg ov(\pi_{\bar{y}}(x), y) \wedge \neg ov(\pi_{\bar{y}}(y), x)$$
$$R_{11}(x, y) := PO'(\pi_{\bar{w}}(x), \pi_{\bar{w}}(y)) \wedge PO'(\pi_{\bar{w}}(x), \pi_{\bar{y}}(y)) \wedge (\pi_{\bar{y}}(x) \prec \pi_{\bar{w}}(y)) \wedge \neg ov(\pi_{\bar{y}}(x), \pi_{\bar{y}}(y))$$

where $PO'(x, y) := ov(x, y) \wedge (x \not\prec y) \wedge (y \not\prec x)$. $R_2(x, y)$ states that the white of the two partially overlap whereas the yolks are disjoint from each other's eggs, and $R_{11}(x, y)$ that x's white partially overlap both y's white and yolk, and x's yolk is contained in y's white.

It is also possible to combine any of the above relation-sets whenever the underlying space-lattice has a natural interpretations for each relation-set's roles. For instance, for spatio-temporal objects one could combine Allen's Interval Algebra and RCC8.

There is, of course, qualitative information that cannot be represented by our bintrees, e.g. unknown data via disjunctions, such as $EC(a, b) \vee PO(a, b)$ but where we do not know which, since our representation is a concrete model (note that we can model certain types of unknown data by introducing appropriate roles, such as done in Example 27); unions such as $a \prec b \oplus c \wedge a \not\prec b \wedge a \not\prec c$, we can only state that $ov(a, b) \wedge ov(a, c)$; space-lattices that require infinite sets of roles, such as fuzzy sets with membership-roles in $[0, 1]$; formulae with role-variables, such as $R(z, x, y) := \pi_z(x) \prec y$; or shape-relations, we have not found a way to express formulae that can state e.g. concavity.

7 Related Work

There has been done much work on efficient representations of transitive relations and structures for reachability queries in directed graphs (see e.g. [19, 27, 9]) which can be used to represent our containment relationships. However, these representations do not facilitate efficient construction or update of these structures from a set of spatial objects. They are also less expressive, as they do not have any concept similar to our roles or the $<$-ordering. In [21] the authors developed a qualitative representation of spatial data based on arrays of representative points. However, this representation has the same drawbacks as above.

There has also been done a lot of work on representing qualitative spatial information as a set of assertions in some spatial logic, whereby the main information extraction method is logical reasoning based on either logical calculi or constraint solving (see e.g. [7, 4] for an overview). These representations are more focused on complex reasoning problems rather than efficient query answering. These reasoning problems are normally at least NP-hard in general, but tractable restrictions exists (see e.g. [23, 26, 17] for RCC8) that can scale to large datasets. However, as the related work above, these approaches presupposes the existence of a constraint network, and does not themselves provide any efficient construction algorithm of these constraint networks, nor any efficient update of already constructed networks.

In [15] the authors construct a compact representation for the RCC8 and CDC (Cardinal Direction Calculus) relations over polygons using a combination of minimum bounding rectangles (MBR) for each polygon and normal relational database tables when a relation cannot be computed from the MBRs. The authors of [16] provide an efficient construction of a representations of RCC8-relationships between spatial objects via sets of rectangular pseudo-solutions. Each pseudo-solution consists of a partial interpretation of spatial objects into rectangles that encodes one part of an RCC8-network. Both of the approaches above give an efficient method for constructing their respective representations from a set of spatial objects, the former using MBRs and the latter using quadtrees. However, they are both limited to RCC8 and CDC relations over two dimensional objects, whereas our approach can handle a more expressive set of relations over elements from any space lattice.

8 Conclusion and Future Work

We have seen that we always can construct a bintree representation for any space-lattice that is correct w.r.t. any predefined set of qualitative relations expressible in our formula language. This formula language is expressive enough to express most of the common qualitative spatial relations. Our bintree representations are compact, can be stored naturally in any tuple-based representation (relational databases, triple-stores, etc.) and allow highly efficient query answering as they can be stored in a relational database and indexed by B-trees.

In the future we want to extend our implementation [11] (that currently handles all relations definable from the formulae of Definition 5) to also handle the role and order extensions and test these against real-world datasets with expressive relation-sets. We also want to compare our approach to the related representations for RCC8 and CDC described in Section 7.

It would also be interesting to try to extend the language of our relations, to for instance allow intersections, unions, or some restricted form of universal quantification in our formulae without effecting the computational properties of the representation.

References

1. S. Abiteboul, R. Hull, and V. Vianu. *Foundations of databases*, volume 8. Addison-Wesley Reading, 1995.
2. J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983. `doi:10.1016/b978-1-4832-1447-4.50033-x`.
3. M. Attene, M. Campen, and L. Kobbelt. Polygon mesh repairing: An application perspective. *ACM Computing Surveys (CSUR)*, 45(2):15, 2013. `doi:10.1145/2431211.2431214`.
4. J. Chen, A. G. Cohn, D. Liu, S. Wang, J. Ouyang, and Q. Yu. A survey of qualitative spatial representations. *The Knowledge Engineering Review*, 30(01):106–136, 2015. `doi:10.1017/S0269888913000350`.
5. A. G. Cohn, J. Bennett, B.and Gooday, and N. M. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1(3):275–316, 1997. `doi:10.1023/A:1009712514511`.
6. A. G. Cohn and N. M. Gotts. The 'egg-yolk' representation of regions with indeterminate boundaries. *Geographic objects with indeterminate boundaries*, 2:171–187, 1996.

7 A. G. Cohn and J. Renz. Chapter 13 qualitative spatial representation and reasoning. In *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*, pages 551 – 596. Elsevier, 2008. doi:https://doi.org/10.1016/S1574-6526(07)03013-1.

8 A. Guttman. *R-trees: a dynamic index structure for spatial searching*, volume 14. ACM, 1984. doi:10.1145/971697.602266.

9 R. Jin and G. Wang. Simple, fast, and scalable reachability oracle. *Proceedings of the VLDB Endowment*, 6(14):1978–1989, 2013. doi:10.14778/2556549.2556578.

10 L. H. Karlsen and M. Giese. A Framework for Constructing Correct Qualitative Representations of Geometries using Mereology over Bintrees. *Annals of Computer Science and Information Systems*, 7:21–33, 2015. doi:10.15439/2015F183.

11 L. H. Karlsen and M. Giese. An Efficient Representation of Qualitative Spatial Information using Bintrees. Technical report, Department of Informatics, University of Oslo, 2017. http://hdl.handle.net/10852/53792.

12 M. Koubarakis. *Spatio-temporal databases: The CHOROCHRONOS approach*, volume 2520. Springer Science & Business Media, 2003. doi:10.1007/b83622.

13 K. Kyzirakos, M. Karpathiotakis, and M. Koubarakis. Strabon: a semantic geospatial DBMS. In *ISWC'12*, pages 295–311. Springer, 2012. doi:10.1007/978-3-642-35176-1_19.

14 G. É. Ligozat. Reasoning about cardinal directions. *Journal of Visual Languages & Computing*, 9(1):23–44, 1998. doi:10.1006/jvlc.1997.9999.

15 Z. Long, M. Duckham, S. Li, and S. Schockaert. Indexing large geographic datasets with compact qualitative representation. *International Journal of Geographical Information Science*, 30(6):1072–1094, 2016. doi:10.1080/13658816.2015.1104535.

16 Z. Long, S. Schockaert, and S. Li. Encoding large RCC8 scenarios using rectangular pseudo-solutions. In *Proceedings of the Fifteenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 463–472. AAAI Press, 2016.

17 Z. Long, M. Sioutis, and S. Li. Efficient path consistency algorithm for large qualitative constraint networks. *IJCAI'16 Proceedings*, 2016. doi:10.1142/S0218213015500311.

18 Y. Manolopoulos, A. Nanopoulos, A. N. Papadopoulos, and Y. Theodoridis. *R-trees: Theory and Applications.* Springer Science & Business Media, 2010. doi:10.1007/978-1-84628-293-5.

19 P. Nicholson. *Space-Efficient Data Structures in the Word-RAM and Bitprobe Models.* University of Waterloo, 2013.

20 R. O. Obe and L. S. Hsu. *PostGIS in Action.* Manning Publications Co., Greenwich, CT, USA, 2nd edition, 2015.

21 D. Papadias and T. Sellis. The semantics of relations in 2D space using representative points: Spatial indexes. In *COSIT'93*, pages 234–247. Springer, 1993. doi:10.1007/3-540-57207-4_16.

22 D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. *KR*, 92:165–176, 1992.

23 J. Renz. Maximal tractable fragments of the region connection calculus: a complete analysis. In *IJCAI'99 Proceedings*, pages 448–454. Morgan Kaufmann Publishers Inc., 1999.

24 H. Samet. Object-based and image-based image representations. *Foundations of Multidimensional and Metric Data Structures*, pages 211–220, 2006. doi:10.1145/1031120.1031123.

25 H. Samet and M. Tamminen. Bintrees, CSG Trees, and Time. *SIGGRAPH Comput. Graph.*, 19(3):121–130, July 1985. doi:10.1145/325165.325211.

26 M. Sioutis, J. Condotta, and M. Koubarakis. An efficient approach for tackling large real world qualitative spatial networks. *International Journal on Artificial Intelligence Tools*, 25(02):1550031, 2016. doi:10.1142/S0218213015500311.

27 S. J. van Schaik and O. de Moor. A memory efficient reachability data structure through bit vector compression. In *SIGMOD'11 Proceedings*, pages 913–924. ACM, 2011. doi:10.1145/1989323.1989419.

Towards a Quantum Theory of Geographic Fields

Thomas Bittner

Departments of Philosophy and Geography, State University of New York, Buffalo, NY, USA
bittner3@buffalo.edu

Abstract

This paper proposes a framework that that allows for the possibility that multiple classically incompatible states are expressed simultaneously at a given point of a geographic field. The admission of such superposition states provides the basis for a new understanding of indeterminacy and ontological vagueness in the geographic world.

1998 ACM Subject Classification J.2 [Physical Sciences and Engineering] Earth and Atmospheric Sciences

Keywords and phrases Vagueness, Quantum Geography, Ontology, Ecoregion classification

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.5

1 Introduction

Classical geography (CG) presupposes that it is possible to identify and to analyze the distribution of geographic qualities on the surface of the Earth in a way that (a) the different kinds of geographic qualities can be analyzed and classified using the Aristotelian method of classification (see below) and (b) the ways in which geographic qualities are instantiated on the surface of the Earth allows for the delineation of regions 'on the ground' at which distinct geographic qualities are instantiated.

The Aristotelian method of classification [2] is based on the assumption that categories/kinds of geographic qualities are structured hierarchically in a tree-like manner. Such trees are called the taxonomic trees or taxonomic hierarchies. Categories farther from the root in the taxonomic hierarchy of geographic qualities are differentiated from categories closer to the root by additional additional more specific qualities. Those additional qualities determine what marks out instances of a more specific category/kind (or species) within the wider parent category (or genus) [2]. Ideally, the Aristotelian method of classification leads classification trees which leaf categories are jointly exhaustive and pairwise disjoint.

In the geographic context there are additional constraints that hold in at least some idealized sense: (i) regions of geographic space at which distinct categories of qualities that are at the same level of the hierarchy tree (e.g., the leafs of the taxonomic tree) are instantiated cannot overlap (ii) jointly the regions with geographic qualities of same level of the taxonomic hierarchy partition the underlying space. That is, the distribution/instantiation of geographic qualities on the surface of the Earth gives rise to geographic fields. At least at sufficiently coarse scales those fields will be smooth and relatively homogeneous.

In non-idealized situations geographic fields have granular features and, in addition, the distribution of geographic qualities displays inhomogeneities. The aspect of granularity is due to the fact that geographic qualities are instantiated at regions of certain scale, i.e., regions that are of some minimal size or larger. Geographic fields are subject to inhomogeneities in the sense that if a quality universal is instantiated at a region of geographic scale then this does not mean that every part of geographic scale of this region is an instance of that quality

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 5; pp. 5:1–5:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

universal. There may be comparatively small (geographic-scale) regions in which different and possibly conflicting quality universals are instantiated.

Consider, for example, the geographic region called ‘Central Great Plains’ at which the land-surface form *Irregular plain* is instantiated. There are comparatively small (geographic-scale) regions in which different land-surface forms are instantiated. Analogously for climate types: There may, for example, exist comparatively small (geographic-scale) regions near larger water bodies where the average temperature (and thus the climate type) is different from the larger surroundings.

All of this seems to indicate that many geographic fields are subject to scale dependency, granularity, and inhomogeneities, all of which seem to be fundamentally vague. It is a fundamental assumption of this paper that an important aspect of understanding geographic fields is to understand the interrelations between scale dependency, granularity, inhomogeneity and the phenomenon of vagueness. This paper aims to contribute to the understanding of those aspects of geographic fields.

In conjunction scale dependency, granularity, inhomogeneities and vagueness of geographic fields seem to lead to a fundamental tradeoff between the classification and the delineation of geographic fields. Consider Fig. 1 in which two ways of delineating the contiguous USA into ecoregions of sub-regional scale are (partly) displayed. The bold black lines depict the delineation formed by the ecoregion sections identified by Bailey [1]. The non-bold black lines depict the delineation of formed by the collection level three regions identified by the EPA [3]. There is an obvious difference between the generalized and coarse character of Bailey’s delineation (bold black boundaries) and the more fine-grainedness of the EPA delineation (non-bold black boundaries).

The smooth and highly generalized boundaries of Bailey’s delineation seem to convey the intuition that there is a large degree of ‘freedom’ to place boundaries by fiat in the way that supports best the purpose of a map as a medium for conveying of information. As [11] puts it, to minimize information decoding error, map designers strive (i) for crisp (non-graduated) boundaries and (ii) for minimizing boundary complexity by drawing boundaries in highly generalized ways. Drawing boundaries in highly generalized ways avoids the misinterpretation of the delineation as realism of a map. By contrast, the non-bold black boundaries are the result of observations ‘on the ground’ that are aimed at identifying local variations of qualities and there by identifying boundaries that separate ecoregions of different kinds [7]. The fine-grainedness of the boundaries conveys the preciseness of the delineation.

Both approaches to identifying ecoregions on the surface of the Earth include classification and delineation operations. Surprisingly, the outcomes of both approaches are very different as can be seen in the maps displayed in Fig. 1.

► **Hypothesis 1.** *The reason for different outcomes of the operations of classification and delineation can be attributed to the fact that the sequence of the application of the operation of classification and delineation is significant. More precisely, Bailey applies the classification operation first and then the delineation on the ground second by contrast, the EPA applies the operation of delineation first and then the operation of classifying the delineated regions second. In technical terms this is to say that the different maps produced by Bailey and the EPA are an indication that the operations of classification and delineation do not commute. The non-commutativity of operations that act on field-like phenomena that are subject to granularity, inhomogeneity, and vagueness is a fundamental aspect of many types of geographic fields.*

If this hypothesis is true then it is a fundamental criterium for the adequacy of a theory of geographic fields that it is able to give a satisfying explanation for the non-commutativity and the tradeoff between classification and delineation.

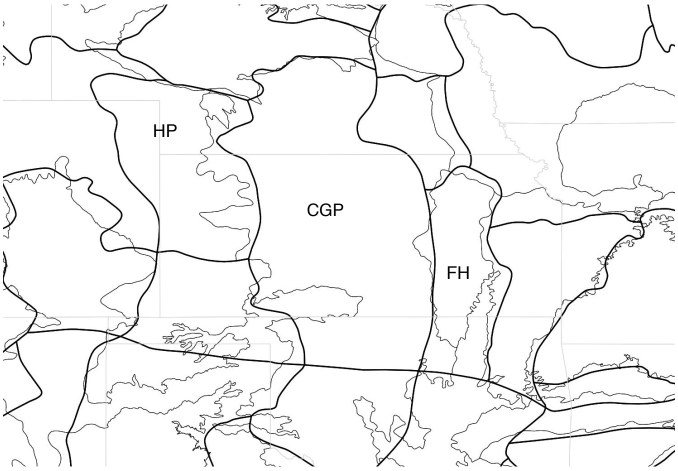


Figure 1 Classification and delineation of the central US into ecoregions [4] according to Bailey [1] (bold boundaries) and [3] (non-bold boundaries).

The non-commutative nature of certain operations is the hallmark of Quantum Mechanics [6]. For example, Heisenberg's uncertainty principle is a consequence of the non-commutativity of operators that determine the position and the momentum of a particle [9]. It is the aim of this paper to present an adequate theory of geographic fields by applying ideas and techniques from quantum mechanics to geographic fields. This set of ideas and techniques applied to geographic phenomena will be called *Quantum Geography* (QG).

The idea of exploring the quantum nature of the geographic world has been discussed previously [13, 5]. This paper goes beyond those discussions by showing how a quantum theory of geographic fields actually could look like in the specific context of Ecoregion classification and delineation. The author feels that in this more specific context it will be easier to go beyond mostly philosophical discussions towards developing scientific theories which predictions can be tested empirically.

2 (Geographic) fields as fiber bundles

In what follows the language of fiber bundles (e.g., [10, 15]) is used to express a theory of geographic fields. The reasons for this choice are as follows:

Firstly, geographic fields can have the form of scalar fields (temperature, elevation, etc.) and vector fields (air flow, hydrological flow, etc.), co-vector (gradient) fields (rates of changes of scalar as well as vector fields) such as slope fields (direction of largest rate of changes of elevation at every point), temperature gradient fields (direction of largest change of temperature at every point), etc. Fiber bundles are general enough to include scalar as well as vector and co-vector fields. Secondly, formalizations of geographic fields need to be able to integrate multitudes of local descriptions of field phenomena into a global framework. For example, fiber bundles provide powerful means to deal with globally curved spaces using locally flat reference systems. Thirdly, a scientific theory of geographic fields must be able to talk about the class of all geographically possible fields in an efficient manner. Fiber bundles provide means to formally characterize what it means for a field to be geographically possible. Finally, 'Classical' descriptions of geographic fields based on fiber bundles can be naturally generalized to the descriptions of geographic fields within the framework of quantum geography. In this respect the paper will mostly follow [10].

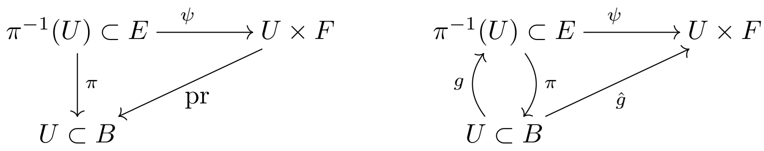


Figure 2 Fibers of a fiber bundle (left); (local) section of a fiber bundle (right).

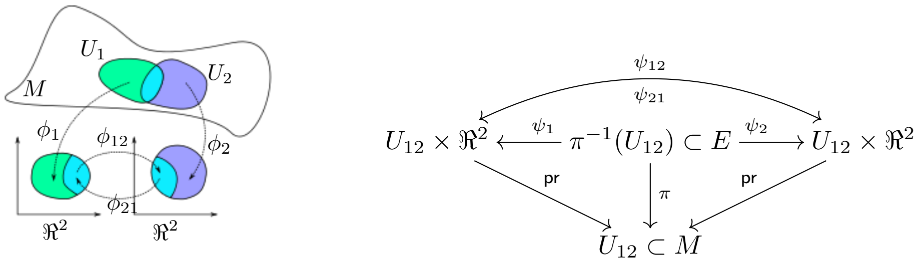


Figure 3 Bundle atlas (left: Image licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons).

2.1 Fiber bundles

The literature on fiber bundles is vast. In this subsection some basic definitions have been collected from [10, 15]. A fiber bundle is a structure (E, B, π, F), where E, B, and F are topological spaces and $\pi : E \to B$ is a continuous surjection satisfying a local triviality condition: for every $e \in E$, there is an open (trivializing) neighborhood $U \subset B$ of $\pi(e)$ such that there is a homeomorphism $\psi : \pi^{-1}(U) \to (U \times F)$ such that the diagram in the left of Figure 2 commutes.

The space B is the base space of the bundle, E is the total space, and F is the abstract fiber. The map π is the bundle projection; B is assumed to be topologically connected; $(U \times F)$ is a product space; $\mathrm{pr} : U \times F \to U$ is the natural projection; and the trivialization map is $\psi : \pi^{-1}(U) \to U \times F$. In a fiber bundle (E, B, π, F) every fiber $\pi^{-1}(x) \in E$ over $x \in B$ is homomorphic to some abstract fiber F. In particular ψ is defined as:

$$\psi : e \in E \mapsto (\pi(e), f(e)) \in U \times F \quad \text{with } f : \pi^{-1}(x) \to F \tag{1}$$

such that f is a homomorphism.

An *open covering* of a fiber bundle (E, B, π, F) is a system $\{U_\alpha\}$ of open subsets of B together with a trivialization maps $\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F$. The system $\psi = \{(U_\alpha, \psi_\alpha)\}$ is a *bundle atlas*. If $\{(U_1, \psi_1), (U_2, \psi_2), \ldots\}$ is an atlas then the trivializations for overlapping members $U_{12} = U_1 \cap U_2 \neq \emptyset$ of the covering are compatible such that for all $e \in \pi^{-1}(U_{12})$: $\psi_{12} \circ \psi_1 = \psi_2^{-1}$ as illustrated in Fig. 3.

A *local section* of (E, B, π, F) is a continuous map $g : U \to E$ where U is an open set in B and $\pi(g(x)) = x$ for all $x \in U$. For a local section g there is a map $\hat{g} : U \to (U \times F)$ such that $\hat{g} = \psi \circ g$. This is displayed in the right of Fig. 2. If (U, ψ) is a local trivialization chart then local sections always exist over U. Given an atlas, local sections can be combined to cover the base space of the fiber bundle as a whole – *covering sections*. The set of all covering sections over the fiber bundle (E, B, π, F) is denoted by $\mathsf{Sect}(E)$.

2.2 Fiber bundles, determinable and determinate qualities

Fiber bundles provide formal means to represent smoothly distributed qualities (quality fields). They also very naturally incorporate the ontological distinctions between quality determinables, quality determinates [14]. Intuitively, a quality field as a whole corresponds to a quality determinable such as energy, temperature, ecoregion domain, etc. At given points of the base space then quality determinates such as 10 Joule, 72 degree Fahrenheit, Dry domain, etc. are instantiated.

Let (E, B, π, F) be a fiber bundle over the manifold B. The set of all possible instances of the universal `field of type` $\mathcal{E}$ `over` B is the set of all covering sections $\mathsf{Sect}(E)$ of (E, B, π, F) such that

- The fiber $\pi^{-1}(x) \subset E$ is the class of quality determinates that fall under the quality determinable $\mathcal{E}$ and that can possibly be instantiated in the neighborhood of $x \in B$.
- If g is a covering section of the base space B then there is a field that is geographically possible and on this possibility it holds that for all $x \in B : \ g(x) = e$ iff the quality determinate $e \in \pi^{-1}(x)$ is instantiated in the neighborhood of $x \in B$. Neighborhood in this context does not mean infinitesimal neighborhood but neighborhood in the geographic sense. Such neighborhoods are specified in the context of the local trivializations.1
- The local representation of the possible $\mathcal{E}$-field g is the function $\hat{g} : U \to (U \times F)$ as depicted in Fig. 2 (right). Here $F = F' \times \Gamma \times \Delta$ where F' is a representation of the quality determinate instantiated in the neighborhood of a given location $x \in U \subseteq B$. Γ is the minimal diameter of a region at which an instance of this quality determinate can be instantiated. Δ is the distance from x at which F' is actually instantiated. (This will be important to capture the possibility of inhomogeneities.)

That is, every $\mathcal{E}$-field-universal in conjunction with its possible instantiating fields (possible $\mathcal{E}$-field-particulars) can be thought of as having the structure of a fiber bundle (E, B, π, F). Possible $\mathcal{E}$-field-particulars correspond to (covering) sections of the underlying fiber bundle. A $\mathcal{E}$-field-particular is such that in the neighborhood of every location $x \in U$ exactly one of the quality determinates of $\pi^{-1}(x)$ is instantiated in a way that is consistent with the constraints that reflect the granular and locally inhomogeneous nature of geographic fields. This will be discussed in the next section.

3 Geographic fields in Quantum geography

It is fundamental to classical geography that geographically possible fields are sections in fiber bundles that take every point in the base space to a member of a set of disjoint possibilities that constitutes the fiber over that base point. By contrast, in *Quantum geography* (QG) the state of a geographic field at a given point of the base space of is not just a point in the space of disjoint possible states collected in the fiber over that point.

3.1 Superposition states

In QG the state of a geographic field at a given point of the base space is represented by a vector which is the sum of the base vectors that span the vector space that forms the fiber over that base point.

▶ **Example 1.** Consider Fig. 4. The quality determinable *Ecoregion domain* has four subclasses as determinates. In classical geography exactly one of those of those four possibilities can be instated at (in the neighborhood of) a given point. In the language of fiber bundles the class of all possible fields of type *Ecoregion domains* are sections in a fiber bundle which

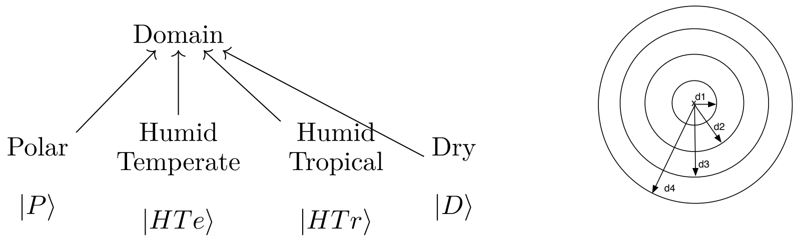


Figure 4 Quality base $\mathcal{Q} = \{|P\rangle, |HTe\rangle, |HTr\rangle, |D\rangle\}$ (left); Localization base $\mathcal{X} = \{|\Delta_1\rangle, |\Delta_2\rangle, |\Delta_3\rangle, |\Delta_4\rangle\}$ (right).

fibers over every point are constituted by the set { Polar, Humid Temperate, Humid Tropical, Dry }.

What are points in the space of possibilities (the fibers) in classical geography are dimensions (base vectors) in a *vector space* of possible states in QG. The fibers of fields of type *Ecoregion domains* in QG form a four-dimensional (complex) vector space with base vectors labeled Polar, Humid Temperate, Humid Tropical, and Dry. In QG a geographic field at a given point of the base space can be in a *superposition* of multiple classically incompatible states. A field of type *Ecoregion domains* at a given location x can be in the state $\sqrt{0.1}$ Polar + $\sqrt{0.4}$ Humid Temperate + $\sqrt{0.3}$ Humid Tropical + $\sqrt{0.2}$ Dry.

▶ **Postulate 1.** *Geographic fields in QG are sections of* vector bundles*. Every fiber has the structure of a vector space. The bases (dimensions) of the vector space of each fiber correspond to what in CG are the points in the space of possible states. In QG states of geographic fields at given points in the base space include superpositions (vector sums) of what in CG are distinct states.*

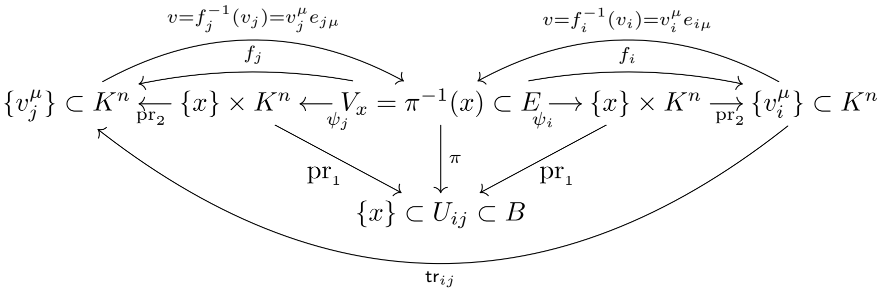
The admission of superposition states constitutes a major departure from CG because it allows for the possibility that multiple classically incompatible states are expressed simultaneously at a given point of a geographic field. In what follows the existence of superposition states will provide the basis for a new understanding indeterminacy and ontological vagueness in the geographic world. To develop the links between ontological vagueness and superposition states some more technical apparatus about vector bundles is needed.

3.2 Vector bundles

A vector bundle (E, M, π, K^n) has fibers with the structure of vector spaces. Local trivializations are of the form $U_{ij} \times K^n$ where K^n is assumed to be R^n or C^n. Consider Diag. 2. If $V_x =_{df} \pi^{-1}(x)$ is a concrete fiber (i.e., an internal space) over M with vector elements v, the linear map $f_i : \pi^{-1}(x) \to K^n$ is equivalent to choosing a base $e_{i\mu}$ and to express the vectors $v \in V_x$ as components with respect to the base $e_{i\mu}$. That is, $f_i(v) = v_i^\mu$, such that

$$v_i^0 e_{i0} + \ldots + v_i^{n-1} e_{in-1} \;\equiv\; \sum_{\mu=0}^{n-1} v_i^\mu e_{i\mu} \;\equiv\; v_i^\mu e_{i\mu} = v \in V_x.$$

Usually, there are a multitude of possible bases for a given vector space. Thus it makes sense to transform the representation of a vector space in one base to a representation of the same space in another base. A change of basis can be defined as $v_j e_j = (v_i e_i)\mathsf{tr}_{ij}$ where tr_{ij} is a linear map that transforms the coordinate v_i in the base e_i to the coordinate v_j in the base e_j.



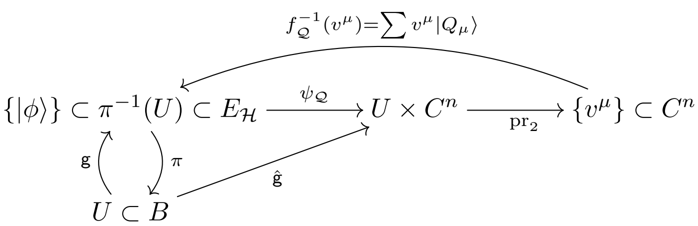
(2)

3.3 Hilbert bundles

Like quantum mechanics, quantum geography requires that the vector spaces that constitute the fibers of the vector bundles are Hilbert spaces [10]. A Hilbert space $\mathcal{H}$ is a *complex* vector space with an *inner product*. In what follows Dirac's notation for vectors in Hilbert spaces [6] is used. The members of a Hilbert space $\mathcal{H}$ are written as ket vectors of the form $|\phi\rangle$ where ϕ is a name/label. As vector spaces Hilbert spaces are closed under vector addition and scalar multiplication. That is if $|\phi\rangle, |\psi\rangle \in \mathcal{H}$ then $\alpha|\phi\rangle + \beta|\psi\rangle \in \mathcal{H}$, where α and β are complex numbers that modify the length of a vector via scalar multiplication and $+$ is the vector addition. The inner product $\langle\psi|\phi\rangle$ of the vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ (defined below) is a complex number.

A base $\mathcal{Q} = |Q_1\rangle, \ldots, |Q_n\rangle$ of a n-dimensional Hilbert space $\mathcal{H}$ is a system of vectors such that every member of $\mathcal{H}$ can be expressed as a vector sum of the base vectors. A base is orthonormal if the inner product of distinct base vectors is zero and all base vectors are of unit length, i.e., $\langle Q_i|Q_j\rangle = 1$ if $i = j$ and $\langle Q_i|Q_j\rangle = 0$ otherwise. If the vector $|\phi\rangle = \alpha_1|Q_1\rangle + \ldots + \alpha_n|Q_n\rangle$ then there exists a dual vector $\langle\phi| = \overline{\alpha}_1|Q_1\rangle + \ldots + \overline{\alpha}_n|Q_n\rangle$ where $\overline{\alpha}_i$ is the complex conjugate of α_i. Thus if $|\phi\rangle = \alpha_1|Q_1\rangle + \ldots + \alpha_n|Q_n\rangle$ and $|\psi\rangle = \beta_1|Q_1\rangle + \ldots + \beta_n|Q_n\rangle$ then the inner product of $|\phi\rangle$ and $|\psi\rangle$ designated by $\langle\psi|\phi\rangle$ is the sum of the products of the components of $\langle\psi|$ and $|\phi\rangle$ computed as $\sum_i \overline{\beta_i}\alpha_i$. In what follows $|\overline{q}_i q_i|$ is an abbreviation for the squared modulus of the scalar product $|\langle Q_i|Q_i\rangle|^2$.[1]

A *Hilbert bundle* is a vector bundle $(E_{\mathcal{H}}, B, \pi, C^n)$ with base space B, bundle space $E_{\mathcal{H}}$ and abstract fiber $\mathcal{H}_{\mathcal{Q}}$ – the Hilbert space $\mathcal{H}$ in the $\mathcal{Q}$-base. A section of a Hilbert bundle $(E_{\mathcal{H}}, B, \pi, C^n)$ is a mapping of signature $U \subset B \to U \times C^n$.



(3)

► **Postulate 2.** *The representation of a geographic field* g *in the* $\mathcal{Q}$*-base of a Hilbert space* $\mathcal{H}$ *is a smooth section* $\mathsf{g} \in \mathsf{Sect}(E^{\mathcal{Q}}_{\mathcal{H}})$ *of a Hilbert bundle* $(E_{\mathcal{H}}, B, \pi, C^n)$*. The section* $\hat{\mathsf{g}}$ *is a vector field such that for every* $x \in B$*,* $\hat{\mathsf{g}}(x) \in \mathcal{H}_x$ *is a vector of unit length of expressed in the base* $\mathcal{Q}$*, i.e.,* $\hat{\mathsf{g}} = \{x \in U \mapsto \sum \alpha_i |Q_i\rangle \in \mathcal{H}_x \mid \sum_i |\overline{\alpha}_i\alpha_i| = 1\}$.

[1] Details can be found in any text book on quantum mechanics. The classic reference is [6].

The constraint $\sum_i |\overline{\alpha_i}\alpha_i| = 1$ ensures that possible states of geographic fields are such that the contributions of all the orthogonal possibilities captured in the system of base vectors $\mathcal{Q}$ jointly add up to 1.

▶ Remark. There are more restrictions needed to ensure that the Hilbert bundles that are intended to represent geographic fields are 'well behaved' in the sense that the Hilbert spaces at neighboring points are compatible, that the section representing the geographic fields are smooth, and others more. In particular the notion of 'connection' [10] is needed to compare vectors in different fibers of the fiber bundle. This goes beyond the scope of this paper. A very good discussion of many relevant aspects can be found in [10]. Whether or not those requirements are necessary/sufficient in the context of geographic fields is still an open question.

▶ **Definition 2.** Consider a geographic field $\hat{\mathsf{g}}$ in the $\mathcal{Q}$-base: $\hat{\mathsf{g}}$ is maximally determinate with respect to the $\mathcal{Q}$-base at x iff $\hat{\mathsf{g}}(x) = 1\,|Q_i\rangle$ for the i-th base vector and $0\,|Q_j\rangle$ for $i \neq j$. The field $\hat{\mathsf{g}}$ is minimally determinate with respect to the $\mathcal{Q}$-base at x iff $\hat{\mathsf{g}}(x) = \frac{1}{\sqrt{n}}\,|Q_1\rangle + \ldots + \frac{1}{\sqrt{n}}\,|Q_n\rangle$.

4 Quality base vs. localization base

Given the structure of the Hilbert spaces that form the fibers of a Hilbert bundle, every geographic field $\mathsf{g} \in \mathsf{Sect}(E_{\mathcal{H}})$ can be expressed in (at least) two *complimentary* systems of base vectors: the quality base and the localization base. Roughly, when g is expressed in the quality base, then the field is a map of signature $\hat{\mathsf{g}} : U \to C^n_{\mathcal{Q}}$ taking locations of the base space to superpositions of qualities that are represented by the n-tuples $C^n_{\mathcal{Q}}$. By contrast, if g is expressed in the localization base, then the field is a map of signature $\hat{\mathsf{g}} : U \to C^n_{\mathcal{X}}$ taking locations of the base space to superpositions of possible deviations from x that are represented by the n-tuples $C^n_{\mathcal{X}}$. That is, in the quality base $\hat{\mathsf{g}}$ maps locations of the base space to information about quality pattern while in the localization base $\hat{\mathsf{g}}$ maps locations of the base space to information about the (metric) closeness to which the information about the quality pattern contained in $\mathsf{g}(x)$ is linked to the location x of the base space (details in Sec. 4.2). In what follows the term 'deviation' is used to describe the distance between the point x to which the quality q is attributed by the field to the closest point at which q is actually expressed. To say that the quality and localization bases are complimentary is to say that if the information expressed in the quality base is maximally determinate then the information expressed in the localization base is minimally determinate and vice versa.

▶ Remark. Examples 3 and 4 below will illustrate that, when compared with the quantum mechanics of a free particle, the quality of a geographic field at a given position in the base space of the underlying fiber bundle is like the position of a free particle. The (degree of) localization a geographic field at that position in the base space is like the momentum of a free particle. Usually the quantum mechanics of a free particle allows for a continuum of possible positions and momenta. For the purpose of this paper it will be sufficient to consider discrete quality spaces and finitely many possible distinct qualities. Roughly, to go from the discrete to the continuous case is to replace sums by integrals.

QG allows for geographic fields that are in indeterminate states and thereby provides means for the expression of ontological vagueness. According to QG ontological vagueness has two interrelated aspects: quality indeterminacy and localization indeterminacy.

4.1 Geographic fields in the quality base

Let $Q_1, \ldots, Q_n$ be quality determinates that are pairwise disjoint and jointly exhaust some quality determinable ξ. A geographic ξ-field is a section in a Hilbert bundle (E_H, B, π, C^n) such that fibers $\pi^{-1}(x) = \mathcal{H}_x$ over each point $x \in B$ have the structure of a Hilbert space. The vectors in this space are expressed in the quality base $\mathcal{Q} = |Q_1\rangle, \ldots, |Q_n\rangle$ of the abstract fiber of the associated local trivialization. The base vectors $|Q_1\rangle, \ldots, |Q_n\rangle$ could, for example, be qualities at the same level of a universal hierarchy identified by the Aristotelian method of classification. Consider the left of Fig. 4. The geographic field associated with the quality determinable *Ecoregion Domain* has at every point of the base space a Hilbert space $\mathcal{H}_x^D$ which vectors can be expressed in the quality base $\mathcal{Q} = \{|P\rangle, |HTe\rangle, |HTr\rangle, |D\rangle\}$.

▶ **Postulate 3.** *A ξ-field of the form $x \in B \mapsto q_1 |Q_1\rangle + \ldots + q_n |Q_n\rangle \in \mathcal{H}_x$ is interpreted as: The quality Q_i is expressed in the neighborhood of x to the degree $|\overline{q}_i q_i|$ in a way such that $\sum_j |\overline{q}_j q_j| = 1$.*

As pointed out above, in QG it is possible that in the neighborhood of a given point x a combination of incompatible qualities are expressed at any given time. This captures the aspect of quality indeterminacy of the underlying ontological vagueness. For example, for the geographic field associated with the quality determinable *Ecoregion Domain* the maximally indeterminate state in the quality base at a given point of the base space is $\frac{1}{\sqrt{4}}(|P\rangle + |HTe\rangle + |HTr\rangle + |D\rangle)$. By contrast, maximally determinate states in the quality base include the state $1\,|P\rangle + 0\,|HTe\rangle + 0\,|HTr\rangle + 0\,|D\rangle$. A state of intermediate indeterminacy is $\sqrt{0.1}\,|P\rangle + \sqrt{0.4}\,|HTe\rangle + \sqrt{0.3}\,|HTR\rangle + \sqrt{0.2}\,|D\rangle$.

Classification

Classification in QG, is the assignment of a determinate classification value to all the locations of the base space of a geographic field. That is, classification is an operation that takes as input a geographic field $\hat{\mathfrak{g}}$ that, when represented in the quality base, is in a state in which superpositions of classically contradicting qualities are expressed in the neighborhoods of the points in the base space. The classification operation then maps $\hat{\mathfrak{g}}$ to a field $\hat{\mathfrak{g}}'$ in the same fiber bundle in which exactly one of the possible qualities is exclusively expressed in the respective neighborhoods.

▶ **Postulate 4.** *A classification operator $\hat{\mathcal{C}}_x$ is an operator on the Hilbert space $\mathcal{H}_x^{\mathcal{Q}}$ with the following properties: (a) $\hat{\mathcal{C}}_x$ is a self-adjoint[2] operator on $\mathcal{H}_x^{\mathcal{Q}}$; (b) the set of base vectors of the quality base $\mathcal{Q} = |Q_1\rangle, \ldots, |Q_n\rangle \subset \mathcal{H}_x^{\mathcal{Q}}$ are eigenvectors of $\hat{\mathcal{C}}_x$ such that $\hat{\mathcal{C}}\,|Q_i\rangle = Q_i\,|Q_i\rangle$ where Q_i is a non-complex number; and (c) $\hat{\mathcal{C}}_x = |\phi\rangle \in \mathcal{H}_x^{\mathcal{Q}} \mapsto |Q_i\rangle \in \mathcal{Q}$ for some i.*

Here Q_i is a classification value and $|Q_i\rangle$ is the state in which the quality determinate that is represented by the number Q_i is exclusively expressed in the neighborhood of x associated with the underlying Hilbert space $\mathcal{H}_x^{\mathcal{Q}}$. That is, the quality values that a field in a determinate state can possibly have at a given location of the base space are given by the eigenvalues of the classification operator. The base vectors of the Hilbert spaces in the quality base are the eigenvectors of the classification operator. If a geographic field is in an eigenstate at

[2] A self-adjoint operator on a complex vector space $\mathcal{H}$ with inner product $\langle\cdot|\cdot\rangle$ is a linear map $\hat{A}$ from $\mathcal{H}$ to itself with a unique corresponding operator $\hat{A}^\dagger$ such that: $(\langle\phi|\,\hat{A}^\dagger)\,|\psi\rangle = \langle\phi|\,(\hat{A}\,|\psi\rangle)$ for all $|\phi\rangle, |\psi\rangle \in \mathcal{H}$. If $\hat{A}$ is represented by a square matrix with complex values, then $\hat{A}^\dagger$ is the matrix obtained from $\hat{A}$ by complex conjugation and transposition.

a given location x of the base space then the field value at this point is the corresponding eigenvalue. According to the classification the quality corresponding to this eigenvalue is exclusively expressed in the neighborhood assigned to x.

▶ **Example 3.** Consider[3] the Hilbert space $\mathcal{H}_y^{\mathcal{X}}$ with the base vectors $|P\rangle, |HTe\rangle, |HTr\rangle, |D\rangle$. In analogy to the position base in QM let $X = \{0, 1, 2, 3\}$ be the set of possible qualities (possible positions in QM) such that 0 stands for Polar Domain, 1 stands for Humid Temperate Domain, 2 stands for Humid Tropical Domain, and 3 stands for Dry Domain. Let the base vectors be functions of the form $|P\rangle : X \to \{0, 1\}$ where $|P\rangle \equiv \lambda x.\ x = 0$, $|HTe\rangle \equiv \lambda x.\ x = 1$, $|HTr\rangle \equiv \lambda x.\ x = 2$, and $|D\rangle \equiv \lambda x.\ x = 3$. For example, $|D\rangle : X \to \{0, 1\}$ is a function that yields 1 if the value of its argument is 3 ($3 = 3$ is true) and 0 otherwise (e.g., $3 = 2$ is false). Functions of this kind form a Hilbert space the members of which are all the functions that can be formed by adding the complex multiples of the functions that save as base vectors. In analogy to the position operator in QM, the classification operator $\mathcal{C}$ is defined as $(\mathcal{C}\ \ |\phi\rangle)$ the multiplication of the state vector $|\phi\rangle$ with the location x at which the operator $\mathcal{C}$ is evaluated. If the operator is evaluated at position x of the underlying quality space then one has $(\mathcal{C}\ \ |\phi\rangle)x \equiv (x\,|\phi\rangle)x$. If x takes its values from the set $X = \{0, 1, 2, 3\}$ then the eigenvalues are $0, 1, 2$ and 3 such that 0 stands for Polar Domain, 1 stands for Humid Temperate Domain, 2 stands for Humid Tropical Domain, and 3 stands for Dry Domain.

On this view, classification is an operation that takes as input a geographic field that is subject to indeterminacy and maps it to a different field that is not subject to indeterminacy. That is, the classification operator $\hat{\mathcal{C}}$ collapses a superposition state $|\phi\rangle$ into one of its eigenstates $|Q_i\rangle$. According to Quantum Mechanics the collapse of a superposition state into an eigenstate as the result of a measurement (in the widest sense) is, on the standard (the 'Kopenhagen') interpretation, inherently indeterministic in nature (e.g., [12]). In QM this indeterminacy is expressed in a probabilistic way. If one were to follow this view in QG one would have:

▶ **Hypothesis 2.** *If $|\phi\rangle = q_1\,|Q_1\rangle + \ldots + q_n\,|Q_n\rangle \in \mathcal{H}_x^{\mathcal{Q}}$ is the state of a geographic field in the neighborhood of x then the probability that the neighborhood of x is classified as Q_i is $|\overline{q}_i q_i|$.*

That is, the more dominant the expression of Q_i in the neighborhood of x the more likely it is that this neighborhood is classified as having the quality Q_i.

Unlike measurement in Physics, classification in geography seems to be more the result of the deliberative actions of cognitive agents that include a certain degree of fiat. To understand the deliberative action of a cognitive agent as a random process may not be appropriate. An alternative way of interpreting the collapse a superposition state $|\phi\rangle$ into one of its eigenstates $|Q_i\rangle$ could be:

▶ **Hypothesis 3.** *If $|\phi\rangle = q_1\,|Q_1\rangle + \ldots + q_n\,|Q_n\rangle \in \mathcal{H}_x^{\mathcal{Q}}$ is the state of a geographic field in the neighborhood of x then this neighborhood is classified as Q_i if a maximum $|\overline{q}_i q_i| = \max\{|\overline{q}_j q_j| \mid 1 \le j \le n\}$ exists.*

On this interpretation one assumes that a cognitive agent is able to perceive the degrees to which certain qualities are expressed at a certain location of a geographic field. One also

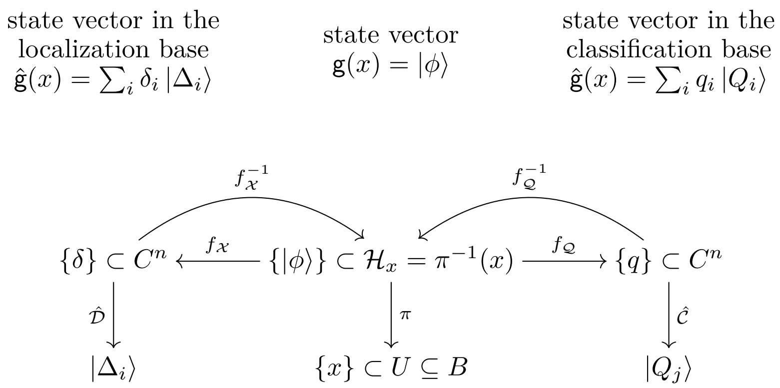
[3] This example in conjunction with Example 4 is intended to illustrate the ways in which the math of QM/QG with their operators, eigenvectors and eigenvalues is designed to achieve a formalism with properties that (at least in the case of QM) yields surprisingly accurate predictions. Details can be found in any introductory textbook on QM (e.g., [8]).

allows for the possibility that the cognitive agent is unable or unwilling to make a judgement when there is no unique maximum.

Whether or not the first or the second hypothesis or neither of them is actually true of the geographic world is in the opinion of the author an empirical question and can (at least in principle) be determined by experiments.

4.2 Geographic fields in the localization base

In QG, a state vector $|\phi\rangle \in \mathcal{H}_x$, which in the quality base takes the form $q_1 |Q_1\rangle + \ldots + q_n |Q_n\rangle$, not only encodes information about the expression of geographic qualities $Q_1, \ldots, Q_n$ in the neighborhood of the position x in the base space. The state vector $|\phi\rangle$ also encodes information about possible inhomogeneities that may affect the expression of these qualities in the neighborhood of x. The information about possible inhomogeneities of the field g at x that is encoded in the state vector $|\phi\rangle \in \mathcal{H}_x$ is accessible when $|\phi\rangle$ is expressed in the localization base. The localization base $\mathcal{X}$ of $\mathcal{H}_x^{\mathcal{X}}$ is a set $\mathcal{X} = |\Delta_1\rangle, \ldots, |\Delta_n\rangle$ such that every field state $|\phi\rangle \in \mathcal{H}_x$ can be expressed in this base as $|\phi\rangle = \delta_1 |\Delta_1\rangle + \ldots + \delta_n |\Delta_n\rangle \in \mathcal{H}_x^{\mathcal{X}}$ with the additional normalization constraint $\sum_i |\overline{\delta}_i \delta_i| = 1$. The idea that the state $|\phi\rangle$ of a geographic field g at a given location $x \in B$ can be described in the quality base $\mathcal{Q}$ as well as the localization base $\mathcal{X}$ is visualized in Diag 4.



(4)

Intuitively, the base vectors $|\Delta_i\rangle$ can be thought of as the various degrees of inhomogeneity that are possible for the field at a given location of the base space. That is, the localization base vectors $|\Delta_i\rangle$ represent states of the field g at $x \in U \subseteq B$ where possible ranges of deviation from x due to inhomogeneities of the field g at that location corresponds to a collection of nested rings that are centered at x (Fig. 4 right). In general the state $|\phi\rangle = \sum_i \delta_i |\Delta_i\rangle$ will be a superposition state that is subject to the constraint $\sum_i |\overline{\delta}_i \delta_i| = 1$. That is, to the degree quantified by the value of the expression $|\overline{\delta}_i \delta_i|$ all of the possible ranges of deviation Δ_i are realized. These superpositions are expressions of the indeterminacy associated with the underlying ontological vagueness of the inhomogeneities of geographic fields. In the localization base a section of the fiber bundle is a mapping of the form

$$\hat{\mathsf{g}} : x \in U \subset B \mapsto \delta_1 |\Delta_1\rangle + \ldots + \delta_n |\Delta_n\rangle \in \mathcal{H}_x^{\mathcal{X}}.$$

Delineation

Delineation in Quantum Geography, is the assignment of determinate localization information to all the points of the base space of a geographic field. That is, delineation is an operation that takes a geographic field $\hat{\mathsf{g}}$ that is represented in the localization base and which is in superposition states at all (or many) points in the base space. The delineation operator $\hat{\mathcal{D}}$ maps the field $\hat{\mathrm{g}}$ to a field $\hat{\mathrm{g}}''$ in the same fiber bundle. At all positions of the base space the field $\hat{\mathrm{g}}$ is in in a state that corresponds to one of the members of the localization base.

▶ **Postulate 5.** *If $\hat{\mathcal{D}}_x$ is a delineation operator on the Hilbert space $\mathcal{H}_x^{\mathcal{X}}$ then: (a) $\hat{\mathcal{D}}$ is a self-adjoint operator on $\mathcal{H}_x^{\mathcal{X}}$; (b) the vectors of the localization base $\mathcal{X} = |\Delta_1\rangle, \ldots, |\Delta_n\rangle$ are eigenvectors of the operator $\hat{\mathcal{D}}$ such that $\hat{\mathcal{D}}\,|\Delta\rangle_i = d_i\,|\Delta\rangle_i$ for the eigenvalues d_i; and (c) $\hat{\mathcal{D}} = |\phi\rangle \in \mathcal{H}_x^{\mathcal{X}} \mapsto |\Delta_i\rangle \in \mathcal{X}$.*

As depicted in Fig. 4 (right), d_i is a distance range from $x \in U$ and $|\Delta_i\rangle$ is the state in which the qualities associated with $|\Delta_i\rangle$ when expressed in the quality base is definitively expressed within the range of distances from $x \in U$ that are associated with d_i. The possible deviation from the base point is due to inhomogeneities of the underlying geographic field. Numerically possible deviation correspond to the eigenvalues d_i of the delineation operator $\hat{\mathcal{D}}$. The base vectors $|\Delta_i\rangle$ of the Hilbert spaces in the localization base are the eigenvectors of $\hat{\mathcal{D}}$. If a geographic field is in an localization eigenstate at a given location x in the base space then the field value expressed in the localization base at this point will be the eigenvalue corresponding to this state.

▶ **Example 4.** Consider the Hilbert space $\mathcal{H}_y^{\mathcal{X}}$ at the point $y \in B$ of the field's base space. Assume that the $\mathcal{H}_y^{\mathcal{X}}$ is expressed in the localization base formed by the vectors $|\Delta_1\rangle, |\Delta_2\rangle, |\Delta_3\rangle, |\Delta_4\rangle$. In analogy to the definition of the momentum eigenstates in QM these vectors have the form $|\Delta_k\rangle \equiv \lambda x.\frac{1}{\sqrt{2\pi\Omega}}e^{\frac{i}{\Omega}d_k x}$ for $1 \leq k \leq 4$ where $i = \sqrt{-1}$ the imaginary unit and x ranges over the possible locations in quality space, i.e., the eigenvalues of the classification operator of Example 3. The constant Ω is a scale factor and plays the role of Plank's constant in QM (more on Ω in Sec. 5). Functions of this kind form a Hilbert space analogous to Example 3. The delineation operator is defined in analogy to the momentum operator in QM as: $(\mathcal{D}\ |\phi\rangle) \equiv (-i\frac{\partial}{\partial x}\,|\phi\rangle)$. The eigenvalues are given by: $\mathcal{D}\,|\Delta_k\rangle = -i\frac{\partial}{\partial x}\frac{1}{\sqrt{2\pi\Omega}}e^{\frac{i}{\Omega}d_k x} = \frac{d_k}{\Omega\sqrt{2\pi\Omega}}e^{\frac{i}{\Omega}d_k x} = \frac{d_k}{\Omega}\,|\Delta_k\rangle$. On the intended interpretation the eigenvalues d_k label the concentric circles in Fig. 4 (right) that indicate possible degrees of deviation due to the inhomogeneities in the underlying field.

Similarly to the collapse of a superposition state in the quality base, the collapse of a superposition state $|\phi\rangle$ into an eigenstate $|\Delta_i\rangle \in \mathcal{X}$ can be understood in a probabilistic way as follows:

▶ **Hypothesis 4.** *If $|\phi\rangle = \Sigma_i\ \delta_i\,|\Delta_i\rangle$ is the state of a geographic field in the neighborhood of x then the probability that the localization deviation associated with $|\phi\rangle$ has the value d_i is $|\overline{\delta}_i\delta_i|$.*

On this hypothesis the value $|\overline{\delta}_i\delta_i|$ specifies the probability that the quality pattern associated with the state $|\phi\rangle$ that is attributed to the point $x \in B$ can actually deviate from x by the distance range corresponding to d_i.

Similar to classification, delineation in geography seems to be more the result of the deliberative actions of cognitive agents that include a certain degree of fiat. The degree of fiat was particularly persuasive in the way in which the boundaries in Fig. 1 were drawn as the result of delineation operations. But, again, despite the indeterminacy there does not seem to be a random process at the heart of the collapse the superposition states. An alternative way of interpreting the collapse a superposition state $|\phi\rangle$ into one of its eigenstates $|\Delta_i\rangle$ could be the following:

▶ **Hypothesis 5.** *If* $|\phi\rangle = \Sigma_i\ \delta_i\,|\Delta_i\rangle$ *is the state of a geographic field in the neighborhood of* x *then the localization deviation associated with* $|\phi\rangle$ *has the degree* d_i *if there is a maximum of the form* $|\bar{\delta}_i\delta_i| = \max\{|\bar{\delta}_j\delta_j| \mid 1 \leq j \leq n\}$.

On this interpretation one assumes that a cognitive agent is able to perceive the localization deviation when drawing boundaries that delineate inhomogeneous regions with distinct qualities. This interpretation also allows for the possibility that the cognitive agent is unable or unwilling to make a judgement.

As in the case of classification, whether or not the first or the second hypothesis or neither is actually true of the geographic world is an empirical question and can (at least in principle) be determined by experiments.

4.3 Classification and delineation do not commute

After performing a classification or delineation operation on a geographic field, intermediate fields are generated (to a certain degree by fiat) that are such that all locations of the base space of the underlying fiber bundle are mapped to eigenstates of the classification or delineation operators that correspond to the classification of delineation values assigned to the point of the base space.

$$\begin{array}{ccccccc} & & & \text{Fourier transform} & & & \\ \{\delta\} \subset C^n & \xrightarrow{f_X^{-1}} & \{|\phi\rangle\} \subset \mathcal{H}_x = \pi^{-1}(x) & \xleftarrow{f_Q^{-1}} & \{q\} \subset C^n & \xrightarrow{\hat{\mathcal{C}}} & |Q_i\rangle \\ \downarrow \hat{\mathcal{D}} & & & & & & \downarrow \hat{\mathcal{D}} \\ |\Delta_i\rangle & \xrightarrow{\hat{\mathcal{C}}} & \sum_k q_k\,|Q_k\rangle & \Longleftarrow \hat{\mathcal{C}}\hat{\mathcal{D}} \neq \hat{\mathcal{D}}\hat{\mathcal{C}} \Longrightarrow & & & \sum_k \delta_k\,|\Delta_k\rangle \end{array} \quad (5)$$

Consider Diag. 5. If classification and delineation operations are performed in sequence (in either order) then the input of the second operation is a field that at all locations of the base space is in eigenstates of the preceding operator. States represented in the localization base are related to equivalent descriptions in the quality base via (discrete) Fourier transforms and vice versa [8]. States related by Fourier transforms are *complementary* in the sense that if one state is maximally determinate then its Fourier transform will be minimally determinate [8]. Consequently, the order of the sequence of the application of classification and delineation operations ($\hat{\mathcal{C}}\hat{\mathcal{D}}$ vs. $\hat{\mathcal{D}}\hat{\mathcal{C}}$) to geographic fields is significant. Performing the classification operation first will result in precise classification and an imprecise delineation. By contrast, performing the delineation operation first will result in a precise delineation and an imprecise classification. That is, the operators representing classification and delineation operations do not commute. That the operators representing classification and delineation operations do not commute is captured in the formalism of QG as follows:

$$\forall\,|\phi\rangle \in \mathcal{H}_x :\ (\hat{\mathcal{C}}\hat{\mathcal{D}} - \hat{\mathcal{D}}\hat{\mathcal{C}})\,|\phi\rangle =_{df} [\hat{\mathcal{C}}, \hat{\mathcal{D}}]\,|\phi\rangle \geq i\Omega \quad (6)$$

As above, $i = \sqrt{-1}$ is the imaginary unit and Ω is a scale factor. Equation (6) a restatement of Heisenberg's uncertainty principle [8] for classification and delineation operations.

5 More open questions . . .

An important open question is how to understand the role of the scale factor Ω in QG. In concluding the following preliminary ideas are offered in contexts where geographic fields are considered at time scales at which they can assumed to be in *stationary states*. In such a stationary state a geographic field does not change in ways that affect its classification and delineation. That is, structurally complex (but real *valued*) terms of the form $|\overline{q}_i q_i|$ and $|\overline{\delta}_i \delta_i|$ remain unchanged in stationary states. By contrast, their constituting terms – the complex *valued* q_i and δ_i – are not constant and change in ways that is governed by the time-independent Schrodinger Equation [8]. According to this equation the phase factors of the complex values q_i and δ_i change and form a standing wave which frequency depends on the scale factor Ω in ways that mirrors Plank's constant in QM. This suggests that the complex values of a geographic field in a stationary state corresponds to the frequency of the processes that give rise to the geographic quality determinables associated with the underlying geographic field. This would fundamentally link the aspects of granularity and inhomogeneity of geographic fields with the frequency of the underlying processes that give rise to those fields.

References

1 R. G. Bailey. Delineation of ecosystem regions. *Environmental Management*, 7:365–373, 1983.

2 J. Berg. Aristotle's Theory of Definition. In *ATTI del Convegno Internazionale di Storia della Logica*, pages 19–30. San Gimignano, 1983.

3 CEC. Ecological Regions of North America. Technical report, Commission for Environmental Cooperation, `ftp://ftp.epa.gov/wed/ecoregions/na/CEC_NAeco.pdf`, 1997.

4 CEC. Ecoregions of the United States. `http://nationalatlas.gov`, September 2009.

5 John E. Chappell. Geography and quantum physics. *The Professional Geographer*, 47(2):220–221, 1995.

6 P. Dirac. *Principles of Quantum Mechanics*. Oxford : Clarendon Press, 1930.

7 EPA. Primary Distinguishing Characteristics of Level III Ecoregions. Technical report, Environmental Protection Agency, `http://www.epa.gov/wed/pages/ecoregions/level_iii.htm`, 2002.

8 David J. Griffiths. *Introduction to Quantum Mechanics*. Addison Wesley, 2004.

9 Jan Hilgevoord and Jos Uffink. The Uncertainty Principle. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2012 edition, 2012.

10 B. Z. Iliev. Fibre bundle formulation of nonrelativistic quantum mechanics (full version). *eprint arXiv:quant-ph/0004041*, April 2000. `arXiv:quant-ph/0004041`.

11 Barry Kronenfeld. Gradation as a Communication Device in Area-Class Maps. *Cartography and Geographic Information Science*, 32(4):231–241, 2005.

12 R. Omnès. *The Interpretation of Quantum Mechanics*. Princeton: Princeton University Press, 1994.

13 William Peterman. Quantum theory and geography: What can Dr. Bertlmann teach us? *Professional Geographer*, 46(1):9, 1994.

14 David H. Sanford. Determinates vs. Determinables. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2011 edition, 2011.

15 R. N. Sen and G. L. Sewell. Fiber bundles in quantum physics. *Journal of Mathematical Physics*, 43:1323–1339, March 2002. `doi:10.1063/1.1447309`.

Urban Artefacts and Their Social Roles: Towards an Ontology of Social Practices

Alessia Calafiore[1,2], Guido Boella[1], Stefano Borgo[3], and Nicola Guarino[4]

1 **Department of Computer Science, University of Torino, Torino, Italy**
calafiore@di.unito.it, boella@di.unito.it
2 **SnT, University of Luxembourg, Luxembourg**
3 **ISTC-CNR Laboratory of Applied Ontology, Trento, Italy**
stefano.borgo@cnr.it
4 **ISTC-CNR Laboratory of Applied Ontology, Trento, Italy**
nicola.guarino@cnr.it

Abstract

Cities can be seen as systems of urban artefacts interacting with human activities. Since cities in this sense need to be organized and coordinated, convergences and divergences between the "planned" and the "lived" city have always been of paramount interest in urban planning. The increasing amount of geo big data and the growing impact of Internet of Things (IoT) in contemporary smart city is pushing toward a re-conceptualization of urban systems taking into consideration the complexity of human behaviors. This work contributes to this view by proposing an ontological analysis of urban artefacts and their roles, focusing in particular on the difference between social roles and functional roles through the prism of social practices.

1998 ACM Subject Classification I.2.0 [General] Philosophical Foundations

Keywords and phrases urban artefact, ontology, social practice, urban planning

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.6

1 Introduction

From a human geography perspective, the notions of *space* and *place* have been considered as the opposite extremes of a continuum which goes from the ideal geometrical abstraction of space to the experiential world of place [5]. Understanding human conceptualization of place entails referring to the meanings which people commonly associate with their spatial experiences [19]. With the emergence of volunteered geographic information and geo-social media, such meanings are encoded in a growing amount of geo-referenced data collected by people who have a non-expert viewpoint on possible place uses [8]. As a consequence, a very important practical concern of Geographic Information System (GIS) and urban planning is to make explicit, for the purpose of mutual understanding and interoperability, people's assumptions about their everyday spatial experiences.

This paper considers a special class of artefacts, *urban artefacts*, designed and built for urban use. We take the view that a city can be seen as a system of urban artefacts, which may play distinct roles when taking part in social practices. Our focus is thus on the interaction between artefacts and their social uses. Different kinds of places will be defined, throughout the paper, as socially constructed concepts that emerge from the interaction between urban artefacts and collective human practices. We introduce here the concepts of social and functional places. It is crucial to consider them as different from the more common

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 6; pp. 6:1–6:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

sense idea of place as geographical space: they are treated as social concepts, particularly, as social roles[1].

The notion of Urban Artefacts is grounded in the DOLCE foundational ontology[2][12, 3]. DOLCE focuses on *particulars*, which differ from *universals* since they cannot be instantiated (i.e. "my car" vs. "car"), and are subdivided in endurants, perdurants, qualities and abstract entities. DOLCE has been explicitly engineered to capture human common sense meanings with a definite cognitive bias, and provides crucial notions to describe socio-technical systems. In the light of this, we consider DOLCE as the most suitable reference framework to ground a representation of social geographical knowledge. The relevant portion of DOLCE employed in the present analysis is reported in Figure 1, which shows how the various notions introduced in this paper are grounded in the DOLCE ontology.

The paper is organized as follows: first, we will provide a general definition of artefact as it has been discussed in the field of formal ontology so far. Then, we will apply the definition of artefact to represent the built environment, introducing the notion of urban artefacts and the needed distinction between their social and planned views. On the one hand, physical qualities resulting from planner design choices are used to classify urban types. On the other hand, the highly dynamic interaction between the built environment and human collective behaviour will be framed using roles theory; particularly, roles played by urban artefacts are considered as depending on their participation in social practices with human agents. Finally, we will exemplify these cases in a real context. We will present expected and unexpected 'meanings' attributed to Piazza del Mercato in Naples (Italy), where roles that an urban artefact plays may diverge from its own ontological status. The scenario shows how to ontologically distinguish different perspectives providing room for the creation of a map of urban artefacts and their multiple roles in an urban environment.

2 Urban artefacts

The goal of this section is to apply the notion of artefact to the field of urban modeling. Here we take the view that a city can be seen as a system including (urban) artefacts of various kinds that may participate to agents' activities by playing multiple roles, possibly at the same time. Our focus is thus on the study of artefacts and their social uses. As we will see in the next section, this leads us to introduce the notion of *mode of deployment* – an idea developed in engineering design for functional studies – into the field of urban studies.

2.1 From Artefacts to Urban Artefacts

The notion of urban artefact that we develop starts from the characterization of ontological artefacts in [4] and artefactual objects in [9]. These approaches discuss broad notions.

In Borgo and Vieu [4] artefacts are modeled as the result of an intentional act of some agents (the artefact's creators) which, by creating a new entity, determine its constitution, capacity, attributed capacity and (implicitly) identity criteria. In their view, an intentional selection *and* capacity attribution is a mental event (possibly associated with physical actions)

[1] At the same time, we do not deny the existence of another ontological level – corresponding to places intended as geographical spaces rather than concepts. However, for the purpose of the current work, in order to be able to distinguish between expected and unexpected – or planned and unplanned – uses of the city, we will only focus on the concepts of social and functional place: two different roles that urban artefacts may play.

[2] `http://www.loa.istc.cnr.it/old/DOLCE.html`

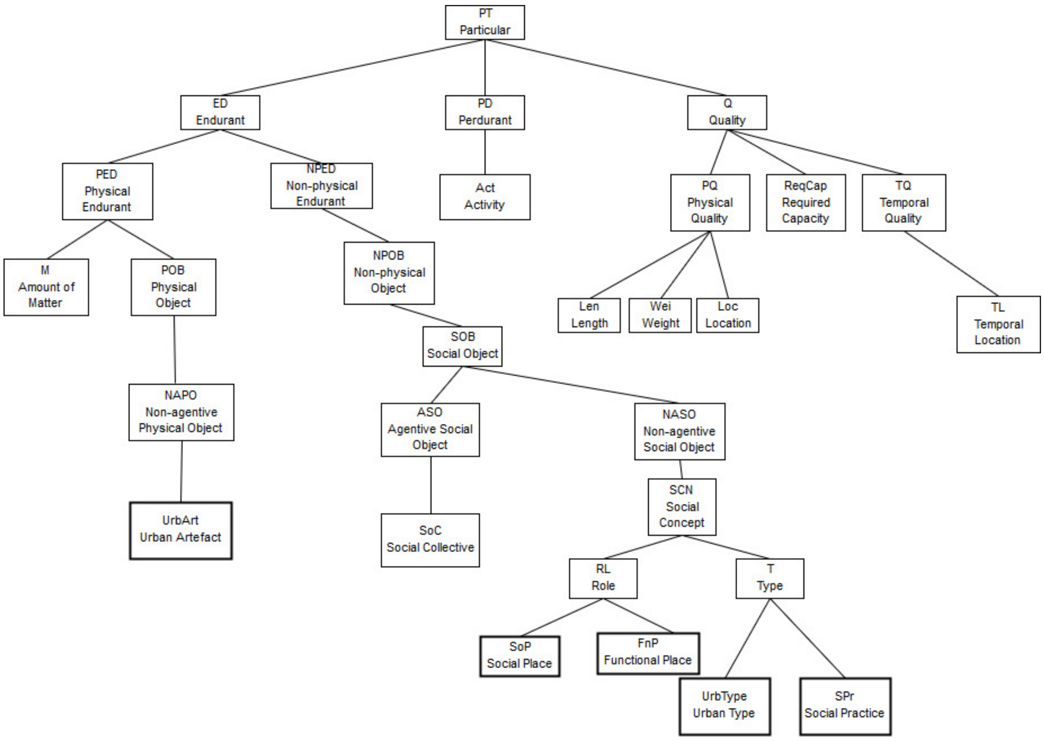


Figure 1 General taxonomy in DOLCE.

that *creates* an artefact: when an agent chooses a pebble as her paperweight, an artefact is immediately created, since an intentional selection and capacity attribution (namely, to behave as a paperweight) have taken place. When the artefact, due to inappropriate attribution or to usage, does not manifest the capacities attributed to it, it is said to be malfunctioning.[3] This view of artefacts has been expanded to account for *technical artefacts* in [1, 2, 10]. In particular, in [2] the authors have analysed different understandings of technical artefacts, isolating, among others, the notion of *technological artefact*.

Guarino [9] introduced the category of *artefactual object* to capture the view of artefact as an entity that realises a given design. This way, artefactual objects are not the result of a direct intentional selection, since the artefactual status is due to the relationship between the object and the design specification. So, the connection between an object and its artefactual status becomes independent of any specific agent act, while the historical property of *being an artefact* is considered as *anti-rigid*, in the sense that it is accidental for all its instances. A naturally fallen tree may be used as a chair (playing therefore an *artefactual role*), but it is not an artefactual object unless it satisfies some chair design. In short, this leads to assume that no artefactual entity is created when certain qualities of the fallen tree are intentionally selected in order to use it as a chair. Artefactual objects are entities whose essence lies not in some attributed capacity to fulfill a certain function, but rather in the match between the object's features and a design specification. Therefore a distinction is needed: to be an instance of an artefactual type, an object is required to match a design, while to be an

[3] This means that artefacts may be disconnected from their actual use. Note that an object stably used to perform some activity is not an artefact until it is identified by its user as the 'tool' for such an activity. In the paperweight example, the user and the creator are the same agent but this is not true in general.

instance of an artefactual role it is enough to be used in a way that is the expected way of use of some artefactual type. Recalling our previous example, in this view the pebble used as a paperweight is not an artefactual object, but only plays an artefactual role unless there is a specific design that is satisfied by that pebble.

Considering urban artefacts, these normally originate from an act of rational design and intentional construction. They fit, therefore, the specific notion of technological artefact defined in [2][4], but they can also be considered as artefactual objects according to [9], since it makes sense to assume for them that a design specification exists.

2.2 Grounding Urban Artefacts in DOLCE

We conceive an urban artefact as a component of an urban system which is formed by physical objects and/or amounts of matter[5], shaped or somehow organized in order to satisfy some design specifications.

Such specifications may deal with different kinds of information, including:

- *Design constraints* concerning the physical structure of the urban artefact and its physical qualities;
- Intended *use scenarios* in terms of *modes of deployment*, i.e., how an urban artefact is supposed to be used or exploited;
- *Normative constraints* concerning forbidden uses or explicit use rights allowed to specific classes of users. For instance, a park may include a playground where children may play, or where only children may play, and a green area where to keep off the grass.[6]

We shall assume that a design specification characterizes an urban type (*UrbType*), which is a category of urban artefacts characterized by a *prototypical* design of architectural and urban interest, which *a-priori* identifies specific physical qualities. Once an artefact is a member of the *UrbType* category it remains so for all its life, i.e., until some disruptive change occurs like the destruction of the physical object that composes the artefact or the modification of one of its core characteristics. In both cases, the original artefact ceases to exist while a new one may appear. It is also possible to have co-located artefacts constituted by the very same physical object. While being an urban artefact of some type (e.g. school type), the same building could belong to another type as well (e.g. polling station type).

Of course, design specifications can be described at different levels of detail: in general each specification is associated to an *urban type*, which may be further specialized in several variations. In principle, once a design is completely specified it may be *realized* by multiple physical objects (say, multiple buildings with the same design), but often urban artifacts are realized just once. According to DOLCE, design specifications are classified as *descriptions*, which are a kind of abstract entities, while urban artefacts are classified as *physical objects*, and more exactly as *non-agentive physical objects* (NAPO), at least in the typical case.[7] In turn, physical objects are a subclass of *endurants* (entities that persist in time by keeping all

---

[4] A technological artefact *a* is a physical object which is, firstly, created by the carrying out by an agent (or group of agents) of a make plan for a physical object with a physical description id, and for which, secondly, a use plan exists.

[5] For instance, it can be a fountain including the water thatu springs from it.

[6] Note that design constraints may in turn result from the obligation to satisfy certain normative constraints that reflect, for instance, quality or safety requirements. Such normative constraints are however very different from those concerning forbidden or allowed use of the artefact.

[7] An urban artefact can also be formed by agentive (APO) physical objects such as robotic systems broadly understood (e.g. traffic control systems). This case is going to become more relevant with the Internet of Things paradigm.

their parts present at each time), and are distinguished from *amounts of matter* (M) since their identity depends on a specific structure, and not just on the parts they are composed of. Like physical objects, all urban artefacts have a spatial location, which is a geo-referenced quality, since its quality space is associated with a geographic coordinate system (GCS).

However, if we aim to model finer changes in the evolution of the city, and in particular changes caused by social practices, this view needs to be enriched with a more flexible classification, where an object of urban interest can change its status depending on the context.

An interesting example of an urban artefact is Piazza del Mercato in Naples. It is formed by some NAPO elements (buildings facades, paved floor, lights, benches and other urban decorations) whose qualities and physical structure satisfy the generic characteristics associated to the concept *square*, and the specific ones described in the design specification according to which *that* urban artefact was originally realized.[8] An urban artefact, which is often a system of both artefacts and natural objects (buildings, benches and trees), can be seen at different levels of granularity: sometimes the square is the focus, in other cases the focus is the neighbourhood of which the square is just a component. In the latter case, the benches or the trees in the square may not be considered as elements of the larger urban artefact just because, at a coarser granularity, the square may be considered as atomic.[9] e is no design specification for it: it just plays the role of a chair.

3 Urban artefacts roles

In real scenarios, the intended use of an urban artefact (described in its design specifications) may not correspond to its actual use in social practices. To model this mismatch, due to the multiple and unexpected ways in which urban artefacts can be used, we introduce the distinction between *urban types* (*UrbType*) and *urban artefact roles*. Urban artefacts, which are instances of urban types, may play several social roles. For example, a *school* can be used as *meetingpoint* during a demonstration. Does this change of use imply a change of identity of the school type?

We believe not. For this reason, we will model uses of artefacts in terms of roles theory. Both types and roles are recognizable by a society or a group of agents, but, according to [13], the former are *rigid* properties (`R`), while the latter are *anti-rigid* (`AR`) and *founded* (`FD`).

Concerning the *way* these social roles emerge from social practices, an important distinction is to be done between the *institutional* roles associated to urban artefacts by their designers, and the *non-institutional* roles that may be actually played. The former are the result of a design choice made at the time the urban artefact is planned, and typically have a functional nature (for instance, the function of Piazza del Mercato might originally have been that of a *market place*); the latter are just related to the fact that an urban artefact may actually be used in a way different from that originally planned. Note that several social roles, especially in an urban scenario, may be played both in an institutional and a non-institutional way: for instance, the same role of market place may be an institutional role for an urban artefact of a certain type (a square) and non-institutional role for an urban artefact of a different type (a church). On the other hand, certain social roles can only

[8] Note that not every square is an urban artefact, since some squares may not have been designed.

[9] The rules that determine which types of physical objects are components of an artefact are not discussed in this paper. Some of these rules combine different aspects like size, function and spatial disposition. Their discussion requires to refer to a set of urban artefact types, but the characterization of such a type is itself an open topic in the literature.

be played in one way, either institutionally or non-institutionally. For instance, the role of *president* is always institutional.

With these categories we can distinguish situations in which the roles played by an urban artefact are compatible with its type from those in which they are not. This is a necessary step to be able to model mismatches between the ‘planned’ and the ‘lived’ city or simply to separate the compliant and the non-compliant aspects of an emergent social pattern in which an urban artefact participates.

Local knowledge expressed by emergent social patterns is crucial to represent the concept of places as recognized by people who live them.

As a consequence, multiple perspectives on urban artefacts are needed: on the one hand, the designer’s perspective, which is based on explicitly stated uses and modes of deployment; the local authority perspective, based on explicitly stated normative constraints, and the users’ perspective based on how the artefact actually contributes to the activities of the inhabitants.

The user’s perspectives may agree with the institutional role of an urban artefact in two cases, namely, when the users’ activities are in agreement with the modes of deployment established by the designer; or when the artefact itself is “re-created” via the recognition, e.g., by the city authority, that the users’ uses are thereafter allowed, i.e., considered as “institutional”. Of course, this latter step may require that the design is rewritten at least in part. The conditions under which this change determines the change of identity of the urban artefact may be subtle. One would like to say that this surely happens when the urban artefact is structurally and functionally modified by changing its topmost type, e.g. when a church is transformed into a hospital.[10] Yet, even these changes do not determine a change of identity when they are assumed to be temporary, like when a church is transformed into a military hospital during a war. We distinguish two types of roles (see Fig. 1):

- the functional place role (FnP) which is compliant with a specific urban type and it can be considered as an institutional role;
- the social place role (SoP) which is related with unexpected uses of the artefact and it can be thought of a non institutional role.

Roles do not depend on an *a-priori* design or mode of deployment but only on the collective use of the artefact at some point in time. The same artefact can play multiple roles at the same time, it can start and stop to play a role several times in its life, and several artefacts (perhaps differing in type) can play the same role. In the case of a school during the demonstration, we can say that the artefact’s identity remains unchanged, it still is a school, but the artefact plays the role of being a meeting point during that day.

It is crucial to recognize the variety of roles that an urban artefact can play depending on actual uses. Social practices, being recurrent activities, define the way in which an urban artefact is actually used: they can be compliant or not with the uses foreseen by the urban type. This observation is key to the difference between functional and social place, the urban artefactual roles modeling the expected (functional) uses of an Urban Artefact and the actual (social) uses which can be unexpected.

To define the roles and to use them to classify urban artefacts, social practices have to be recognized, thus a definition of social practices needs to be introduced.

[10] Here we assume that church and hospital are general urban types and that there is no higher type (in particular, building would not be a type).

3.1 Defining Social Practices

In a society there are established and accepted behavioral patterns which characterize traits of human organizations. We find this especially in complex social organizations like cities. Inspired by the social geography field of study [18, 11], we claim that the actual meaning attributed to an urban artefact depends on social practices. The notion of social practice has been object of several studies, mostly in sociology [6, 15] and social geography [18], but never clearly formalized, with the exception of Tuomela [20], who proposed a complete formal representation of social practices, albeit with a strong bias on groups intentionality and a low level of awareness of their spatial aspects.

In the following, we propose an interpretation of social practices (SPr) based on the DOLCE ontology (see Fig. 1), which we shall use to define the social role an urban artefact plays depending on certain urban acts (UrbAct). Through the notion of social practice we can move our focus from individual to collective behavior. As stated in [7] the basic domain of study to interpret social structures is *neither the experience of individual actor, nor the existence of any form of societal totality, but social practices ordered across time and space.* A social practice models the way performed activities are situated in time and space and organized in a skilled and knowledgable fashion by groups of human agents [7], collected here in the DOLCE's category of *social collective* (SoC). To capture these intuitions, we shall use a primitive predicate $Rec(spr, y, t)$ standing for 'a social practice spr is recognized by the social collective y at time t. We have:

$$UrbAct(x) \rightarrow PD(x) \wedge \exists y(PC(y,x) \wedge UrbArt(y)); \tag{1}$$

$$Rec(spr, y, t) \rightarrow SPr(spr) \wedge SoC(y) \wedge Time(t); \tag{2}$$

$$SoC(y) \rightarrow ASO(y); \tag{3}$$

$$SPr(spr) \rightarrow \exists x, t, y\,(CF(x, spr, t) \wedge Act(x) \wedge SoC(y) \wedge Rec(sty, y, t)); \tag{4}$$

$$Act(x) \wedge Spr(y) \wedge CF(x,y,t) \wedge PC(z,x) \wedge DF(y,w) \wedge SoP(w) \rightarrow CF(z,w,t) \tag{5}$$

In the formulas above, which use the primitives introduced in [12] and [13], $PD(x)$ stands for 'x is a perdurant', $PC(x,y)$ for 'x participates to y', $DF(x,y)$ for 'concept x defines role y' and $CF(x,y,t)$ for 'x is classified by the concept y at time t'. Axiom (4) says that every social practice is a concept that classifies some activity taking place in a given space, so this means we are focusing on *urban* social practices. Also, a social practice requires the existence of a social collective that recognizes it, so we can say that it is specifically dependent on a social collective. To be recognized by a social collective we shall also assume that the social practice is a recurrent activity[11]. For example, the *urban* social practice of having dinner at a particular place is not recognized in the same way by everyone: different groups would characterize it as constrained by different temporal or spatial conditions as well as action patterns depending on how that practice is usually performed in that social and cultural system. Therefore, we can say that there are many possible types of the social practice of having dinner at that place (say, for certain special occasions, or as a regular habit every week), each recognized by one or more social groups.

We can also have distinct social practices emerging from different spatial patterns. For example, in the last decades large shopping malls have been considered as places of entertainment as a result of a change in urban lifestyles in the so-called post-modernist

[11] This assumption means that an activity type, to be a social practice, has to be iterated several time by the members of the same social collective. Such constrain is not included in our formalization yet.

society [21]. The practice of entertainment counts as a new type of practice specifically located in a particular space. It follows that social practices are conceptualized differently depending on temporal and spatial locations. On the contrary, we may consider an emerging activity type as a social practice only relatively to the fact that the activities are recognized by a social collective. Finally, Axiom (5) states that the social place role defined by a given social practice classifies the urban artefact(s) which is used in the associated social practice activities.

3.2 How social practices define urban places

As we have seen, urban artefacts may play institutional or non-institutional roles. Institutional roles are those that are played in compliance with the intended function of the urban artefact. Playing such institutional role may be recognized as a recurrent activity by a certain social collective, so that a social practice emerges. Whenever such a social practice exists, we can say that the urban artefact marks an *institutional place*.

For instance, urban artefacts of type *school* are designed for teaching activities. Teaching is an intended use typically included in the design specifications for this type of artefact. When the artefact plays its institutional role, this is also a recurrent activity of the social collective of educators, so that a social practice emerges from such recurrent collective behaviour. When this holds, we can say that the school is a *teaching place* – more exactly, an *institutional teaching place*. From a different perspective, the same school can be seen as an *institutional learning place*, given the social practice of learning recognized by the social collective of students. In conclusion, we shall say that *being an institutional place* is an emergent and dynamic property of an urban artefact that depends on the iteration of some collective behavior, which can be ascribed to an urban artefact since such collective behavior is compliant with its planned use. So, being institutionally intended to be used for a certain purpose is not enough for an artefact to mark an institutional place.

Classifying schools as *teaching places* is rather natural, but the teaching collective behavior might be manifested also outside schools, e.g., in hotels, bookstores, factories and parks, which are not sub-types of school. Likewise, we may use a school just as a meeting place or as a place for recreational activities.

These unforeseen uses of an urban artefact are quite common: a square may be used for praying, a church may be turned into an hospital, a school yard may be used for sports tournaments. All these uses involve playing *non-institutional* roles, in a way that is not compliant with the uses foreseen in the urban artefact's design specifications. In these cases, the urban artefacts marks a *non-institutional* place.

So, we shall define a non-institutional place as a place marked by an urban artefact participating in social practices, recognized from a recurrent activity, such that the artefact is not used in compliance with its designed uses. Non-institutional places are often called *social places*, so in the following we shall use both terms interchangeably.

Note that, differently from an urban artefact, social place and institutional places are not generally recognizable *per se*. They are recognized as such only by the members of a social collective since these share the knowledge of (and may also participate to) that social practice. Indeed, the so-called *local* knowledge, related to urban places, is shared in specific communities and is generally the most difficult knowledge to gather, since it is not the subject of standard geography classifications. At the same time, awareness of the different groups which use an urban artefact in distinct ways is essential to identify the actual stakeholders. Our approach to conceptualize social place is theoretically grounded in social geography theories [11, 16, 14, 18], where it is claimed that the *social content* of the city is the basis for reading it. The study of social content focuses on the social structures, generated by collective behaviors, in connection with the forms of the city where they appear [17]. Differently from institutional places, where the social practice is compliant with the planned uses, the essence of a social place is of being related to unexpected social practices in which urban artefacts participate. By using this notion, we can exploit the local knowledge about the lived city expressed by different social collectives to describe the urban social environment.

To show some implications of this modeling choice, a real example is discussed in the following section.

4 Piazza del Mercato, Naples – A real context example

Piazza del Mercato is a square in Naples which represents an interesting as well as rather complex example of the differences between typologies and typicalities characterizing urban artefacts. In the framework developed above, a typology refers to $UrbType$ while typicalities are expressed by the performed social practices. These types and practices determine the social places and the institutional places marked by urban artefacts.

Several urban regeneration plans of Piazza del Mercato have been discussed in the last decade and the process is far from being over. Since it is not possible to illustrate here all the aspects related to this area, we will focus on two situations that exemplify how to use the framework for the representation of the urban environment. Note that each situation presents different processes which in turn provide multiple perspectives of Piazza del Mercato.

4.1 First situation

Piazza del Mercato in Naples (pdm) is an urban artefact classified by the urban type *square* and constituted by a number of non-agentive physical objects such as fountains and lampposts, and with specific physical qualities, e.g., location, size and delimitations, and normative constraints like the no-parking restriction over the whole area.

$$UrbArt(pdm) \tag{6}$$

Although parking was not allowed, pdm has been used as a parking area until 2006, when fences and a CCTV system were installed. This is an example of an urban artefact that was not designed to be a parking area, but a typicality of its use is related to the social practice of parking. As long as the social practice existed, pdm was classified by the local social community as a *parking place*.

This social practice conceptualizes a recurrent activity of parking that was performed in pdm. The activity was recognized specifically by users that participate in that use of pdm. Also, in pdm's design specifications it was (and currently is) required not to use the area for parking. Therefore, *being a parking place* was a non-institutional role of pdm, being in contrast with the normative constraints of its design specifications. A social place therefore emerged.

Given the following conditions:

$$Act(parkingInPdM); \tag{7}$$

$$ActIn(parkingInPdM, pdm, 2005); \tag{8}$$

$$Spr(pdm_{parking}); \tag{9}$$

$$CF(parkingInPdM, pdm_{parking}, 2005); \tag{10}$$

$$SoP(parkingPlace); \tag{11}$$

$$DF(pdm_{parking}, parkingPlace) \tag{12}$$

Figure 2 The image shows Piazza del Mercato, in Naples, where a group of muslims are praying and a football field with mobile goalposts is drawn.

We can now infer that *pdm* is classified by the social place role *parkingPlace*:

$$CF(pdm, parkingPlace, 2005) \tag{13}$$

4.2 Second situation

After 2006 the parking practice was eliminated and new social practices emerged. As we can see in Figure 2, *pdm* became a place where Muslims meet to pray and young people to play football. These are unexpected uses of the square that are not specifically ascribed to its type nor identifiable through design specifications. However, the knowledge Muslims or young people have about *pdm* is related to their experiences of *pdm*. Let us say that Muslims meet in *pdm* to pray on Friday morning and youngsters play football on Sunday.

Only from the activities that members of the two social collectives (recurrently) perform we can identify the two non-institutional roles played by *pdm*: *being a praying place* and *being a football place*.

$$SPr(pdm_{praying}); \tag{14}$$
$$Act(prayingInPdM); \tag{15}$$
$$SoC(muslims); \tag{16}$$
$$CF(prayingInPdM, pdm_{praying}, fridays\text{-}in\text{-}2017); \tag{17}$$
$$Rec(pdm_{praying}, muslims, fridays\text{-}in\text{-}2017); \tag{18}$$
$$SoP(prayingPlace); \tag{19}$$
$$SPr(pdm_{playingFootball}); \tag{20}$$
$$Act(playingFootballInPdM); \tag{21}$$
$$SoC(youngPeople); \tag{22}$$
$$CF(playingFootballInPdM, pdm_{playingFootball}, sundays\text{-}in\text{-}2017); \tag{23}$$
$$Rec(pdm_{playingFootball}, youngPeople, sundays\text{-}in\text{-}2017); \tag{24}$$
$$SoP(footballPlace)\,. \tag{25}$$

Therefore we conclude:

$$CF(pdm, prayingPlace, fridays\text{-}in\text{-}2017); \tag{26}$$
$$CF(pdm, footballPlace, sundays\text{-}in\text{-}2017)\,. \tag{27}$$

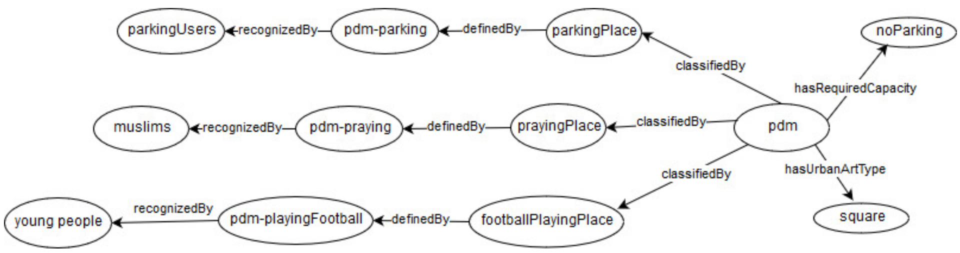


Figure 3 The figure shows some of the instances (in ellipses) and their relations that can be used to represent the situations characterizing Piazza del Mercato. It can be seen that the urban artefact *pdm* is classified by different social places; this is expression of the multiple social collective's point of views which experience *pdm* through different social practices.

4.3 Final remarks

Piazza del Mercato in Naples has been used to exemplify possible implementations of our framework. Some final remarks are needed:

- urban artefacts cannot be classified exclusively on the basis of their designed uses, since they are changeable depending on how people actually experience them;
- social and functional places, as defined in this paper, are emergent roles (properties);
- places, as understood here, are expression of different social collectives' perspectives;
- we believe that the conceptual system we developed helps to better identify the stakeholders of an area and to establish how a new design may impact them;

Given these observations Piazza del Mercato, which here we studied only in minimal part, seems to be better represented in its complexity via a role theory. In Figure 3 we sketch how the notion of social place can support the recognition of different stakeholders through the analysis of social collectives. Also, taking into consideration social practices to define an urban artefact's roles allows an evaluation of the compliance between actual and required uses.

5 Future work and Conclusions

The social use of urban artefacts and their type are two interacting aspects of the built environment. The way urban artefacts are designed strongly influences the way we use them, and forces the construction of specific social contents related to recognized collective behaviours. Cities, indeed, can be interpreted as the result of the political, economical and social organization of contemporary societies.

However, the emergent semantics of an urban artefact's social roles goes far beyond the one addressed by a specific type. Also, making explicit the social characterization of the urban areas is the grounding for a better contextualization and evaluation of the way these are planned and how they can be changed.

The formal framework described in the previous sections defines only a static representation of the urban environment. It answers the following question: How can ontological analysis help modeling the *social content* [17] of the build environment besides its planned characterization?

Clearly, the dynamics of the existing mutual interaction between the social roles and the urban types needs further analysis and specifications. Even if it is not the focus of this contribution, being able to represent dynamic processes that arise from this mutual interaction is crucial to support decision making in the definition of re-design patterns that want to be more contextualized in the social dimension of an area.

In a typical design studio scenario, groups of professionals such as architects, urban planners and landscape architects generally start their work with site visits and intense analysis to deeply describe the current state of an area. This work is all on the shoulders of these professionals which use both qualitative and quantitative methods. Involving citizens via a participatory process to collectively build a *social* knowledge base about an area is generally considered too expensive. Nowadays, with the growing availability of geo-based mobile tools it has become much easier to collect data via crowdsourcing [12], and with the Internet of Things (IoT) paradigm every physical object in the city can potentially have a digital identity and sense the environment around itself as well as interact with the others.

Recording the dynamics of place uses and the way they are recognized by different social collectives will allow a better understanding of design choices and evaluations. Grounding the wealth of data collected through crowdsourcing, geosocial media, and the IoT in our framework, can be extremely useful – especially in the first phase of participatory urban planning, in order to identify the relevant stakeholders and critical areas.

References

1 Stefano Borgo, Massimiliano Carrara, Pawel Garbacz, and E. Pieter Vermaas. A formal ontological perspective on the behaviors and functions of technical artifacts. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23:3–21, 2009.

2 Stefano Borgo, Maarten Franssen, Pawel Garbacz, Yoshinobu Kitamura, Riichiro Mizoguchi, and E. Pieter Vermaas. Technical artifacts: An integrated perspective. *Applied Ontology*, 9:217–235, 2014.

3 Stefano Borgo and Claudio Masolo. Ontological Foundations of DOLCE. In Roberto Poli, Michael Healy, and Achilles Kameas, editors, *Theory and Applications of Ontology: Computer Applications*, chapter 13, pages 279–295. Springer, 2010.

4 Stefano Borgo and Laure Vieu. Artefacts in formal ontology. In *Handbook of phylosophy of technology and engineering sciences*, volume 9, pages 273–307. 2009.

5 Helen Couclelis. Location, place, region, and space. *Geography's inner worlds*, 2:15–233, 1992.

6 Anthony Giddens. *Central problems in social theory: Action, structure, and contradiction in social analysis*, volume 241. Univ of California Press, 1979.

7 Anthony Giddens. *The constitution of society: Outline of the theory of structuration.* Univ of California Press, 1984.

8 Michael F. Goodchild. Formalizing place in geographic information systems. In *Communities, Neighborhoods, and Health*, pages 21–33. Springer, 2011.

9 Nicola Guarino. Artefactual systems, missing components and replaceability. In *Artefact Kinds: Ontology and the Human-Made World*, pages 1–11. Springer International Publishing Switzerland, 2014.

10 Gilles Kassel. A formal ontology of artefacts. *Applied Ontology*, 5:223–246, 2010.

11 Henri Lefebvre. *The production of space*, volume 142. Oxford Blackwell, 1991.

12 C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, and L. Schneider. Dolce: a descriptive ontology for linguistic and cognitive engineering. *WonderWeb Project, Deliverable D17 v2.1*, pages 1–38, 2003.

13 Claudio Masolo, Laure Vieu, Emanuele Bottazzi, Carola Catenacci, Roberta Ferrario, Aldo Gangemi, and Nicola Guarino. Social roles and their descriptions. In *KR*, pages 267–277, 2004.

[12] This kind of information is generally known as Volunteered Geographic Information.

14 Doreen Massey. *For space*. Sage, 2005.
15 Andreas Reckwitz. Toward a theory of social practices a development in culturalist theorizing. *European journal of social theory*, 5(2):243–263, 2002.
16 Edward Relph. *Place and placelessness*, volume 67. Pion London, 1976.
17 Aldo Rossi. *The architecture of the city*. The MIT Press, 1984.
18 Edward W. Soja. *Postmodern geographies: The reassertion of space in critical social theory*. Verso, 1989.
19 Yi-Fu Tuan. *Space and place: humanistic perspective*. Springer, 1979.
20 Raimo Tuomela. *The philosophy of social practices: A collective acceptance view*. Cambridge University Press, 2002.
21 Sharon Zukin. Urban lifestyles: diversity and standardisation in spaces of consumption. *Urban studies*, 35(5-6):825–839, 1998.

An Ontological Framework for Characterizing Hydrological Flow Processes

Shirly Stephen[1] and Torsten Hahmann[2]

1 **School of Computing and Information Science, University of Maine, Orono, ME, USA**
shirly.stephen@maine.edu

2 **School of Computing and Information Science, University of Maine, Orono, ME, USA**
torsten@spatial.maine.edu

Abstract

The spatio-temporal processes that describe hydrologic flow – the movement of water above and below the surface of the Earth – are currently underrepresented in formal semantic representations of the water domain. This paper analyses basic flow processes in the hydrology domain and systematically studies the hydrogeological entities, such as different rock and water bodies, the ground surface or subsurface zones, that participate in them. It identifies the source and goal entities and the transported water (the theme) as common participants in hydrologic flow and constructs a taxonomy of different flow patterns based on differences in source and goal participants. The taxonomy and related concepts are axiomatized in first-order logic as refinements of DOLCE's participation relation and reusing hydrogeological concepts from the Hydro Foundational Ontology (HyFO). The formalization further enhances HyFO and contributes to improved knowledge integration in the hydrology domain.

1998 ACM Subject Classification I.2.4 Knowledge Representation Formalisms and Methods, I.2.1 Applications and Expert Systems

Keywords and phrases hydrology, flow processes, formal ontology, participation, semantic roles

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.7

1 Introduction

Much progress has been made towards formal semantic representation of concepts in the water domain, though mostly by representing static physical and spatial hydrogeological features (e.g., rock bodies, water bodies, voids) while neglecting important dynamic aspects, such as the transport of water between the various stages of the water cycle and different places where water is stored. One of these dynamic aspects is movement of water via hydrological flow processes on the surface of the earth (such as runoff and stream flow), in subsurface rock formations (such as percolation), as well as movement of water between surface and subsurface entities (such as infiltration). The presented work is a step towards filling this gap by laying out an ontologically rigorous formal framework of hydrological flow processes that spans surface and subsurface flow and links the two. The framework distinguishes different kinds of hydrological flow processes based on their participants and organizes them taxonomically. The taxonomy is formalized using semantic participatory roles, which are played by different hydrogeological entities in the different kinds of flow processes, as refinements of DOLCE's participation relation. The participating static hydrogeological entities are expressed using concepts from the work towards the Hydro Foundational Ontology (HyFO) [4, 13, 14, 16] as a domain reference ontology for the water domain.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 7; pp. 7:1–7:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Hydrological flow consists of several spatio-temporal components that describe the movement of water above, on, and below the surface of the earth, and these are individually influenced by physical phenomena such as gravity, porosity and permeability of the soil zone, and capillary pressure. Existing standards for the hydrology domain model flow of water at general levels without refining them according to their participants, and focus on flow either in surface or ground water systems. However, the water cycle is a constant physical interaction between surface and subsurface water features, for example, surface water bodies may be fed from or discharge water to aquifers beneath and surface water may infiltrate the ground to become groundwater. These interactions, which are flow processes themselves and have properties in their own right (e.g., flow volume or speed), must be explicitly represented and thus require a unifying representation of surface and subsurface water flow. Their formal representation would benefit a range of applications, such as assessing hydraulic connectivity between surface streams and groundwater bodies, or determining flow paths between aquifers within aquifer systems or seepage of water through confining beds. Representation of flow processes also informs regional, agricultural, and urban planning, where information about where water comes from and where it flows to is required to maintain adequate water supplies and trace the path of water-born pollutants. Future incorporation into HyFO will improve HyFO's overall utility for analyzing, refining, and integrating flow concepts across existing hydro ontologies. For example, because of the lack of a flow concept in HyFO, GWML2's flow module [3], which models the flow of groundwater, was the only GWML2 part that could not yet be ontologically analyzed and logically specified using HyFO concepts [16].

Objective. The specific objective of this work is to ontologically analyze and categorize different flow patterns in the hydrological domain and to formally represent them as an extension to HyFO. In the process, we aim to at least partially address a number of challenging ontological questions about the nature of hydrological flow, including: What precisely are hydrological flow processes? What is common to all of them? How do they differ? Can they be clearly delineated? How are they related to hydrogeological entities? What are their spatial and temporal properties?

Scope. We limit our study to flow processes that (1) occur directly on or below the surface of the Earth and that (2) do not involve physical changes in the state of water matter, thus excluding other hydrological processes that transport water, such as precipitation, condensation, evaporation, and evapotranspiration. The aim is to represent spatio-temporal dynamic aspects of flow, leaving aside qualities and quantifiable properties that flow processes might exhibit, such as water pressure in an aquifer, or the speed or volume of flow, though the representation should be extendable by such parameters in the future. As such, the ontology is not intended to serve as a mathematical model for calculating flow quantities, but rather to express interactions between water bodies contained in different rock bodies and to capture the general physical pattern of different hydrological flows, including how flow processes are physically manifest in the different hydrogeological units/zones.

Approach. Each occurrence of a flow process manifests itself in specific hydrogeological entities from and to which water flows. Based on DOLCE's upper level classification, *hydrological flow* (HF) is modeled as a *perdurant* that can have temporal parts (e.g., sub-processes) and that is related to physical endurants, such as rock formations and water bodies, via DOLCE's time-indexed *participates* relation $PC(x, y, t)$. This approach permits different entities to participate at different times during the duration of a flow process.

We gather common kinds of flow processes from the hydrological literature and identify their participating hydrological and geological endurants and associated physical aspects, such as the spatial configuration and connectivity of geological formations or their porosity (the presence of connected voids), which capture minimal requirements for where and when different kinds of flow can occur. We then use Fillmore's case roles [5] to identify three types of participants common to all hydrological flow processes and formalize them as refinement of DOLCE's participation relation. They are subsequently used to develop a taxonomy of hydrological flow processes, each indicating what static hydrogeological entities – selected from HyFO's hydrogeological entities and newly axiomatized hydrological *subsurface zones* – must or can participate in each of the roles. In addition, basic temporal constraints between the participants are identified and formalized.

Sec. 2 surveys how flow is represented in existing hydro ontologies and Sec. 3 reviews HyFO's hydrogeological entities and DOLCE's participation relation as basis for our formalization. A preliminary analysis of flow processes leads to a unifying high-level pattern for hydrological flow and participatory roles in Sec. 4, which is formalized more fully in Sec. 5.

2 Related Work

Devaraju and Kuhn advocate in [6] a process-centric approach to relate processes to observable properties. While not the focus of their work, the authors highlight the benefits of identifying the entities that participate in processes to distinguish different types of processes (e.g., infiltration and percolation) and to identify processes that only differ in label (e.g., surface runoff and overland flow), which we pursue more systematically here. The formalization in [6] focuses on the hydrological processes of precipitation and evapotranspiration rather than the flow processes that we are concerned with.

Most closely related to our work is [9]. It identifies hydrogeological entities (e.g., sources, sinks, channels) that participate in flow processes and relates them to perdurant processes via BFO's [21] *involvement* relation, which is the inverse of DOLCE's *participation* relation. The identified processes include seven basic kinds of flow processes [9]: *overland flow*, *channel flow*, *infiltration*, *return flow*, *through flow*, *percolation* and *base flow*. Certain flows are identified a aggregate processes, which are a composition of these seven processes. However, no formal analysis to constrain the endurants that can participate in different processes is undertaken, neither is a full taxonomy of flow processes developed, nor are different participatory roles in hydrological flow processes formally distinguished.

Existing hydrology-related data models and ontologies include the Groundwater Markup Language (GWML2) [3], INSPIRE's data specification on hydrography [17], HY_Features [8], the hydrogeology extension to SWEET (SWEET-HG) [23], and the surface water ontology pattern (SWOP) [20]. HY_Features [8] and INSPIRE's hydrography specification [17] do not model flow processes at all. The others model hydrological flow to varying extents, but either model abstract flow paths rather than processes with spatio-temporal properties (e.g., GWML2, SWOP), or concentrate only on properties of flow, such as transported amounts of water or volumetric rates of flow, rather than hydrological or hydrogeological participants (e.g., SWEET-HG). Moreover, the data models tend to represent flow only within the surface (e.g., SWOP) or subsurface water domains (e.g., GWML2, SWEET-HG), leaving out important flow concepts, such as infiltration, that cross the surface-subsurface boundary. SWEET-HG [23] further models some kinds of flow (e.g., recharge and interflow) as processes, and others (e.g., baseflow) as phenonema, but all of them lack sufficiently detailed representations and none explicitly model the participants.

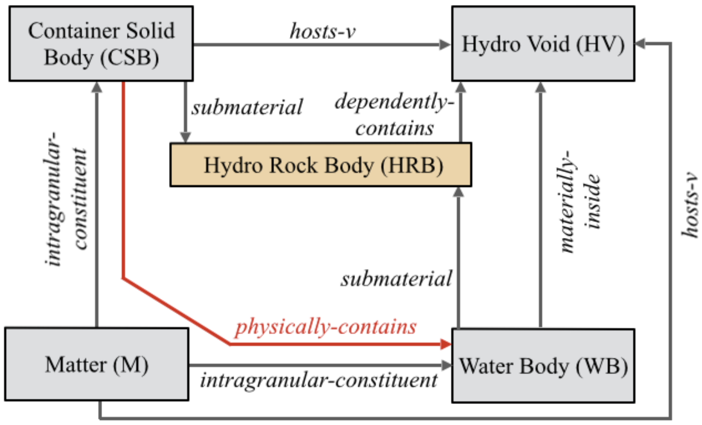


Figure 1 Key HyFO concepts and the physical relationships between them.

3 Background

The formalization that we present in Section 5 builds on prior work on the Hydro Foundational Ontology (HyFO), on DOLCE's upper ontology, and on Allen's interval algebra.

HyFO. HyFO is an effort to develop a domain reference ontology for the hydrology domain that represents key semantics of surface and ground water concepts in a unified framework in first-order logic [4, 13, 14, 16]. At the highest level, it describes the following five types of hydrogeological entities, depicted in Figure 1, and their interrelationships:

- **Physical containers** such as rock formations constituted of solid material and containing empty spaces (voids).
- **Physical void** such as a depression or channel in the ground surface, or microscopic pores between grains of solid matter.
- **Bodies of water** that are either *surface water bodies* (SWB), or *subsurface water bodies* ($SSWB$) [16] typically located in the voids of physical containers.
- **Rock and water matter** that constitute container and water bodies, respectively.
- **Hydro rock bodies** that are rock bodies with water bodies therein (i.e., physical containers with contained water bodies), such as aquifers (a geological unit + the stored water) or rivers (a riverbed + the river's water). *Hydrologic units* (HU) and *hydrogeo units* (HGU) (following GWML2 terminology) are its surface and subsurface variants (see Figure 3)

These hydrogeological concepts are specialized *physical endurants* (using DOLCE's PED concept) and are assigned abstract spatial locations (*space regions*, S, in DOLCE) using the $\mathrm{r}(x)$ function. The configuration of the spatial regions is described using mereological and topological relations from [15, 12] similar to the RCC and 9-intersection relations. The new axioms presented in this work only make use of the relations of parthood $P(x, y)$ (i.e., x is a spatial part of y) and partial overlap $PO(x, y)$ (i.e., x and y share some spatial part), and the intersection operation $(\cdot)$ that identifies the spatial region shared by two regions. HyFO axiomatically relates the hydrogeological entities using foundational physical relations: physical containment [14], constitution, and hosting a void [13]. Most relevant to the work here is $submaterial_t(x, y, t)$, which denotes that at time t, x is a physical part of y whose removal would alter y (i.e., x and y are materially-spatially interdependent). It is a temporally indexed variant of the physical containment relation $submaterial(x, y)$ from [14].

Relevant DOLCE Concepts: Processes and the Participation Relation. Upper-level ontologies such as DOLCE [19] or BFO [21] help ground domain ontologies in formal ontological distinctions and can facilitate interoperability across ontologies [4, 16]. A key distinction in DOLCE is between *endurants* (called *continuants* in BFO) that are wholly present at any time they exist (e.g., the hydrogeological entities from HyFO) and *perdurants* that are necessarily temporally extended objects such as processes or events (called *occurents* in BFO). Perdurants are characterized by the endurants that *participate* in them and by temporal characteristics, such as when an instance of a perdurant starts, pauses or terminates. This is captured by DOLCE's time-indexed *participation* relation $PC(x, y, t)$ that expresses that an endurant x participates in a perdurant y at time t (AD-33 from [19]), further implying that the perdurant y occurs, among other times, at time t. Other types of participation, such as *constant participation* and *temporary participation* are also defined in DOLCE but are of lesser importance here.

(Ad-33) $PC(x, y, t) \rightarrow ED(x) \wedge PD(y) \wedge TR(t)$ ***(x participates in y during t)***

In this paper, we treat hydrological flow processes as perdurants and, more precisely, as processes rather than events in line with [11], assuming that they occur fairly steadily (though possibly variable over seasons), that is, they are typically not time-bounded and can be decomposed spatially or temporally into smaller processes of the same kind (e.g., infiltration over a longer period of time consists of many shorter lasting infiltration processes). We do not consider specific water-related events, such as a specific time-bound event of flooding associated with a specific hurricane or with intense rainfall in a confined location.

Temporal Relations. The temporal parameter in the participation relation refers to instances of DOLCE's temporal region concept $TR(x)$, which encompasses both extended time intervals as well as time points. To temporally compare temporal regions, we rely on $beforeEq(t1, t2)$ as the only temporal relation. It denotes that $t1$ occurs entirely before or at the same time as $t2$, which is satisfied by any of the three qualitative relations "precedes", "meets", or "equals" from Allen's interval algebra [2]. The relation equally applies to time intervals and time points (as special kind of intervals with the same start and end), but is most useful when – as we assume – the start and end of each interval is a time point. Then the existence of overlapping intervals requires the existences of start times that can be properly ordered temporally.

4 Analysis of Physical Endurants in Hydrologic Flow Processes

As a process, hydrological flow specializes DOLCE's concept of a perdurant (HF-A1). Next we analyze common hydrological flow processes discussed in the literature to develop refined participation relations that deal with key hydrological participants.

(HF-A1) $HF(x) \rightarrow PD(x)$ ***(Hydrological flows are perdurants)***

4.1 Kinds of Hydrologic Flow Processes

Fundamental to designing ontologies is identifying key concepts and their intended semantics. Here, refined *hydrological flow* concepts are based on the terminology frm USGS coupled groundwater and surface water model (GSFLOW model) [18] and related models, including the Integrated Water Flow Model (IWFM) [7]. GSFLOW models hydrological flow within and between three coarse geophysical regions: (1) the *ground surface*, which encompasses the topsoil zone and is a mix of rock and organic matter; (2) *water bodies* contained or supported by the *ground surface*, represented in HyFO as *surface water bodies* that includes both standing and flowing water bodies; (3) subsurface zones, including *zones of unsaturation* and *saturation*. The following summarizes the most important types of hydrological flow identified from these models that will guide our analysis and formalization.

Runoff is movement of water above the *ground surface*, typically caused by precipitation or melting of snow and ice. It may take the form of surface runoff, channel flow, or subsurface runoff.

Infiltration is the vertical movement of water through the *ground surface* or the groundwater table.

Overland flow is movement of water where it travels between two points on the *ground surface* without infiltrating the *ground surface*.

Percolation is the movement of water through pores and voids in *subsurface zones* driven by gravity and capillary forces.

Throughflow is the downhill percolation of water through the *zone of unsaturation* under the influence of gravity until it infiltrates the water table.

Channel flow is the movement of water within a river channel.

Recharge and Discharge is the movement of water between two *water bodies*, such as the flow of water from a tributary into another stream segment.

Interflow is the flow of water from the *zone of unsaturation* into a *water body*.

Baseflow is the flow of water from an aquifer into a connected *surface water body*.

Leakage is water moving from a *surface water body* into the connected subsurface rock unit or a *subsurface water body*.

4.2 Modeling Physical Participants in Hydrological Flow Processes

Hydrological endurants evolve continually during hydrological flow processes, e.g., water levels change and containers erode. Despite such physical changes, each flow process relies on a specific set of participating entities. Different kinds of flow processes can be discriminated based on how and what kinds of hydrogeological endurants can participate. This is formalized using semantic roles, which capture and distinguish participants based on their function within events or processes. Semantic roles were originally developed to assign participatory roles to language predicates in English grammar, but more generally identify domain independent thematic roles for endurants that participate in perdurants [10]. We utilize Fillmore's case roles [10] as summarized in [1, p. 93] but other semantic role frameworks, such as Sowa's thematic roles [22], would yield similar formalizations.

Out of the nine semantic roles from [1, p. 93], we have identified the following four as relevant for describing how hydrogeological entities participate in flow processes:

Theme participant is the moving entity that continuously participates throughout the flow process' duration. It always is some amount of water matter, but may include other, smaller amounts of matter, such as sediments, organic materials, or pollutants.

Source participant is the entity from which water moves, such as a subsurface water body from which water seeps into a surface stream.

Goal participant is the entity where the water is moved to, such as the surface water body that receives water from a baseflow process.

Locative participants are physical endurants where a process occurs and that undergoes change in physical or spatial qualities. All source and goal participants are locative participants as well.

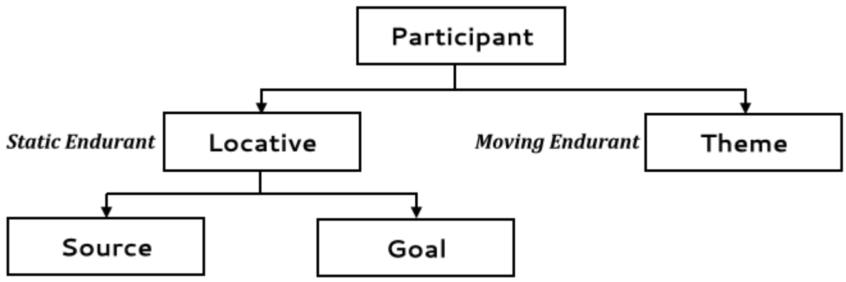


Figure 2 Kinds of participants in hydrological flow derived from Fillmore's case roles.

The reviewed kinds of hydrological flows all transport of water from some source entity to a goal entity. For example, infiltration is a flow process that transports water from the ground surface to some subsurface entity (e.g., a zone of unsaturation or a hydrogeo unit). For example, the infiltration process can consist of distinct temporal stages identified by different participating endurants, such as percolation of water matter through the zone of unsaturation, followed by percolation through a physical container, and followed by entry into a subsurface water body. As such, hydrological flows are inherently spatial processes where water changes location from a *source* to a *goal* participant, both refining the more general concept of a *locative participant*. The transported water matter itself is the *theme* of any hydrological flow process. Temporal aspects in participation arise from the spatial ones and do not need to be modeled separately. Furthermore, hydrological flow generally does not involve intentional agents, except maybe for the designers of flow enhancing, altering, or impeding measures such as dams, culverts, or environmental engineering. But such agents are not directly involved in flow processes but rather in water planning processes.

5 Formalization

The first set of axioms (HPC-A1–10) captures the three kinds of specialized participation relations and ontological and temporal dependencies between them. Section 5.1 captures what static hydrogeological entities can fill the participant roles. Afterwards, the zones of saturation and unsaturation, which are not yet part of HyFO, are axiomatized. Subsequently, we formalize flow within a single hydrogeological entity (intraflow; Sec. 5.3) and across two distinct entities (interflow[1]; Sec. 5.4). All axioms for the ontology in this paper are expressed in full first-order logic, and are available at `https://github.com/gruninger/colore/tree/master/ontologies/hyfo_flow` encoded using the ISO standard Common Logic.

Figure 2 presents the hierarchy of participants distinguished by their roles in *hydrological flow*. These roles are formalized as refined variants of the general participation relation (*PC* from DOLCE) as *TPC*, *SPC* and *GPC* (HPC-A1–3). The term *locative participant* is used to denote either *source* or *goal participants* (HPC-A4). HPC-A5–8 capture basic ontological dependencies between the participants: disjointness of the locative particpants from the theme participant (HPC-A5) and the existence of some source, goal, and theme participant for each flow process (HPC-A6–8), though at possibly different times. HPC-A9 and A10 express basic temporal constraints on the participants: some *goal participant* must exist at the same time or after any *source participant* (HPC-A9) and, vice versa, some *source participant* must exist at the same time or before any *goal participant* (HPC-A10). In reading

[1] This concept is distinct from the term **Interflow** described in Section 4.1.

these two axioms, let the reader be reminder that the temporal parameters denote either extended time intervals or time points, and any time interval at which a participation relation holds can be broken into many contained time points, for which the participation relation must also hold. Thus, HPC-A9 and A10 are satisfied if the last/earliest possible timepoint during a *source/goal participation* satisfies the axiom.

(HPC-A1) $TPC(x, y, t) \rightarrow PC(x, y, t) \wedge HF(y) \wedge PED(x)$ ***(Theme participation)***

(HPC-A2) $SPC(x, y, t) \rightarrow PC(x, y, t) \wedge HF(y) \wedge PED(x)$ ***(Source participation)***

(HPC-A3) $GPC(x, y, t) \rightarrow PC(x, y, t) \wedge HF(y) \wedge PED(x)$ ***(Goal participation)***

(HPC-A4) $LPC(x, y, t) \leftrightarrow SPC(x, y, t) \vee GPC(x, y, t)$
(Locative participation generalizes source and goal participation)

(HPC-A5) $LPC(x, y, t) \rightarrow \neg TPC(x, y, t)$
(Locative and theme participation are disjoint)

(HPC-A6) $LPC(x, y, t) \rightarrow \exists z\, TPC(z, y, t)$
(At any time when something locative participates in a hydrological flow process, then there must also be some theme participant)

(HPC-A7) $HF(x) \rightarrow \exists y, t SPC(y, x, t)$
(Any HF process has some source participant)

(HPC-A8) $HF(x) \rightarrow \exists y, t GPC(y, x, t)$ ***(Any HF process has some goal participant)***

(HPC-A9) $HF(x) \wedge SPC(y, x, t1) \rightarrow \exists z, t2[GPC(z, x, t2) \wedge beforeEq(t1, t2)]$
(Any source participant has a goal participant at the same or a later time)

(HPC-A10) $HF(x) \wedge GPC(z, x, t2) \rightarrow \exists y, t1[SPC(y, x, t1) \wedge beforeEq(t1, t2)]$
(Any goal participant has a goal participant at the same or a later time)

5.1 Hydrogeological Participants

Every *hydrological flow* process requires the involvement of some *source*, *goal* and *theme participants* (HPC-A7,8, together with HPC-A4,6). *Water matter* is the only entity that is transformed by virtue of being moved and becomes the sole *theme participant* (HPC-A11) of hydrological flow, while *geologic units* and *water bodies* stay in place. The movement of water from a *source participant* indicates that it must contain some amount of *water matter* at the beginning of the process (HPC-A13). The most common *source* and *goal participants* are *hydro rock bodies*, which are hybrid entities that consist of a surface or subsurface *water body* and a geologic unit that serves as the physical container. Other important *locative participants* are the *ground surface*, which is a layer of soil and rock and acts as a boundary between surface and subsurface flow (e.g., in infiltration and overland flow), and *zones of saturation* and *unsaturation*, which participate in subsurface flow, e.g., throughflow (HPC-A12). The *zone of saturation* lies within the region of a *hydro rock body* that is a *locative participant*, and is therefure not separately mentioned in HPC-A12.

(HPC-A11) $HF(x) \wedge TPC(y, x, t) \rightarrow WM(y)$
(Water matter is always the theme participant in hydrological flow)

(HPC-A12) $HF(x) \wedge [SPC(y, x, t) \vee GPC(y, x, t)] \rightarrow HRB(y) \vee GS(y) \vee ZOU(y)$
(Source and goal participants in hydrological flow are a HRB, GS, or ZOU)

(HPC-A13) $HF(x) \wedge TPC(y, x, t2) \rightarrow \exists s, t1[SPC(s, x, t1) \wedge submaterial_t(y, s, t1) \wedge beforeEq(t1, t2)]$ ***(Any water matter that is a theme participant is submaterial of the source participant at the same or an earlier time point)***

5.2 Subsurface Zones

Hydrological flow can occur almost anywhere beneath the ground surface. As evident from Section 4.1, a complete categorization of subsurface flow processes requires a more detailed description of two hydrologically distinct regions, the *zones of saturation* and *unsaturation*. These have previously not been represented but can be described using existing HyFO concepts and relations from [16]. The presented formalization closely follows their informal descriptions from GWML2 [3].

The zone immediately below the soil surface containing pore spaces that can potentially accommodate *water matter* is called the *zone of unsaturation* (ZOU), often also referred to as zone of aeration or vadose zone. This zone includes the capillary fringe where the moisture content is less than saturation, that is, water may flow through this zone but does not reside there for extended periods of time. The *zone of saturation* (ZOS) is the zone that lies below the *zone of unsaturation* and is bounded at the top by the water table. The zone of saturation is the region of a *hydrogeo unit* whose void spaces are entirely filled with *water matter* (Z-A1). *Subsurface water bodies* such as those in aquifers are situated in this zone (Z-A2), but the zone excludes regions occupied by *confining beds* (Z-A3). *Zone of unsaturation* is the spatially complementary region of the *ZOS* within a *geologic unit* (Z-A4), meaning they may be spatially connected but do not overlap (Z-A5).

For completeness, we also refine the definition of a *ground surface* from [13] to describe it as a relevant part feature that is hosted by a *geologic unit*.

(Z-A1) $ZOS(z) \rightarrow HGU(z) \wedge \exists c, w[P(\mathrm{r}(z), \mathrm{r}(h)) \wedge CSB(c) \wedge submaterial(c, z) \wedge \neg ZEX(\mathrm{r}(z) \cdot con\text{-}voidspace(c)) \wedge WM(w) \wedge P(\mathrm{r}(z) \cdot con\text{-}voidspace(c), \mathrm{r}(w))]$

(Zone of saturation is a hydrogeologic unit that includes some connected, non-empty voidspace in the unit's container – denoted as the intersection between the zone z and the container's connected voidspace – and the voidspace is completely filled with water matter)

(Z-A2) $SSWB(x) \rightarrow \exists z[ZOS(z) \wedge P(r(x), \mathrm{r}(z))]$

(Every subsurface water body is located in a ZOS)

(Z-A3) $CB(x) \wedge ZOS(z) \rightarrow \neg PO(r(x), r(z))$

(A ZOS does not overlap with any confining bed)

(Z-A4) $ZOU(z) \rightarrow GU(z) \wedge P(\mathrm{r}(z), \mathrm{r}(h)) \wedge \forall y[ZOS(y) \rightarrow \neg PO(\mathrm{r}(z), \mathrm{r}(y))]$

(Zone of unsaturation is a geologic unit that does not overlap a ZOS)

(Z-A5) $GS(x) \wedge [ZOS(z) \vee ZOU(z)] \rightarrow \neg PO(\mathrm{r}(x), \mathrm{r}(z))$

(Zones of saturation or unsaturation do not overlap the ground surface)

(GS-A1) $GS(x) \rightarrow RPF(x) \wedge \exists y[GU(y) \wedge hosts(y, x)]$

(Ground surface is a relevant part feature hosted by a geologic unit)

5.3 IntraFlow

Our preliminary analysis in Section 4.1 suggests a classification based on whether the water flow is confined to a single entity or between entities. This results in the top-most refinement of *hydrological flow* based on whether the participating *source* and *goal participants* are distinct: (1) *intraflow* is flow within a single endurant object, and (2) *interflow* is flow between two endurant objects. Figure 4 illustrates the full taxonomy of the different types of *hydrological flow* processes along with the constraints that we use to hierarchically organize the concepts.

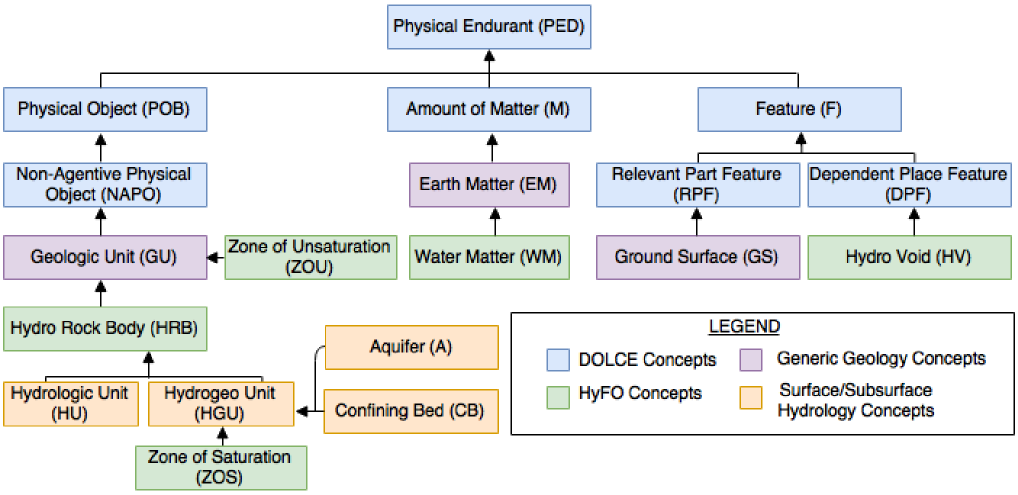


Figure 3 HyFO's geological and hydrogeological concepts that are relevant to hydrological flow.

Intraflow represents the flow of water within a single (hydro)geological endurant, hence the *source* and *goal participants* are identical (HF-A2). This high-level concept captures three types of water movement: (1) the (mostly horizontal) movement of water within the *ground surface*, (2) the movement of water within a single *water body* that is contained in a *hydro rock body*, and (3) the movement of water through the pores and fractures of a *geologic unit* that lacks a water body. Because *source* and *goal participants* are identical, any *water matter* that acts as a *theme participant* in an *intraflow* process remains submaterial of the static *locative participant* over the entire duration of the flow process (HF-A3).

(HF-A2) $intraFlow(x) \rightarrow HF(x) \wedge \forall y,t[SPC(y,x,t) \leftrightarrow GPC(y,x,t)]$
(***The source and goal participants are identical in intraflow***)

(HF-A3) $intraFlow(x) \wedge TPC(y,x,t) \wedge LPC(z,x,t) \rightarrow submaterial(y,z,t)$
(***The water that is the theme participant in an intraflow process is submaterial of the locative participant at any time t during the process***)

Surface-intraflow and ***subsurface-intraflow*** are distinct and disjoint subclasses of *intraflow* (HF-A4–A7), denoting flow processes that occur above/within the *ground surface* and below the *ground surface*, respectively.

(HF-A4) $surfaceIntraFlow(x) \rightarrow intraFlow(x)$ (***Specializing intraflow***)

(HF-A5) $surfaceIntraFlow(x) \wedge LPC(y,x,t) \rightarrow HU(x) \vee GS(x)$ (***The locative participant in surface-intraflow is either a surface HRB (a HU) or the GS***)

(HF-A6) $subsurfaceIntraFlow(x) \rightarrow intraFlow(x)$ (***Specializes intraflow***)

(HF-A7) $subsurfaceIntraFlow(x) \wedge LPC(y,x,t) \rightarrow HGU(y) \vee ZOS(y) \vee ZOU(y)$
(***The locative participant in subsurface-intraflow is either a subsurface HRB, a ZOS, or a ZOU***)

(HF-A8) $intraFlow(x) \leftrightarrow \neg surfaceIntraFlow(x) \vee \neg subsurfaceIntraFlow(x)$
(***Disjoint and exhaustive subclasses of intraflow***)

Overflow specifically describes the lateral flow of water on or within the *ground surface* (HF-A9,10) that does not infiltrate it. This includes surface runoff where flow precipitation or excess water from a surface *water body* flows over the Earth's surface.

(HF-A9) $overFlow(x) \rightarrow surfaceIntraFlow(x)$ (***Specializing surface-intraflow***)

(HF-A10) $overFlow(x) \wedge LPC(y, x, t) \rightarrow GS(y)$
(***The locative participant in an overflow process is the ground surface***)

Water matter may move inside a *water body* or a *hydro rock body* that contains the *water body*. For example, water flows within a river and between different parts of a river (e.g., from one section of rapids to into a more even flowing section) until it eventually gets discharged at the river's mouth or at a junction with another river. This kind of flow occurs above and below the *ground surface*: ***surface-withinflow*** occurs within *water bodies* contained in a *hydrologic unit* (HU), which is a surface *hydro rock body* (HF-A11,12), while ***subsurface-withinflow*** occurs within a *subsurface water body* hosted by a *hydrogeo unit* (HGU), which is a subsurface *hydro rock body* (HF-A13,14).

(HF-A11) $surfaceWithinFlow(x) \rightarrow surfaceIntraFlow(x)$
(***Specializing surface-intraflow***)

(HF-A12) $surfaceWithinFlow(x) \wedge LPC(y, x, t) \rightarrow HU(y)$
(***Locative participant in surface within flow is surface*** HRB***, i.e., a HU***)

(HF-A13) $subsurfaceWithinFlow(x) \rightarrow subsurfaceIntraFlow(x)$
(***Specializing subsurface-intraflow***)

(HF-A14) $subsurfaceWithinFlow(x) \wedge LPC(y, x, t) \rightarrow HGU(z)$ (***Locative participant in a subsurface-withinflow is a subsurface HRB, i.e., a HGU***)

(HF-A15) $surfaceIntraFlow(x) \leftrightarrow overFlow(x) \vee surfaceWithinFlow(x)$
(***Overflow and surface within flow are exhaustive classes of surface-intraflow***)

(HF-A16) $\neg overFlow(x) \vee \neg surfaceWithinFlow(x)$
(***Overflow and surface-withinflow are disjoint classes***)

Through flow is a specialization of *subsurface-intraflow* (HF-A17) that represents flow of *water matter* through a *zone of unsaturation*. Once water infiltrates the *ground surface*, gravity and other forces cause it to move through the unsaturated zone until it eventually reaches the *zone of saturation* or a surface or subsurface *water body*, neither of which are themselves participants in the *through flow* process. This flow depends on the properties of the rock matter constituting the *zone of unsaturation*, such as porosity, permeability and hydraulic conductivity, necessitating that it occurs in porous *geological units*, that is, those with non-empty connected voidspace (HF-A18).

(HF-A17) $throughFlow(x) \rightarrow subsurfaceIntraFlow(x)$
(***Specializing subsurface-intraflow***)

(HF-A18) $throughFlow(x) \wedge LPC(y, x, t) \rightarrow ZOU(y) \wedge \exists z\,[\,GU(z) \wedge \neg ZEX(r(y) \cdot con\text{-}voidspace(z))]$ (***Locative participant in through flow is zone of unsaturation that lies in some porous geological unit***)

(HF-A19) $subsurfaceIntraFlow(x) \rightarrow throughFlow(x) \vee subsurfaceWithinFlow(x)$
(***Exhaustive subclasses of subsurface-intraflow***)

(HF-A20) $\neg throughFlow(x) \vee \neg subsurfaceWithinFlow(x)$
(***Throughflow and subsurface-withinflow are disjoint classes***)

5.4 Interflow

In *interflow* processes, such as infiltration, baseflow, or leakage as described in Sec. 4.1, water flows between distinct *source* and *goal participants* (HF-A21). Often, such flow processes may consist of several heterogeneous temporal parts that are *inter-* or *intraflow* processes, but we epitomize them as single *interflow* processes as long as they occur in spatially connected *source* and *goal participants* (HF-A22). More complex *interflow* processes across multiple participants can be easily composed. The temporal constraints on source and theme participants from HPC-A13 are expanded to *goal participants* for *interflow* by stating that any *theme participant* will eventually become submaterial of a *goal participant* (HF-A23).

(HF-A21) $interFlow(x) \rightarrow HF(x) \wedge \forall s, g, t1, t2[SPC(s, x, t1) \wedge GPC(g, x, t2) \rightarrow s \neq g]$
(*Any source and goal participants in interflow are distinct*)

(HF-A22) $interFlow(x) \wedge SPC(s, x, t) \wedge GPC(g, x, t) \rightarrow C(r(s), r(g))$
(*Source and goal participants in an interflow process are spatially connected*)

(HF-A23) $HF(x) \wedge TPC(y, x, t1) \rightarrow \exists g, t2[GPC(g, x, t2) \wedge submaterial_t(y, g, t2) \wedge beforeEq(t1, t2)]$ **(*Any water matter that is a theme participant is submaterial of the goal participant at the same or a later time point*)**

Three specializations of *interflow* processes are identified: (1) *surface-interflow* where water moves between hydrogeological endurants above the Earth's surface such as surface *hydro rock bodies* or the *ground surface* (HF-A24); (2) *subsurface-interflow* where water moves between subsurface hydrogeological endurants such as *hydrogeo units* and *zones of saturation* (HF-A25); (3) *surface-subsurface-interflow* where water moves between a surface and a subsurface endurant (HF-A27).

(HF-A24) $surfaceInterFlow(x) \rightarrow interFlow(x) \wedge \forall l[LPC(l, x, t) \rightarrow (HU(l) \vee GS(l))]$
(*Surface-interflow is the flow of water between HU's or GS*)

(HF-A25) $subsurfaceInterFlow(x) \rightarrow interFlow(x) \wedge \forall l[LPC(l, x, t) \rightarrow (HGU(l) \vee ZOU(l))]$
(*Subsurface-interflow is the flow of water between HGU's or ZOU*)

(HF-A26) $surface\text{-}subsurfaceInterFlow(x) \rightarrow interFlow(x) \wedge \forall l[LPC(l, x, t) \rightarrow (HRB(l) \vee ZOU(l) \vee GS(l))]$
(*Surface-subsurface-interflow is the flow of water between HRB's, ZOU or GS*)

Prototypical *interflow* processes are *recharge* and *discharge* where water flows into or out of a *hydro rock body* and fills or drains the contained *water body*. Thus, in *recharge* the goal participant is a *hydro rock body* and in *discharge* the source participant is a *hydro rock body*.

6 Summary

The absence of a formal semantic representation for different kinds of *hydrological flow* processes inhibits an integrated view of how water moves and is stored above and below the surface of the Earth. This paper presents a general schema for analyzing *hydrological flow* patterns, identifying *water matter* as a definite participant, but with varying *source* and *goal participants* (see Figure 2). Thus, three refinements of DOLCE's participation relation are proposed to model this formally: the participation of the transported water as *theme participant* and two *locative participants*, namely the *source* and *goal participants* that indicate the hydrogeological entities that loose or gain water.

A taxonomy of common hydrological flow relations, depicted in Figure 4, is developed using the roles and the participating physical endurants as described by HyFO (see Figure 3). The highest-level distinction between hydrological flow patterns is based on whether water moves within a single *locative participant* (*intraflow*) or between two distinct *locative participants* (*interflow*). *Intraflow* is further specialized based on the single participant: whether it is the *ground surface* (e.g., overflow), a surface *hydrologic unit* (e.g., channel flow), a subsurface

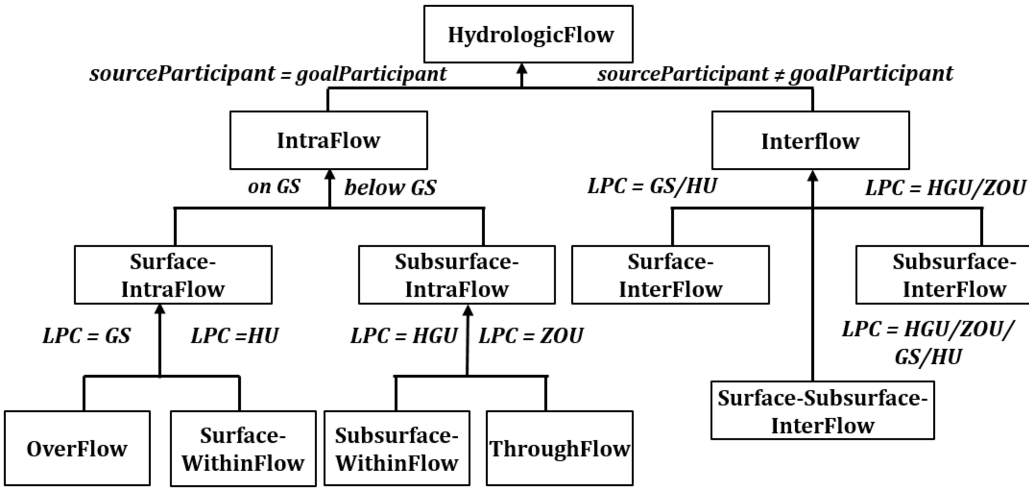


Figure 4 The taxonomy of *hydrological flow* concepts.

hydrogeologic unit (e.g., flow within an aquifer), or a subsurface zone (e.g., throughflow, percolation). Similarly, *interflow* can be distinguished based on three combinations of *source* and *goal participants*: (1) both are surface features (one possibly the *ground surface*), (2) both are subsurface features (a hydrogeological endurant or a zone), (3) one is a surface and the other is a subsurface feature (e.g., infiltration, leakage, and base flow).

The different flow processes have been formalized only to the extent necessary for communicating the taxonomy's basic distinctions and associated temporal constraints. It is a proof-of-concept that shows how one could extend HyFO with dynamic hydrological aspects. More work is required to complete the formalization of all discussed flow concepts and to test the formalization's internal consistency and its consistency with how hydrological flow is represented in related hydro data models and ontologies.

References

1 Bas Aarts. *English Syntax and Argumentation.* Palgrave & Macmillan, 3rd edition, 2008.

2 James F. Allen. Maintaining knowledge about temporal intervals. *Comm. ACM*, 26(11):832–823, 1983.

3 Boyan Brodaric. GroundWaterML2 – GW2IE Final Report. Technical report, Open Geospatial Consortium, 2015. Engineering Report 15-082, version 2.1.

4 Boyan Brodaric and Torsten Hahmann. Towards a foundational hydro ontology for water data interoperability. In *11th Int. Conference on Hydroinformatics (HIC-2014)*, 2014.

5 Walter A. Cook. *Case grammar: development of the matrix model (1970–1978).* Georgetown University Press, 1979.

6 Anusuriya Devaraju and Werner Kuhn. A process-centric ontological approach for integrating geo-sensor data. In *International Conference on Formal Ontology in Information Systems (FOIS-10)*, pages 199–212. IOS Press, 2010.

7 Emin C. Dogrul. Integrated water flow model (IWFM v3.1): Theoretical documentation. Dept. of Water Resources, State of California, 2009.

8 Irina Dornblut and Robert Atkinson. Hy_features: a geographic information model for the hydrology domain. Technical Report GRDC 43r1, Global Runoff Data Centre, 2013.

9 Chen-Chieh Feng, Thomas Bittner, and Douglas M. Flewelling. Modeling surface hydrology concepts with endurance and perdurance. In *International Conference on Geographic Information Science (GIScience-04)*, pages 67–80. Springer, 2004.

10 Charles J. Fillmore. The case for case. In Emmon Bach and Robert T. Harms, editors, *Universals in linguistic theory*, pages 1–88. Holt, Rinehart and Winston, 1968.

11 Antony Galton. Outline of a formal theory of processes and events, and why GIScience needs one. In *International Conference on Spatial Information Theory (COSIT-2015)*, pages 3–22. Springer, 2015.

12 Torsten Hahmann. *A Reconciliation of Logical Representations of Space: from Multidimensional Mereotopology to Geometry.* PhD thesis, Univ. of Toronto, 2013.

13 Torsten Hahmann and Boyan Brodaric. The void in hydro ontology. In *International Conference on Formal Ontology in Information Systems (FOIS-12)*, pages 45–58. IOS Press, 2012.

14 Torsten Hahmann and Boyan Brodaric. Kinds of full physical containment. In *International Conference on Spatial Information Theory (COSIT-13)*, pages 397–417. Springer, 2013.

15 Torsten Hahmann and Michael Grüninger. A naïve theory of dimension for qualitative spatial relations. In *Symposium on Logical Formalizations of Commonsense Reasoning (CommonSense 2011) at the AAAI Spring Symposium (AAAI-SS 2011)*. AAAI Press, 2011.

16 Torsten Hahmann, Shirly Stephen, and Boyan Brodaric. Semantically refining the GWML2 with the help of a reference ontology. In *International Conference on Geographic Information Science (GIScience-16)*, 2016.

17 INSPIRE Thematic Working Group Hydrography. D2.8.I.8_v3.1 INSPIRE Data Specification on Hydrography – Technical Guidelines. Technical report, INSPIRE, 2014.

18 Steven L. Markstrom, Richard G. Niswonger, R. Steven Regan, David E. Prudic, and Paul M. Barlow. GSFLOW-coupled ground-water and surface-water FLOW model based on the integration of the precipitation-runoff modeling system (PRMS) and the modular ground-water flow model (MODFLOW-2005). Technical report, Geological Survey (US), 2008.

19 Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, and Alessandro Oltramari. Wonderweb deliverable D18 – ontology library (final report). Technical report, National Research Council – Institute of Cognitive Sci. and Technology, Trento, 2003.

20 Gaurav Sinha, David Mark, Dave Kolas, Dalia Varanka, Boleslo E. Romero, Chen-Chieh Feng, Lynn E. Usery, Joshua Liebermann, and Alexandre Sorokine. An ontology design pattern for surface water features. In *International Conference on Geographic Information Science (GIScience-14)*, pages 187–203. Springer, 2014.

21 Barry Smith et al. *Basic Formal Ontology 2.0 Draft Specification Guide*, 2012.

22 John F. Sowa. *Knowledge Representation: Logical, Philosophical, and Computational Foundations.* Brooks Cole, 1999.

23 A. Tripathi and H. A. Babaie. Developing a modular hydrogeology ontology by extending the SWEET upper-level ontologies. *Comput. Geosci.*, 34(9):1022–1033, 2008.

Classification, Individuation and Demarcation of Forests: Formalising the Multi-Faceted Semantics of Geographic Terms

Lucía Gómez Álvarez[1] and Brandon Bennett[2]

1 **School of Computing, University of Leeds, Leeds, UK**
sc14lga@leeds.ac.uk

2 **School of Computing, University of Leeds, Leeds, UK**
B.Bennett@leeds.ac.uk

Abstract

Many papers have considered the problem of how to define *forest*. However, as we shall illustrate, while most definitions capture some important aspects of what it means to be a forest, they almost invariably omit or are very vague regarding other aspects. In the current paper we address this issue, firstly by providing a definitional framework based on spatial and physical properties, within which one can make explicit the implicit variability of the natural language forest concept in terms of explicit parameters. Our framework explicitly differentiates between the functions of *classification*, *individuation* and *demarcation* that comprise the interpretation of predicative terms. Whereas ontologies have traditionally concentrated predominantly on classification, we argue that in many cases (especially in the case of geographic concepts) criteria for individuation (i.e. establishing how many distinct individual objects of a given type exist) and demarcation (establishing the boundary of an object) require separate attention, involve their own particular definitional issues and are affected by vagueness in different ways. We also describe a prototype Prolog system that illustrates how our framework can be implemented.

1998 ACM Subject Classification I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Forest, Definition, Vagueness, Ontology, GIS

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.8

1 Introduction

There are many definitions of forest in the literature [29, 9]. While most of these capture some important aspects of what it means to be a forest, they almost invariably omit or only very loosely specify other important requirements. In particular, constraints on the overall size, shape and topology of a forest are often omitted. In the current paper we address this issue, firstly by providing a framework within which one can give formal definitions which combine both spatial and physical properties; and secondly, by making explicit the implicit variability of the natural language forest concept in terms of explicit parameters; and thirdly, by differentiating between the functions of *classification*, *individuation* and *demarcation* that comprise our interpretation of predicative terms. Whereas formal definitions have traditionally concentrated predominantly on classification, we argue that in many cases criteria for individuation (i.e. establishing how many distinct individual objects of a given type exist) and demarcation (establishing the boundary of an object) require separate attention and involve their own particular definitional issues, particularly in the domain of geographical objects. In the current paper we propose the use of a *supervaluation semantics*, which we

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 8; pp. 8:1–8:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

suggest provides a natural and flexible framework within which to represent the wide variety of different interpretations of the term 'forest'.

2 Defining 'forest'

How much forest is there in the world? This is a surprisingly difficult question to answer [1]. A broad range of forest concepts and definitions (more than 600 were reported in [29]) have been specified for different purposes, thereby leading to very different estimates [19]. Thus, it is not surprising that confusion arises over both global forest extent and its spatial distribution [17, 18].

Traditionally, two main categories of forest definitions have been discussed in the literature: *land cover* and *land use* [28]. While the former defines forest in terms of the ecological layer and the physical characteristics of the land, the latter does it with regard to the purpose to which the land is put to use by humans [32]. Definitions favouring one or another (or both) approaches, together with other relevant features, are linked to different perspectives and management objectives, the most relevant ones being: timber management, conservation of ecosystems, increasing carbon stocks and landscape restoration [9]. For example, definitions used for the analysis of carbon stocks generally focus on land cover, but ignore aspects like connectedness or distinctions between natural or planted forests because they are not relevant to describe the carbon potential. The opposite happens when defining forest for landscape restoration purposes, where together with land use information, they become crucial aspects to understand the effects on ecosystem services and forest-based livelihoods. These differences are linked to scale and disciplinary compartmentalization and, besides responding to the purpose for which the definitions were created [9], they pose limitations on the construction of global knowledge [20] and data interoperability.

Moreover, beyond the semantic ambiguity of the concept itself, reflected in a wide variety of specifications, most definitions of forest (and other geographic features), both in the academic literature and in administrative regulations, are not precise [26, 34] which questions consistency even within a particular research community or monitoring project. This limits understanding of data and may impair management decisions or distort research findings.

Awareness of these issues exists and there is an extensive literature reporting it, both in academia [7, 6, 29] and in policy [25, 14]. However, global agreement on the meaning of such words has not been reached. Part of the community has focused on precisely defining forest together with other natural resource terms. Other research has focused on examining the reasons for definitional problems. Many papers have advocated accommodating a variety of definitions [9] and pointed to consequent challenges, particularly for data integration and multidisciplinary work [20]. The current investigation follows these lines and proposes a framework that enables the coexistence, analysis and comparison of different precise definitions of forest.

3 Forest definitions from a logical and ontological point of view

Concepts and definitions of forests and other geographical features have received wide attention in the domain of Knowledge Representation, Geographical Information Science and Philosophy, and the pervasive vagueness that affects them has been highlighted [13, 15]. In this section we first briefly discuss different aspects of vagueness. We then introduce supervaluation semantics and discuss its ability to express the variety of possible meanings of concepts, which will provide us with a flexible framework for implementing forest definitions.

Finally, we analyse how the research on ontology and, more specifically, on geographical ontology, can contribute to the discussion on appropriate forest definitions, providing insight into aspects of geographical features which are often overlooked on most characterisations.

3.1 Vagueness and logics of vagueness for the study

The topic of vagueness has a long philosophical history, dating back nearly two and a half thousand years [38]. Although not often explicitly mentioned in the literature, it is important for our approach to make the distinction proposed in [5] between *sorites* vagueness and *conceptual* vagueness:

Sorites[1] *vagueness* occurs when the applicability of a predicate depends on specific measurable parameters but their thresholds are undetermined, thus creating borderline cases. For example, if we define forest as a large expanse densely populated by trees, there is no exact specification of exactly how many trees it must contain. In fact, giving a precise definition seems to be contrary to the way that the term 'forest' is used. It would seem very odd to claim that a particular number of trees is insufficient to form a forest, whereas if there were one more we would have a forest, as described by *Fisher* in [15].

Another kind of vagueness arises when there is a lack of clarity on which attributes or conditions are essential to the meaning of a given term, so that it is controversial how it should be defined. Thus, there is indeterminism regarding to which property or logical combination of properties is relevant to determining whether a concept is applicable. We call this *conceptual vagueness.* It is akin to ambiguity in that it occurs where a term has multiple meanings that are qualitatively distinct, except that with ordinary ambiguity the different meanings apply to completely different objects or situations, whereas with conceptual vagueness there is considerable overlap between the sets of cases to which the different meanings apply. This kind of vagueness underlies the controversy about whether to define forest in terms of land cover or land use, and this case is a good illustration of the overlapping of the applicability of the two interpretations, since the use to which land can be put depends to a large extend on the material and ecological properties of its land cover.

In order to provide formalisations of vague predicates and vague concepts we follow the lines of [3] by taking a supervaluationistic approach. Supervaluation semantics is based on the idea that a vague language can be interpreted in many different precise ways, each of which can be logically conceptualised in a precisification, which determines precise truth conditions for each predicate of the language. By incorporating a set of many possible precisifications into the semantics, a supervaluation model can accommodate major differences between perspectives of multiple disciplines and management objectives as well as more minor variations in the use of vague terms. Although the truth of propositions describing properties of a situation will typically vary from precisification to precisification, there will be some propositions that are true within the whole set of what are considered to be reasonable (a.k.a. *admissible*) propositions. Such propositions may be called *supertrue* and constitute a common consensus of accepted facts.

3.2 From an Ontological point of view

The value of ontologies in developing advanced information systems is now well established [21]. A formal ontology can provide both a solid conceptual layer to clarify the forest

[1] Referred to a puzzle known as The Heap: Would you describe a single grain of wheat as a heap? No. Two grains? No. ... You must admit the presence of a heap sooner or later, so where do you draw the line? (Stanford Dictionary of Philosophy).

definitions and computational support for the integration of information coming from remote sensing and traditional survey data. More importantly, it gives us a framework within which we can systematically compare formal forest definitions and their spatial projections by making implicit assumptions as explicit as possible.

One of the main advantages of ontologies is that they improve the interoperability of the information systems that use them, acting to enforce a consensus view reached by a community regarding a certain domain. By formalising the semantics of terminology, ontologies provide a well defined framework for the interchange of information between different systems. Moreover, logic-based ontologies support automated reasoning (the ability to infer logical consequences) over the formal definitions and axioms of the domain. These advantages have been reported for ontologies in information systems [31, 39] in general and in GIS [16, 27] in particular.

In this paper we suggest that, by enriching a Forest Ontology with supervaluation semantics, we provide the basis for a flexible, adaptive and unbiased representation of vague terms, such as *Forest* or *Tree*. This brings the opportunity of moving from the need for a (more or less) consensual and precise but artificial definition of a vague term like *Forest* to a more realistic formalisation capturing a whole set of views coming from multiple disciplines and stakeholders. This is, overall, a first step towards a framework explicitly supporting the coexistence of different forest definitions within information systems. Moreover, the use of supervaluation semantics, which largely preserves the inferential capabilities of first order logic, should facilitate reasoning both within particular precise interpretations and with regard to comparisons between different interpretations.

Moreover, in order to provide a sound characterisation of forests we can learn from the literature on formal ontology in general and in the geographic domain in particular. Given the complexity of the geographic space, its special characteristics [12] and the variety of 'things' it can include [33, 36], much discussion has been raised when trying to answer concisely the question of what is a geographic concept [35]. Some of the most interesting characteristics pointed by ontologists and geographic information scientists include location, topology, boundaries and mereology. It is surprising thus that when defining forest, little attention has been paid to those aspects.

Among the most challenging issues affecting forest definitions are the following:

- *Existence of vague objects.* There is an ongoing philosophical discussion on whether the vagueness exhibited by geographic names and descriptions is merely linguistic or ultimately ontological, that is, its terms are vague because they refer to vague objects, objects with fuzzy boundaries [37]. Beyond the philosophical discussion, fuzzy and crisp representations of forest have different uses both in academia and research.
- *The dichotomy of the object-field representations.* In order to represent geographic phenomena, ontologies have to encapsulate not only the meanings linked to specific concepts but also the way these meanings are handled and represented [2]. Thus, a precisification of a concept such as forest must embed information about its mode of specification, typically either in terms of an object model or a field.
- *Scale and granularity.* A conceptualization of geographic space may have several levels of granularity, each of which will be appropriate for problem solving at different levels of detail [13].
- *Endurants and perdurants.* Endurants are entities that persist through time and are regarded as wholly present at each moment of their existence (matter and objects), while perdurants are never fully present at any one given moment in time, but instead 'unfold' themselves in successive phases or temporal parts (processes and events) [2]. Although

geographic entities are usually described as endurants, they may also be considered as processess. This is especially the case for forests, which undergo continual change, as trees grow and die.

- *Individuation criteria.* How are entities such as mountains, rivers and forests individuated within a landscape? Although the possession of a boundary is one mark of individuality, in the geographical domain boundaries give rise to a number of ontological conundrums and may themselves be difficult to individuate [8].
- *Topology, Mereology and Location.* In the geographic space, topology is considered to be first-class information, whereas metric properties, such as distances and shapes, are used as refinements that are frequently less exactly captured [13]. A general theory of spatial location is necessary to relate an entity with the spatial region that it occupies and, finally, topology is crucial as mereology alone cannot account for some very basic spatial relations, such as the relationship of continuity between two adjacent objects or the relation of one thing being entirely inside or surrounding some other thing [8].
- *Identity over time.* A simple model of composite objects as mathematical sets (where sets are identical if and only if they have exactly the same elements) does not account for more complex unity and identity criteria allowing one to accept the continued existence of geographic objects even after the loss of or gain certain parts [8].

4 A framework for the formalisation of forest definitions

In order to provide a precise account of what a forest is, many aspects of its nature must be specified. We propose a framework that compiles a collection of relevant features to characterise a forest, which tries to unify different perspectives. The framework, shown in Table 1, is by no means exhaustive.

Moreover, we divide these aspects into three main groups, using the notions of classification, individuation and demarcation. We suggest that there are a large number of natural language predicative words and phrases for which a reasonably clear distinction can be made between some or all of these different aspects, and each has different ways of being made precise. Thus, we review the general notion of individuality as it has been discussed in Philosophy, complemented by the notions of identity and unity [22, 23] and by certain existential conditions. We discuss the difference between *individuation* and *classification*, and how this affects information modelling for in GIS and Forest monitoring. We propose a framework that clearly differentiates these aspects and also distinguishes the aspect of *demarcation*, which is particularly relevant for GIS. We suggest that the use of these notions as guiding tools for building ontology-driven GIS can provide clarity on the ontological commitments assumed in research and forest monitoring.

We finally show that in considering precisification of a vague predicate, these aspects are often seperable, in that it may be possible to make one or two of the aspects precise without committing to a precisification of the other aspects(s). For example we might individuate and classify a forest entity without demarcating a precise boundary; or we might demarcate a region of vegetation land cover without committing to whether it should be classified as a forest. We argue that such distinctions are necessary to clarify the ontological commitment made when using vague concepts in research.

5 Main purposes of forest definitions

The features included in the framework in Table 1 are grouped in three main categories, namely classification, individuation and demarcation. These categories refer to three main

Table 1 Compilation of forest definitions from different sources.

Aspects of forest concept definitions	
1. Classification	
1.1. Qualitative characteristics (global to the whole object)	A typical example would be the land use
1.2 Presence (or absence) of features	E.g. roads, trees of more than 5m, shrubs, ...
1.3 Density, uniformity and scale of features	Measures like the canopy cover should be measured not only in terms of the density but also the uniformity and or scale, given that the predicate can be applied to regions with different extensions.
1.4 Location restrictions	Some definitions are contextualised in one area, like tropical forest
2. Individuation	
2.1 Morphological restrictions.	Such as shape or minimal area
2.2 Metrical restrictions	We may want to evaluate the proximity of constituents
2.3 Topological restrictions	Is the forest necessarily self-connected? Does the forest have holes?
2.4 Mereological restrictions	Is the forest the same forest (whole) if it loses one part?
2.5 Rough location	Part of the identity of the object is linked to its geographical rough position
3. Demarcation	
3.1 Fine grained threshold	Determines the precise boundary of the forest
3.2 Fuzzy threshold?	We may allow for fuzzy boundaries

purposes of definition identified as particularly relevant for the field of naive geography. Moreover, our understanding of *classification* and *individuation* match with the general notions used to evaluate and validate ontologies suggested in [23], and the additional category of *demarcation* is specific to the geographical domain of the current research.

5.1 Questions from a naive geographer

We consider three simple questions using the term forest and the different aspects of its meaning to which they relate. These questions are used as guidelines to link cognitive conceptualizations of the geographical space with relevant notions in the domain of philosophy and ontology, as well as with actual research questions around the topic of global forest monitoring.

(a) How much forest is there?
(b) How many forests are there in this region?
(c) What area is occupied by *this* forest?

Question (a) should be interpreted as *How much forestland is there?*, where the mass noun 'forestland' is typically interpreted in terms of a field conceptualisation of the geographical space. A possible answer to that question can be found in [24], where the global land-area is organised as a grid of land-pieces which are systematically classified in terms of the canopy cover (it must be stressed that the classified entities in this case are not forests but land parcels). A relevant precisification of 'forest' or 'forestland' for this study is exclusively concerned with fixing the characteristics that a piece of land needs to display in order to satisfy this classification.

Question (b), however, requires the *individuation* of forests in order to be able to count them, thus taking an object model approach. Characterising individuation criteria of objects is hard and, as seen in Table 1, relies strongly on spatial (and also temporal) factors. As we can see in [9], identity criteria are necessary in order to characterise forest for many management objectives. For example, for landscape restoration purposes and conservation of natural ecosystems, it is important to understand the dynamics of individual forests, and to track whether they merge, split, appear or disappear, even if the global amount of forestland is preserved.

Question (c) asks for the demarcation (or extension) of *a forest.* To give a precise answer, appropriate thresholds and footprint algorithms need to be selected. Demarcation criteria is particularly relevant for assessing forest loss and gain among others. It is often presumed that forests need to be demarcated in order to answer question (b). However, as will be shown in section 5.4, this need not necessarily be the case: we may be able to differentiate (and hence count) forests, without fully establishing their exact boundaries.

The ability of an intelligent agent to interpret these questions and select the relevant aspects of forest definitions, relies on a capacity to understand that geographic terms may present different aspects of a complex multi-faceted semantics in different contexts. In other words, it requires certain understanding of a naive geography.

5.2 Classification

Almost all predicates, and certainly all the notions of forest, incorporate some kind of classification. In this paper we consider that an object x is classified under a predicate ϕ if it satisfies the necessary and sufficient conditions that govern ϕ's applicability. This can be formalised as

$$\forall x[\phi(x) \rightarrow \Phi(x)] \text{ and } \forall x[\Psi(x) \rightarrow \phi(x)] (\text{where } \phi \text{ does not occur in either } \Phi \text{ or } \Psi)\,.$$

Here we understand that the predicate ϕ does not carry identity or unity criteria on its own. Rather, any object denoted by x to which ϕ may be applied is a member of a domain of individuals that has already been fixed (either by direct stipulation of the domain or by axioms involving other predicates). Following the nomenclature in [23] we express it as ϕ**-O-U**. Moreover, classificational predicates are either semi-rigid -R or anti-rigid ~R[2].

Common examples of classification tasks in the geographic domain include both the assignation of a category to an already individuated geographical object, such as classifying a particular forest *forest#23* into a forest type *tropicalForest*(*forest#23*) or a tree *tree#5* into its species *oak*(*tree#5*), and the assignation of a category to a portion of a mass term, typically a certain region or land area *landpiece*#8, for example into *forestland*(*landpiece#8*). The latter, which focuses on the properties that characterise whether the concept *foresthood* is applicable to a given land parcel, is certainly the one that has received more attention within the Forestry literature. This kind of characterisation does not incorporate any specification of individuality; and it seems this is not required to answer the naive question (a). It assumes that an appropriate division of land into parcels has already been made (e.g. as raster cells) and characterises 'forest' or 'forestland' as a mass term. Thus, the predicate is not concerned with forest *objects* as such and does not provide a means to answer the question of how many

[2] The notion of rigidity is introduced in [23]: 'A property is rigid if it is essential to all its possible instances. There are also properties that are essential to some entities and not essential to others (semi-rigid), and properties that are not essential to all their instances (anti-rigid)'

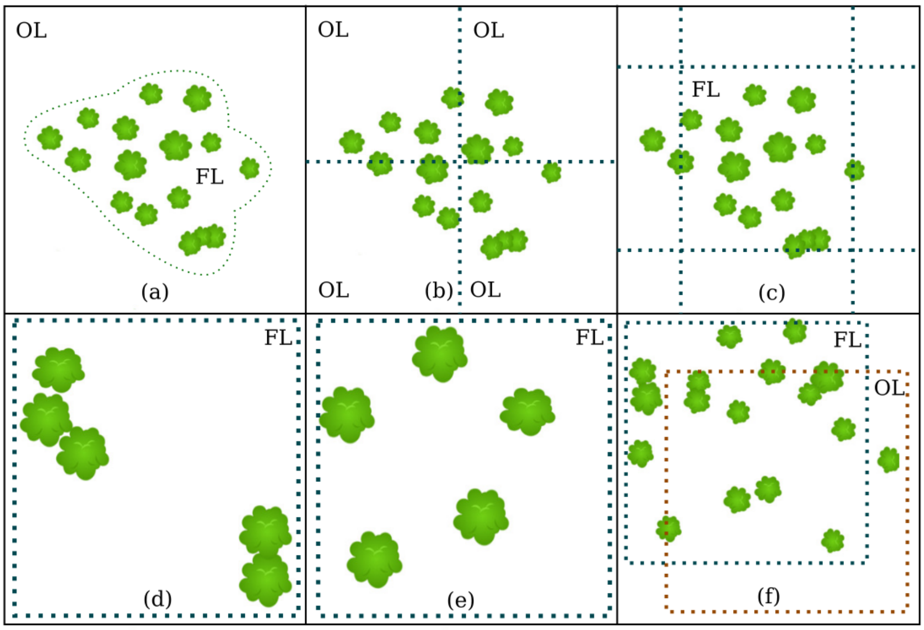


Figure 1 Different land parcels (dashed lines) into forestland (FL) or otherland (OL) following FAO's requirement of a 10% of canopy cover.

forests there are. Moreover, although ϕ can be used to determine the total area of forestland over the entire domain under consideration, it does not determine the extension of individual forests, and, in some cases, the total area of forestland will be different from the total area of the forests contained, for example when some parcels of forestland are isolated from any major forest.

5.2.1 Classificational characterisation of forests

In order to provide a background for any specification of the classification criteria for a precisification of *forestland*, some basic characteristics are discussed below, especially those relating to the points listed in Table 1.

Consider the task of classification applied over portions of the mass concept of land. In the framework proposed in Table 1 we consider certain aspects of the classification that tend to be overlooked even in attempts to provide concise definitions, such as [14]. Following the framework, a precisification of forestland combines, first, (1.1) qualitative attributes, such as the legal land use of the area. Then, (1.2) the presence of certain features, such as the trees and the absence of other elements such as roads or buildings (these classificatory features may of course make reference to other kinds of object or land cover defined in the ontology, which in turn may be also subject to issues of vagueness and of finding an appropriate individuation). Following, (1.3) the density, uniformity and scale of features. As illustrated in Figure 1 and described below, differences in scale and uniformity may result in substantial variations in the meaning of the categories. Finally, (1.4) some location restrictions (e.g. the area must be within the tropics for *tropical-forestland*(x)) may be added in order to improve contextual adaptation.

Figure 1 shows some classifications of pieces of land in terms of their canopy cover, set to the 10% as required in the FAO definition of forest [14]. Figures (a), (b) and (c) display the same piece of land. (a) shows a common sense demarcation and, while (b) and (c) classify that same land using a grid of the same size, both result in different information: (b) is fully covered by 'otherland' (OL) while (c) is mostly covered by 'forestland' (FL). Similarly, two pieces that greatly overlap can be classified as different land types, as in (f). A change in scale such as from (c) to (e) implies a gain on precision but also a variation on the meaning of the 'forestland'. In an extreme scenario, a single tree could constitute 'forestland', which could be misleading. Finally, uniformity issues can arise when the density is concentrated in clusters (d) instead of being evenly distributed across the piece of land (e). This could lead to the confusion of open forests with small and isolated clusters of trees in grasslands.

While any of the previous observations challenges the application of these techniques, they highlight the level of uncertainty of the produced information and inform of the need for the specification of the admissible scale and uniformity requirements for the classification to be meaningful under a certain precisification. A more detailed account of the ontological issues involved on the mapping of land cover is provided in [10].

5.3 Individuation

The notion of individuality is fundamental in the study of ontology and essential when adopting an object model. However, formally characterising the full criteria for the individuation of particular classes of object tends to be extremely hard [23] and is not addressed in the majority of actual ontologies. Studies in Cognitive Science show that humans identify and individuate objects using at least three sources of information: spatiotemporal information, property (featural) information, and sortal information [40]. Moreover, among them, spatial features such as shape are typically more salient than other properties [40].

Within the philosophical literature it is considered that individuation requires both identity and unity, where the former is related to the problem of distinguishing a specific instance of a certain class from other instances (by means of a characteristic property unique to that object) and the latter is related to the problem of distinguishing the parts or constituents of an instance from the rest of the world, which are bound together creating a whole. We now propose a particular form of existence condition to axiomatically specify the domain of individuals that are instances of a given concept.

Existence conditions differ from classification in that the latter express the necessary and sufficient conditions for an object to be an instance of a class while the former explicitly specify the necessary and sufficient conditions to infer the existence of an object. Below is a constructive existential axiom that specifies that whenever a set of conditions $\Phi(x_1, \ldots, x_n)$ are satisfied for some original concepts $x_1, \ldots, x_n$, then an object of kind K exists and a relation holds between the original group and the existent object $\Psi(x_1, ..., x_n, y)$.

$$\forall x_1 ... \forall x_n [\Phi(x_1, \ldots, x_n) \rightarrow \exists y [K(y) \wedge \Psi(x_1, \ldots, x_n, y)]],$$
$$\forall y [K(y) \rightarrow \exists x_1 ... \exists x_n [\Phi(x_1, \ldots, x_n) \wedge \Psi(x_1, \ldots, x_n, y)]].$$

The identity criteria I of a concept determine the conditions under which it can be established that two references refer to the same object, that is, the characteristics that are unique to a single specific instance [22].

$$\forall x \forall y [(K(x) \wedge K(y)) \rightarrow (I_k(x, y) \leftrightarrow (x = y)).$$

Finally, the notion of unity refers to the problem of describing the parts of objects and the specific conditions (UC) under which the object constitutes a whole. A general axiomatic characterisation of this, in terms of a unifying relation among the parts of a whole is given in

[23]. In modelling a particular domain or type of object, it is likely that more specific unity criteria will be required. For instance, a forest may be regarded as a spatially connected region of forested land, which is of maximal extent (i.e. is not part of a larger spatially connected forested region). Thus, assuming a predicate Forested has been defined and applied to all parcels of forestland, then the following axiom (where P is the parthood relation and SCON is the property of being spatially self connected) captures a unity condition for a possible precisification of forest:

$$\mathsf{Forest}(x) \rightarrow \mathsf{Forested}(x) \wedge \mathsf{SCON}(x) \wedge \neg\exists y[\mathsf{P}(x,y) \wedge \neg(x = y) \wedge \mathsf{Forested}(y) \wedge \mathsf{SCON}(y)]\,.$$

Another approach to the unity of forests becomes available if we have forest detailed information at the level of the location of individual trees. We can then define a forest as a maximal collection of proximal trees, parameterised by some threshold of proximity. This is the approach taken in our prototype software implementation, which will be described in Section 6.

A variety of situations can hinder the specification of a criteria for identity and unity. Some of them are drastic evolutions of objects through time, situations in which objects merge or split and objects whose boundaries are ill defined or affected by sorites vagueness thus creating confusion about self-connectedness and parthood.

Underlying axiomatisations of the space and mereotopology are key to provide appropriate notions of parthood. In some cases, a set theoretical view where two sets are the same if and only if they have exactly the same elements is appropriated to model the space. However, for most objects a looser identity criteria that allows one to accept the continued existence of an object even after the loss of certain sorts of parts is necessary.

5.3.1 Individuation of geographical objects

The consideration of the individuation of forests entangles in the extensive bibliography about the ontology of geographical features, their characterisation and their boundaries. Difficulties tend to arise both regarding the unification of geographical features (e.g. deciding whether something is part or not of a forest) and their indentity (e.g. deciding whether a forest now is the same forest as one that existed 100 years ago) particularly if there have been substantial changes in vegetation or location [4]. Moreover, while most of the objects in the physical world have a *bona fide* boundary that acts as one of the main marks of their individuality, geographical boundaries are often fuzzy or otherwise indeterminate [8], which makes the individuation even more challenging and the demarcation of most geographical objects non trivial.

It may seem that a demarcation is required in order to individuate a feature such as a *forest*. However, this is not necessarily the case. We show how individuation is possible without committing to a specific boundary, particularly in cases in which we don't encounter borderline cases. We first describe the example of *mountain*, widely used in the literature, to show how individuation and demarcation are not necessarily related. We then analyse the case of *forest* and discuss how in certain cases it may be appropriate to use different thresholds for the same parameters in the different modes of predication. Although *mountain* and *forest* involve very different attributes, the two cases are analogous in that individuation can be achieved prior to demarcation, and, if needed, a demarcation can subsequently be obtained by further precisification.

In a prototypical case of *mountain* individuation, the main characterisation for both the existence and the individuation can be done in terms of the peak and a minimum

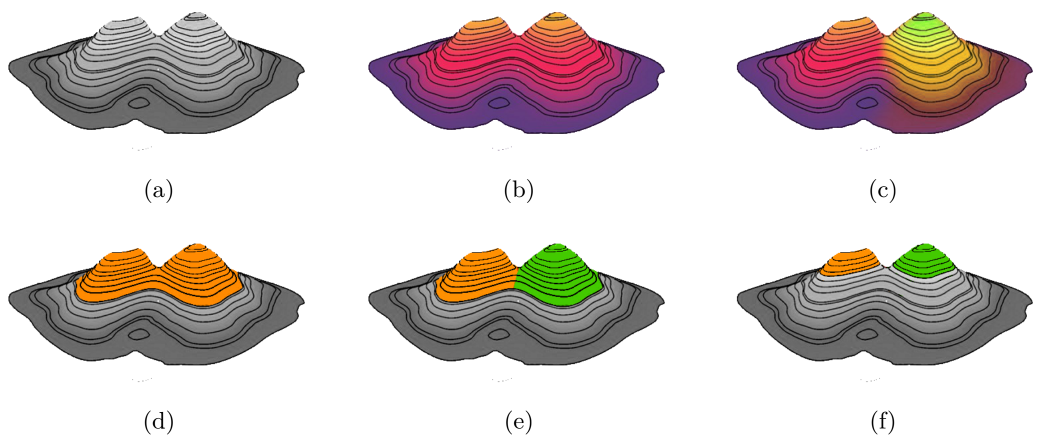


Figure 2 Mountain individuation and demarcation. Each individual is shown with a gradient of different colours, where the colour on the peak (orange and green) are the ones characterising the object and used in the demarcated images (d), (e) and (f).

prominence[3]. Thus, in Figure 2 we see (a) a landscape in which we can identify, depending on the precisification, (b) one or (c) two mountains. In (b) the minimum prominence required is not satisfied and thus the whole ensemble is assumed to be a single mountain, and the contrary occurs in (c). It is not until (d), (e) and (f) that we specify a demarcation strategy. (d) and (e) apply the same demarcation strategy to the two different individuations (b) and (c), and an alternative demarcation for (c) is shown in (f). The identity criteria for the mountain is to share the highest point, that is, the peak. In terms of unity, mountain can be characterised as a self-connected whole extending from the peak to whichever precise boundary.

The identification of *forest* without relaying on the demarcation is harder because the most salient features are its shape, size and parts. However, as we can see in Figure 3 different looser metrics can be established, in that case nearness between the members, to detect rough shape, location, size and fragments. Different nearness thresholds and minimal size criteria determine the individuation of two forests in Figure 3(a) and three forests in Figure 3(b). In addition, different precisifications of tree are used in Figure 3(a) and Figure 3(b), implying that in the former some elements are not considered members of the collective. The actual strategy for the demarcation (green shade connecting the trees) is not done until a later stage and depends on the selection of a suitable footprint strategy. In the figure, different thresholds for the demarcation are shown in different tones of green. Finally, it must be noted that it is not necessary to analyse forest in terms of a collective of trees in order to be able to individuate it. A similar approach to the one in mountains could be done for forests by interpolating a field of canopy cover in which similar measures to the ones used in Figure 2 can be used.

Although identity is likely to be characterised mainly through the rough location of the object, both identity and unity criteria are less intuitive for *forest* than for *mountain*, and are expected to vary between precisifications according to the particular management objectives.

[3] characterizes the height of a mountain or hill's summit by the vertical distance between it and the lowest contour line encircling it but containing no higher summit within it. It is a measure of the independence of a summit.

Thus, whether forest must be self-connected or not and whether it must comply with either morphological or metrical restrictions must be made explicit for all the precisifications that attempt to refer to *forests* as individuals.

5.4 Demarcation

Finally, by *demarcation* we mean the act of determining the spatial extension of an object, or equivalently of establishing its boundary. Once this extension/boundary is established, it may be referred to as ‘the demarcation’ of the object; or, in cases where the boundary is unclear or debatable, it may be regarded as one of many possible demarcations of the object.

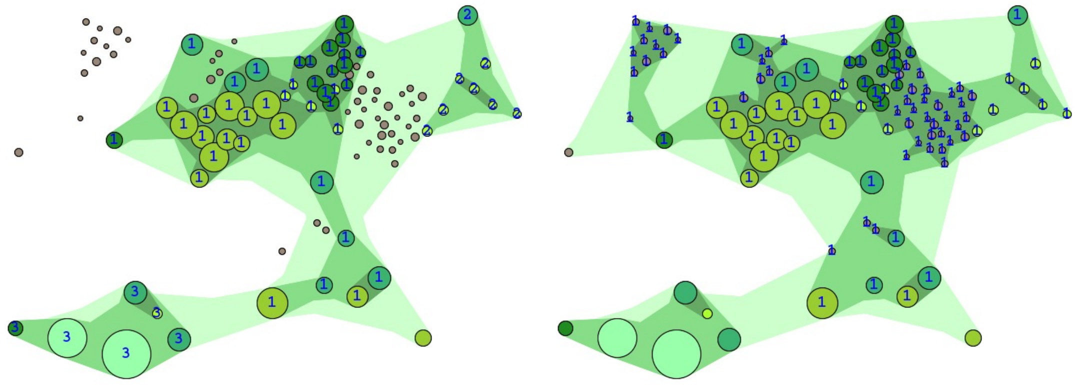
Establishing an object’s demarcation may be straightforward or extremely problematic. Moreover, although this may be because of the characteristics of the particular object under consideration, it is usually strongly related to the ontological category of object. For example, physical artifacts (e.g. cups, tables) are typically easily demarcated because they consist of solid matter forming an integral whole that is not physically connected to any other matter. Even with artifacts, the demarcation may not be completely clear. For example, some tables have a glass top that just rests on the wood below. It is debatable whether the glass is actually part of the table. For aggregate objects, such as a school of fish, or indeed a forest, demarcation is very often problematic, both because it may not be clear which entities should be counted as constituents and because there is no unique way to determine the spatial extension of something that is made up of many disconnected constituents (e.g., in demarcating a forest, one may want to include the space that lies between the trees within the demarcation of the forest, as can be seen in Figure 3). Distinguished regions within field like objects, which are again common in geography (e.g. soil type regions) also give rise to significant demarcation problems [30].

Although the study of suitable algorithms for the demarcation go beyond the scope of this paper, it must be noted that, in most of the studies, the specification of these strategies are often reduced to a set of thresholds [14, 24]. In some cases, particularly in human drawn maps, it has even been guided by intuitive and aesthetic judgements. We consider that a more careful analysis and specification of the algorithms used can be important for an appropriate characterisation of forests. For instance, a survey on the strategies that can be used to demarcate the forests identified in Figure 3 can be found in [11].

5.5 Interactions among classification, individuation and demarcation

We do not see classification, individuation and demarcation as cleanly separable aspects of a predicate’s semantics. As we have seen, a reasonable classification often presupposes an appropriate individuation. Individuation of an object typically goes some way towards determining its demarcation; and conversely by demarcating objects we usually also individuate them. Consequently, these different aspects may coexist within a definition and be entangled among different terms within an ontology. This may result in a complex albeit comprehensive ontology. But if one does not pay some attention to ensuring that all aspects are accounted for, they may be omitted. As we have seen, it is typically classification that receives the focus of attention, whereas individuation is taken for granted.

From a purely logical point of view, the difference between individuation and classification is also not entirely clear cut. This is because we often have choices as to what we take to be the universe of quantification. What is considered to be an individuating criterion in one choice of universe could be considered as a classification criterion applied to a larger universe of entities from which we want to select a significant subset.



(a) (b)

Figure 3 Forest individuation and demarcation.

6 Implementation of a Supervaluationistic Geographic Query System

We have implemented in Prolog a prototype system for individuating, demarcating and classifying forest regions. The system interprets a set of spatially located plant objects in terms of a precisification specification of the form $\langle \Theta(\mathsf{tree}),\ \Gamma(\nu),\ \sigma \rangle$, where $\Theta(\mathsf{tree})$ is a classification of tree objects in terms of more basic properties (e.g. species, height), $\Gamma(\nu)$ is a tree grouping algorithm parametrised by a nearness threshold ν, and σ gives the minimum number of trees for a tree group to be considered as a forest region.

The images in Figure 3 were generated by this system and illustrate how altering the threshold for 'nearness' used in an aggregation algorithm and the classification criteria of constituents affect individuation and demarcation of aggregates. The mid green region indicates the grouping obtained using the nearness threshold, ν, that has been applied in grouping and determines the number of tree groups counted. (We may regard the tree groups as 'forests', although to keep the computation simple and the images clear, we are individuating much smaller groups of trees than would normally be considered to be forests.) The light green area is computed with a nearness threshold 2ν, which incorporates all trees into one group, and the dark area with $\nu/3$ shrinks the tree group demarcations to include only the more dense areas. In visualisation (a) the small brown circles (thorn bushes) are not counted as forest constituents according to the chosen version of $\Theta(\mathsf{tree})$. Hence, we get a split between forest region 1 and forest region 2. In visualisation (b), using an alternative theory $\Theta'(\mathsf{tree})$, thorn bushes are treated as forest constituents, so regions 1 and 2 become merged. Also, the minimum number of trees required to count as a forest region, σ, has been increased in the (b) precisification. Because of this, the group counted as 3 in (a) is no longer considered to be a forest region in (b).

7 Conclusions

We have discussed several challenging problems that obstruct the task of giving precise definitions of geographic terms, such as 'forest'. To address these challenges we have provided a framework within which one can specify a range of possible interpretations of the 'forest', and which makes explicit how the semantic aspects of *classification*, *individuation* and *demarcation* interact and combine within possible definitions. We have indicated how this variability can be modelled within an ontological theory augmented with supervaluation semantics incorporating explicit specification of precisifications. The proposed framework has been implemented within a prototype Prolog-based GIS. In future work we intend to give a fully formalised theory, which will form an ontology module within our system or could equally be used within a different (non-Prolog) implementation based on the same general principles. Because of its generality and flexibility, this framework could be applied to characterising a wide range of geographic and other spatial objects, even where significant vagueness and ambiguity is present and where complex individuation and demarcation criteria may be required.

Since spatial properties and relationships often play an essential role in specifying individuation and demarcation criteria, we believe that theoretical study of these aspects of predicate semantics will play a key role in establishing more comprehensive and robust ontologies of geographic and other spatially related terminology. Moreover, the development of foundational spatial information theory, which unites both geometrical and cognitive aspects of space, will play a key role in addressing this challenge.

References

1 F. Achard. Forest definition and extent. In *Vital Forest Graphics*, pages 6–9. UNEP/Earthprint, 2009.

2 P. Agarwal. Ontological Considerations in GIScience. *International Journal of Geographical Information Science*, 19(5):501–536, 2005.

3 B. Bennett. What is a Forest? On the vagueness of certain geographic concepts. *Topoi*, 20(2):189–201, 2001.

4 B. Bennett. Physical objects, identity and vagueness. In *Principles of Knowledge Representation and Reasoning: Proceedings of KR2002*, pages 395–406, San Francisco, CA, 2002. Morgan Kaufmann.

5 B. Bennett. Modes of concept definition and varieties of vagueness. *Applied Ontology*, 1(1):17–26, 2005.

6 D. Bruce and L. C. Wensel. Modelling forest growth: approaches, definitions, and problems. *USDA Forest Service general technical report NC*, 1988.

7 D. C. Buckland, D. R. Redmond, and R. Pomerleau. Definitions of Terms in Forest and Shade Tree Diseases. *Canadian Journal of Botany*, 35(5):675–679, 9 1957.

8 R. Casati, B. Smith, and A. Varzi. Ontological Tools for Geographic Representation. *Formal Ontology in Information Systems*, pages 77–85, 1998.

9 R. Chazdon, P. Brancalion, and L. Laestadius. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration, 2016.

10 A. Comber, P. Fisher, and R. Wadsworth. What is land cover? *Environment and Planning B: Planning and Design*, 32(2):199–209, 2005.

11 M. Dupenois and A. Galton. Assigning Footprints to Dot Sets: An Analytical Survey. In *COSIT 2009*, pages 227–244. Springer, Berlin, Heidelberg, 2009.

12 M. Egenhofer. What's special about spatial? *ACM SIGMOD Rd*, 22(2):398–402, 6 1993.

13 M. Egenhofer and D. Mark. Naive Geography. *International COSIT*, pages 1–15, 1995.

14 FAO. *Global Forest Resources Assessment.* 2015.

15 P. Fisher. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113:7–18, 2000.

16 F. Fonseca and M. Egenhofer. Ontology-driven geographic information systems. *Proceedings of 7th ACM ISA in GIS*, 35(4):14–19, 1999.

17 D. Fuller. Tropical forest monitoring and remote sensing: A new era of transparency in forest governance? *Singapore Journal of Tropical Geography*, 27(1):15–29, 2006.

18 C. Giri, Z. Zhu, and B. Reed. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, 94(1):123–132, 2005.

19 A. Grainger. Difficulties in tracking the long-term global trend in tropical forest area. *Proceedings of the NAS of the USA*, 105(2):818–823, 2008.

20 A. Grainger. Uncertainty in the construction of global knowledge of tropical forests. *Progress in Physical Geography*, 34(6):811–844, 2010.

21 N. Guarino. Formal Ontology and Information Systems. *Proceedings of FOIS*, 98:6–8, 1998.

22 N. Guarino and C. Welty. Identity, Unity, and Individuality: Towards a Formal Toolkit for Ontological Analysis. *Ecai 2000*, pages 219–223, 2000.

23 N. Guarino and C. Welty. An Overview of OntoClean. In *Handbook on Ontologies*, pages 201–220. Springer Berlin Heidelberg, 2009.

24 M. C. Hansen et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 850(27):123–134, 2013.

25 P. Holmgren and R. Persson. Evolution and prospects of global forest assessments. *UNASYLVA-FAO*, 2003.

26 Z. Hu, G. Tang, and G. Lu. A new geographical language: A perspective of GIS. *Journal of Geographical Sciences*, 24(3):560–576, 2014.

27 M. Kavouras and M. Kokla. A method for the formalization and integration of geographical categorizations. *International Journal of GIS*, 16(5):439–453, 2002.

28 H. Gyde Lund. When Is a Forest not a forest? *Journal of Forestry*, 100(8):21–27, 2002.

29 H. Gyde Lund. Definitions of Forest, Deforestation, Afforestation and Reforestation. *Forest Information Services*, pages 1–159, 2007.

30 D. M. Mark and F. Csillag. The nature of boundaries on 'area-class' maps. *Cartographica*, 26:65–78, 1989.

31 E. Mena and V. Kashyap. Domain Specific Ontologies for Semantic Information Brokering on the Global Information Infrastructure. In *Proceedings of FOIS'98*, pages 269–83. 1998.

32 W. Meyer and B. Turner. *Changes in land use and land cover: a global perspective*. Cambridge University Press, 1994.

33 B. Smith and D. M. Mark. Ontology and geographic kinds. In *Proceedings of the 8th International Symposium on Spatial Data Handling (SDH'98)*, 1998.

34 D. Tagliapietra et al. A review of terms and definitions to categorise estuaries, lagoons and associated environments. *Marine and Freshwater Research*, 60(6):497–509, 2009.

35 E. Tomai and M. Kavouras. From Onto-GeoNoesis to Onto-Genesis: The Design of Geographic Ontologies. *GeoInformatica*, 83(281), 2004.

36 A. Varzi. Philosophical Issues in Geography – An Introduction. *Topoi*, 2:119–130, 2001.

37 A. Varzi. Vagueness in geography. *Philosophy & Geography*, 4(1):49–65, 2001.

38 A. Varzi. Vagueness, Logic, and Ontology. *Yearbooks for Philosophical Hermeneutics*, 1:135–154, 2001.

39 G. Wiederhold. Interoperation, Mediation, and Ontologies. In *FGCS 1994 Workshop on Heterogeneous Cooperative Knowledge Base*, volume 3, pages 33–48. ICOT, 1994.

40 Fei Xu. Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*, 11(9):400–406, 2007.

Sense of Direction: One or Two Dimensions?

Clare Davies[1], Lucy Athersuch[2], and Nikki Amos[3]

1 **Department of Psychology, University of Winchester, Winchester, UK**
`clare.davies@winchester.ac.uk`
2 **University of Winchester, Winchester, UK**
3 **University of Winchester, Winchester, UK**

Abstract

The Santa Barbara Sense of Direction scale (SBSOD) has been an invaluable research tool for over 15 years. Previous studies with non-US populations, despite supporting the scale's internal validity, suggested national differences in individual item responses and possibly the factor analytic structure, although translation differences were confounded with cultural and environmental factors. Using a pooled British sample (N=151) – avoiding linguistic translation, yet reflecting 'old world' environmental experience and strategies – this paper revisits the SBSOD's validity and structure. While largely supporting the scale's internal validity across cultures and spatial environments, findings from this population suggest at least a two-factor structure underlying the scores, with the first factor explaining less than half of its variance, supporting the oft-discussed division between survey- and route-oriented strategies. We conclude by proposing a more nuanced, efficiency-based theory of 'sense of direction'.

1998 ACM Subject Classification H.1.2 Human Information Processing

Keywords and phrases sense of direction, spatial ability, cognitive mapping

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.9

1 Introduction

Since its publication in 2002 [11], the Santa Barbara Sense of Direction Scale (SBSOD) has proved invaluable in a wide range of studies whose authors wanted to include a simple measure of human spatial cognition. Thanks to the generosity of the University of Santa Barbara research team who created the SBSOD, the scale has been freely available online for the past 15 years. Used by cartography and GI technology researchers (e.g., [21, 2]), psychologists (e.g., [14]) and neuroscientists ([9, 23]) alike, to try to link its basic concept of 'sense of direction' to other behaviours and variables, like many psychometric measures it has shown differing degrees of predictive power to its expected correlates in different studies. Despite its authors' careful separation of its key construct from others in spatial cognition ([11, 10, 3]), the scale is often taken on trust as a general unitary measure of what is vaguely imagined to be 'large scale spatial ability' (e.g., [8]).

Beyond the SBSOD, however, studies have often suggested that human spatial performance involves a range of different strategies, influenced by a number of predictors, and that tasks reflecting it do not always correlate particularly well with one another [24]. This suggests that sense of direction may not be quite such a unitary construct. In parallel, over the decades since the 1970s discoveries concerning hippocampal place cells, neuroscience has shown that spatial navigational abilities actually involve a battery of complementary cell types, locations and pathways within the brain ([26, 6]). This tallies with decades of behavioural evidence that people typically solve environmental-scale spatial problems by drawing on and integrating multiple cues in different ways, depending upon the task and individual differences [25].

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore;
Article No. 9; pp. 9:1–9:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Often, however, research in this field still seems to be influenced by Siegel and White [22], in making the implicit assumption that an accurate survey representation of environmental-scale[1] space is still the ultimate goal for spatial cognition, even if we no longer subscribe to a similarly 'staged' theory of spatial learning where this is the final level. Yet it appears likely that for many purposes, short-cut strategies such as encoding a simpler topological cognitive map may not necessarily be inferior to a full topographic and metric, cardinal direction-aligned, survey. (Such a topological map would be based on landmarks and/or route knowledge, rather than more accurately metric 2D spatial topography.) High SBSOD scores have been repeatedly argued to indicate the latter, rather than the former, type of representation [13].

For the sake of modelling and predicting human spatial performance more closely, which in turn must inform spatial information theory and provision, we need to be clear as to how many dimensions or factors might underpin individual differences in 'sense of direction'. Recent work in Germany ([19, 18]) has suggested that the types of self-rated abilities in the SBSOD actually split into egocentric versus allocentric representation or knowledge, and separate knowledge of cardinal directions. This contrasts with the unitary 'sense of direction' construct generally claimed for the SBSOD. The contradiction was noted in the German work, but without offering clear explanations for it.

One potential reason might be the observation of Montello and Xiao [17],that people dealing with different types of environment might show more affinity with alternative problem-solving strategies that work better in their situation. In Europe and other 'old world' cultures, regular grid-pattern environments are far less common; where existing, they are not necessarily as predictable or north-aligned as many US cities. People from such cultures might find survey representations and cardinal directions too difficult to apply, yet still develop what they imagine to be a good mental map. If certain aspects of 'sense of direction' were thus of greater or lesser relevance in different populations, then in statistical terms we might expect the SBSOD's value for Cronbach's alpha to vary, implying more or less internal reliability for the psychometric scale. We may also expect, as the above German research team indeed found, that the scale splits into more than one subscale or underlying factor; different individuals choose different spatial strategies in the absence of simple cues to the survey layout.

One problem with interpreting such cross-cultural studies, as also noted by Montello and Xiao [17], is that language is often confounded with culture and environment. It can be difficult, then, to determine whether some differential responding to subsets of questions may be due to subtle differences in understanding of the concepts involved, even where translation has been skilfully made. Indeed, the translation process itself may be difficult when English phrases such as 'sense of direction' may not have exact equivalents in other cultures, or may themselves mean something different already. Thus we have a confound between culture and language, when trying to understand the contradictions in the sense-of-direction literature.

The present study, therefore, is an initial attempt to examine the SBSOD's underlying construct(s) of 'sense of direction' in another English-speaking country and culture, where environments and hence (perhaps) optimal strategies tend to be 'old world' like much of Germany, but the language does not (or at least, not the vocabulary used within the SBSOD). The British context has previously been shown to create different expectations and strategies in urban spatial tasks, relatively to US research participants [5]. Therefore, it is reasonable to expect that the SBSOD's items may also be subject to those differences. Below we reanalyse

[1] Based on Montello's scale distinctions, [16].

previously collected SBSOD scores from a series of research projects, all of them using British participants and conducted within England. First, however, we should look a little more closely at what we might expect from applying the SBSOD's specific items in this context.

2 The SBSOD and its items

As shown in Table 1, the SBSOD questions cover a range of self-rated abilities. Items 1 and 11 are about giving (presumably route, though this is unstated) directions to others; item 8 is about receiving and using them. Items 10 and 14 cover memory for routes, while items 7, 9 and 13 focus on using and liking maps (and possibly GPS or other technology, in the case of 'planning' in item 13). The remaining items cover other specific aspects of spatial awareness: distance (#3), cardinal directions (#5), novel environments (#6), location awareness (#12), holding some kind of 'mental map' (#15) and finally, actual 'sense of direction' as interpreted by the respondent (#4).

A common observation in environmental-scale spatial cognition studies has been that some people often demonstrate either a 'survey' knowledge of the space (with some integrated metric knowledge of distances and directions, as if in a topographic map). Others – or the same people in a less familiar or more constrained space – seem to rely more on landmarks and route topology. What SBSOD score could be obtained by a person in the latter situation, who was generally competent in most spatial situations yet never created a completely metric survey map?

This hypothetical competent-yet-topologically-constrained respondent could easily give themselves a high score on items 1, 2, 4, 6, 8, 10, 11, 12 and 14 of the SBSOD – 9 of the 15 items – because they may well have little trouble with routes or route directions in everyday life. Furthermore, at least in theory, that person could potentially interpret item 3 as pertaining to route distances, and item 15 as implying a good memory for key routes and some ability to link them up (though not necessarily into a metrically accurate 'map'), so the respondent might also score quite generously on those as well. Having thus shown confidence in 11 out of 15 items, the respondent might only score themselves more poorly on items 7, 9, 13 and 5 (although it is also conceivable to mentally align known route segments to an imagined north, even if the 'north' they imagined was not geographically true).

The above thought experiment is intended to show that a fairly high score on the SBSOD could reflect self-confidence in a landmark- or route-based spatial strategy, because it happened to work well enough in the respondent's everyday experience. It is worth remembering, of course, that in most major studies of the SBSOD to date the scale has nevertheless tended to show a significant[2] correlation with tasks which clearly required some metrically accurate cognitive mapping for successful completion. Thus in general, the most confident SBSOD respondents do seem to be those most capable of creating such mappings. Nevertheless, the varying and sometimes surprisingly low correlations with actual tasks (typically between 0.1 and 0.45) do leave room for other interpretations and strategies to be playing some part in people's SBSOD scores.

If, as the German data suggested, the SBSOD can reflect different common strategies for learning and navigating environments, then we may expect similar results with the British sample. Route-related and survey-related questions might thus show different response patterns. If the key difference between the US and German respondents is in fact local

[2] It should be noted, though, that these correlations are still usually no more than moderate, as is typical in psychology.

Table 1 SBSOD items and descriptive statistics from current sample, with (n=151) and without (n=132) the surveyor group.

[0.5ex] No.	Item (simplified)	Mean(sd) With	Mean(sd) Without
1	Good at giving directions	4.17(1.72)	3.93(1.68)
2	Good memory for where I left things	4.24(1.70)	4.27(1.70)
3	Very good at judging distances	3.97(1.70)	3.70(1.59)
4	My 'sense of direction' is very good	4.26(1.81)	4.02(1.75)
5	Think in terms of cardinal directions	2.57(1.77)	2.24(1.52)
6	Don't easily get lost in a new city	3.83(1.76)	3.67(1.75)
7	Enjoy reading maps	3.54(2.08)	3.08(1.79)
8	No trouble understanding directions	4.28(1.73)	4.07(1.65)
9	I am very good at reading maps	3.83(2.00)	3.43(1.81)
10	Remember routes very well as car passenger	4.05(1.99)	3.95(1.99)
11	I enjoy giving directions	3.61(1.68)	3.42(1.62)
12	Important to me to know where I am	5.30(1.56)	5.15(1.57)
13	Do navigational long-trip planning	4.13(2.03)	3.80(1.96)
14	Usually remember new route first time	4.30(1.84)	4.12(1.81)
15	Very good 'mental map' of my environment	4.48(1.81)	4.23(1.77)
Total	Mean score across items	4.04(1.16)	3.81(1.03)

environment and culture, rather than translation issues, we might expect a UK sample to resemble the Freiburg and Saarbruecken samples more closely than the US data.

Therefore, the present analysis had three aims: (1) to complement the Montello and Xiao cross-cultural analysis [17] with a British sample, to help disambiguate linguistic from environmental or cultural differences between the previous samples; (2) to see whether the British data supported the unitary (one underlying factor) nature of the SBSOD claimed by previous studies by the creators at UCSB, or whether it indicates the additional factors claimed by the German studies; and (3) if not, to try to tease out whether this could indicate different strategic approaches – but equal spatial confidence – by separate subgroups within our participant sample.

3 Participants, Age and Sex

The pooled sample used here was collated from five datasets collected in three UK locations over the past few years. Two, comprising two-thirds of the total sample (97 participants), were from student research projects at the University of Winchester, in the south of England, and used undergraduate psychology students as participants. Any students who took part in both studies were omitted, to avoid overlap in the data. Two more (totalling 35 participants) were collected previously at the University of Huddersfield, in the north of England, consisting of one staff and one student psychologist sample. The fifth sample consisted of 19 professional surveyors employed by Ordnance Survey and located in various locations around Great Britain, but mostly in the southern half of England. Thus the total N was 151.

Due to the predominance of psychology students in the sample, as in previous studies, the sample was largely female (124), with 27 males (including all of the OS surveyors). While the mean age was 29 (sd 13.3), this reflected an inevitable skew towards college-age participants: 61 per cent (92 participants) were aged 18-22. (Even so, this means that almost 40 per cent were above the typical age, making this sample perhaps more representative of the population than in many studies.)

Across the whole sample, age was found to strongly correlate with total SBSOD score: Spearman's rho (151) = .38, $p<.001$. When this analysis was repeated without the group of surveyors, however, the correlation became small and non-significant: Spearman's rho (132) = .14, p=.11. The group of surveyors may also have been largely responsible for a found sex difference in the responses: $t(149) = 7.84$ (with equal variances), $p < .001$. There were only 8 males in the non-surveyor sample, so the professional status of the all-male surveyors was inevitably confounded with gender in this case.

4 Analyses and Results

4.1 Analysis 1: Group Differences

The data was first assessed for homogeneity, by performing an independent analysis of variance on the SBSOD total scores. Study (group) was the independent variable. While Levene's test showed that homogeneity of variance could be assumed, the omnibus F test was highly significant: $F(4,146) = 16.34$, $p < .001$, partial eta squared = .31. Tukey post hoc comparisons showed that this was entirely due to the surveyors having a far higher mean SBSOD score (5.65, 95 per cent CI[5.21,6.10]) than all other groups, the highest of which was the older group of Huddersfield participants (psychology staff: mean SBSOD = 4.24, 95 per cent CI[3.77,4.71]). There were no significant differences between the university-based groups. For this reason, most of the analyses below were run at least once without the surveyors as well as with, to check that the results were not skewed by this unsurprisingly (but of course, unusually) expert group.

4.2 Analysis 2: Descriptives

European cultures such as Britain tend to place far less emphasis on cardinal directions and regular grid-pattern urban layouts than 'new world' countries such as the US. Table 1 was visually compared with Montello and Xiao's cross-cultural analysis [17], to see which pattern of responses within it was most similar to ours.

Overall, even with the surveyors included, the mean score obtained from our sample appears lower than both the Santa Barbara (USA) and Freiburg and Saarbrücken (Germany) samples analysed by Montello and Xiao – but higher than their Tokyo sample. In other respects, however, the pattern of responses between items seems quite similar to both the US and German samples, showing particular dips in score for both items 3 (distance judgement) and 5 (cardinal directions), and notably higher-than-average scores for items 12 (knowing where I am) and 15 (good mental map). Thus the pattern of descriptives in itself cannot distinguish which of the previous samples is most resembled by the current data. This is hardly surprising, as response patterns were also not very clearly differentiated between the samples in the original Montello and Xiao analysis.

The descriptives in Table 1 do suggest, however, that including the surveyor group inflated most of the mean scores (and hence also their standard deviations), despite making up less than 20 per cent of the total sample. An exception here is item 2, on remembering where one had left things; this probably reflects the older age profile of the surveyors, many of whom were nearing retirement age. To the extent that we would expect far higher self-rated abilities in professionals for whom spatial awareness and navigation are essential elements of their job, the increased scores help to validate the data, and show that participants were responding as expected.

4.3 Analysis 3: Scale Reliability

Using the university samples only (N=132), the SBSOD was assessed for inter-item reliability using Cronbach's alpha. As in previous studies ([11, 17]), the scale showed good reliability: alpha = .868.

It is obviously interesting to see whether any items in the SBSOD make a non-productive contribution to the scale – i.e., if they reduce rather than increase its inter-item reliability. For some reason, to our knowledge this statistic has not previously been reported for the individual SBSOD items. With the present sample, one might have expected item 5 (about reliance on cardinal directions) to show a less strong relationship to the rest of the scale: in most British towns it is impossible to rely on such absolute spatial cues, so much so that even when they are present locals appear to stick to other spatial strategies ([5]). Montello and Xiao [17] explicitly raised the likelihood that such environmental differences might affect cultural tendencies towards different spatial strategies.

Surprisingly, though, the item-total statistics showed that Cronbach's alpha would be slightly lower (.866) if item 5 had been omitted, suggesting that it was still contributing to the overall coherence of the scale. Item 12 ("It's important to me to know where I am") contributed similarly marginally (again at .866). Perhaps inevitably (and as noted by Montello and Xiao without quoting this statistic), item 2 on locating objects was the only item shown to reduce Cronbach's alpha, which would have been .872 without it. Similarly, item 10 (on memory for routes when a car passenger) made no apparent difference to alpha: it would still have been .868 without that item. All other items would have reduced alpha if omitted, although in most cases only marginally; the largest potential loss of reliability was from item 4 (the actual item on 'sense of direction'), reducing alpha to a still-respectable .847. (This is not surprising, since in the original development of the scale, apparently its authors deliberately chose items for their correlation with that one [11].) Thus we can confirm previous findings that although item 2 would be best dropped in any revision of the scale, in general the rest shows robust inter-item reliability.

Such results might be problematic, if they showed generalised responding by participants – a tendency to give a similarly high or low score to all questions, perhaps due to inattention to the wording. Fortunately, however, other reliability statistics (ANOVA with Tukey's test for nonadditivity, and Hotelling's T-squared test) showed that questions did tend to be quite strongly distinguished from each other by participants – while still showing some response consistency as above.

Does good internal reliability in itself imply that the scale must measure a single underlying cognitive ability across all participants? Not necessarily. In our view this would falsely imply, *post hoc ergo propter hoc*, that we could assume a single common cause for any set of items which happened to intercorrelate. Carroll [4] demonstrated how in an imaginary multi-item measure, despite strong overall consistency across the items, nevertheless different participants could be adopting different strategies or displaying multiple relevant strengths to differing degrees. Thus a psychometric scale may often have high internal reliability overall, yet also show different *patterns* of responding (reflecting different styles or strategies) by different individual respondents. Therefore, factor analysis is also helpful as below, if our goal is to test whether there may be multiple ways to achieve good scores on a given psychometric scale – as opposed to ensuring that its questions do all contribute to the overall concept.

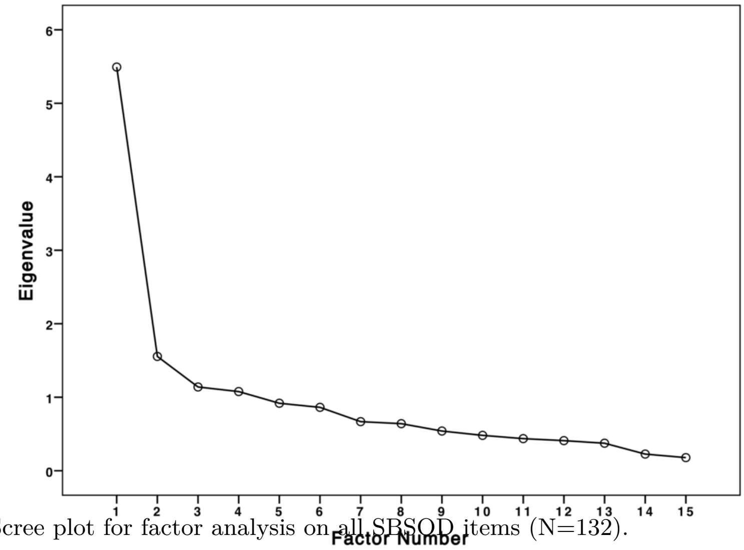


Figure 1 Scree plot for factor analysis on all SBSOD items (N=132).

4.4 Analysis 4: Factor Analysis

The data from the 132 English university participants was submitted to a factor analysis. All common assumptions for factor analysis appeared to be met by the dataset.

Principal axis factoring (PAF) was used, as this attempts to produce 'clean' factors which optimise the grouping of loadings. Whilst some previous papers (e.g., [17]) did not always specify the extraction method used in previous factor analyses of the SBSOD, they usually appear to have used a similar factorising method rather than principal components analysis. This makes logical sense: while we can expect a lot of individual variance within items in the SBSOD, the only goal of our analysis here is to assess the common links between them, not to try to explain every item-specific issue within people's responses.

The scree plot from this analysis is shown in Figure 1. While showing a strong primary factor, as is typical of factor analysis on any psychometric scale, unlike previous studies the scree plot is ambiguous about the potential role of further factors. Four factors had eigenvalues above 1.0 (the so-called Kaiser criterion), and the first one only explained 36.6 per cent of the variance in the data, with the second factor explaining 10.4 per cent, the third 7.6 and the fourth 7.2. Thus the evidence for a unitary psychological construct underpinning the scale seems to be weakened in this population, as with the previous German studies.

To enable interpretation of the factors, the analysis was repeated using orthogonal (Varimax) rotation, limiting the number of factors to the two which had explained more than 10 per cent of the data. The pattern matrix (see Table 2) suggested that while eleven items of the fifteen loaded at above .3 on the first factor, items 10, 14 and 15 loaded more strongly on the second, and loadings above .3 on that second factor were also seen in items 4, 6, 8 and 13.

Translating this into plain language, the highest loadings on the first factor were the items which one might think of as most 'surveyish' in the scale: items 7, 9 and 13 which concerned using maps and related technologies, items 1 and 11 on being able and confident in giving directions to others (which may often require more complete spatial knowledge than individual route topology), and items 3 and 5 on distance estimation and use of cardinal directions. Item 8, on understanding other people's directions, also loaded on this factor. But the latter actually loaded quite strongly onto both factors – along with general 'sense of direction' (item 4), and both the self-orientation items 6 and 12. The second factor's

Table 2 Pattern matrix for the two main factors in both the orthogonal (varimax) and oblique (direct oblimin) rotated factors, showing only loadings at 0.3 and above.

No.	SBSOD item	Orth F1	Orth F2	Obl F1	Obl F2
1	Good at giving directions	.629		.656	
2	Good memory for where I left things				
3	Very good at judging distances	.497		.515	
4	My 'sense of direction' is very good	.588	.569	.567	.389
5	Think in terms of cardinal directions	.446		.483	
6	Don't easily get lost in a new city	.431	.361	.424	
7	Enjoy reading maps	.759		.835	
8	No trouble understanding directions	.507	.338	.511	
9	I am very good at reading maps	.864		.938	
10	Remember routes very well as car passenger		.717		.759
11	I enjoy giving directions	.477		.496	
12	Important to me to know where I am	.320		.318	
13	Do navigational long-trip planning	.585	.353	.595	
14	Usually remember new route first time		.667		.630
15	Very good 'mental map' of my environment		.561		.518

strongest loadings, however, were from items 10, 14 and 15. The first two of these are about memory for routes, and the third is on possession of a 'good mental map'.

Could it be that for some participants, 'good sense of direction', self-orientation and a 'good mental map' are linked to a route-based rather than a survey-based strategy? (After all, this does not say whether the mental 'map' in question is more like a street plan or a subway map in its content.) The results seem to imply a clear dissociation between at least some of the questions about routes, and those about maps and cardinal directions, so the possibility seems plausible.

This pattern of loadings, with its two-factor overlap and dissociation, appeared even more strongly when the factor analysis was repeated twice more with just the two factors, and using first orthogonal but then oblique (Oblimin) rotation – the latter to allow the two factors to correlate (as, typically, psychometric aptitude factors tend to do). In the latter case, as shown on the right side of Table 2, item 4 (the actual 'sense of direction' item) was the only item to load on both factors; the other loadings on factor 2 were items 10, 14 and 15. Once again, this implies that some participants may consider themselves to have a good sense of direction and mental map, yet rely on route memory rather than constructing a survey representation. Correlation between the factors was .448.

Figure 2 shows the factor plot in rotated factor space (though showing the factors orthogonally, and thus slightly distorting the actual shape). The 'route' and the 'map'-related questions seem to lie along two separate dimensions, but with 'sense of direction' loading similarly on both, and with 'good mental map' closer to the route-memory items than to the items about reading cartographic maps. For these participants, perhaps, a 'good mental map' is apparently a collection of well-memorised routes and landmarks, possibly linked into a framework which is more topologically than topographically accurate – but apparently so efficient for everyday use that participants consider it to be a 'good mental map' and showing 'good sense of direction'.

When the analysis was rerun including the group of surveyors, it showed a very similar pattern of communalities and loadings, but with slightly higher correlation and less distinct-

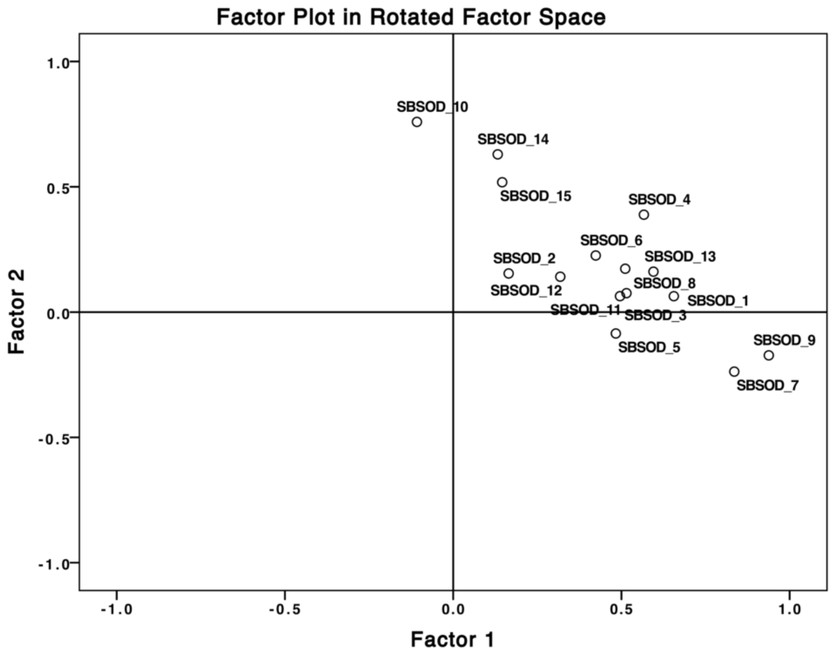


Figure 2 Factor plot for the two factors from the final factor analysis (N=132).

iveness between the factors – probably because the surveyors tended to have high scores on all items anyway.

The inclusion of both orthogonal and oblique rotation results above was usefully questioned by one reviewer of this paper. In brief, it seems more intuitively plausible (as with any psychometric scale) that some of the unique variance of any item will be due to conceptually irrelevant 'error' variance, which is why factor analysis is generally favoured. However, at present there seems to be insufficient evidence to make a theoretical decision as to whether or not the underlying factors in sense of direction would be expected to correlate; e.g., some people may have strong knowledge of both the topology and topography of familiar spaces (given the above-cited neuroscience evidence for multiple spatial representations in the brain). Yet alternatively, these might be viewed as competing strategies, of which only one might be used in a given scenario. Without a theoretical reason to select either option, both have been explored here. Obviously, further confirmatory and theory-based studies are needed. In the meantime, we turn to a comparison of participants who tended to score more highly on one, rather than the other, subset of questions as identified above.

4.5 Analysis 5: Comparing route- and map-oriented participants

To check the above arguments, participants with a 'route' pattern of responses (scoring higher on items 10 and 14) were identified, and compared to those showing more affinity with cartographic survey maps (items 9 and 7), the two highest-loading items on the primary factor above. Participants were grouped according to whether their scores were higher on the former two questions (summed together) or on the latter two. 76 participants scored higher on the two 'route' questions, 55 on the two 'map' questions, and 20 had scored equally on the two question pairs. Excluding the surveyors, these figures were 76, 39 and 17 respectively.

Excluding the group of surveyors, whose obvious map bias and generally high scores across the board would be likely to skew the results, the 'route' and 'map' groups' scores on individual SBSOD items and on its total score were compared using t-tests. As this made for 16 tests, familywise error was corrected for using the Bonferroni heuristic – i.e., the tests were only counted as significant if $p < .05/16$, i.e. $<.003125$. (However, as this is rather a conservative correction, p values close to this value would also be reported.) Equal variances could be assumed in all tests reported below, except where stated.

Results showed that the 'map'-oriented group scored significantly higher on item 1 (good at giving directions: $t(113) = 3.29$, $p = .001$, mean difference = 1.02 with 95pc CI[0.41,1.64]). They also scored higher, of course, on the two items used as the basis for the grouping: item 7 (enjoy reading maps: with unequal variances, $t(56.2) = 7.17$, $p < .001$, md = 2.26, 95pc CI[1.63,2.89]) and item 9 (good at reading maps: $t(113) = 7.66$, $p < .001$, md = 2.17, 95pc CI[1.61,2.74]). However, they differed only very slightly and non-significantly on their overall SBSOD score: $t(113) = 2.16$, $p = .033$, md = .39, 95pc CI[.03,.75].

An additional t-test to compare the two groups' age profiles suggested that, even without the surveyor group, the more map-oriented group was generally slightly older: with significantly unequal variances due to the skew in the age distribution, $t(50.4) = 3.16$, $p = .003$, md = 7.6 years, 95pc CI[2.8,12.5].

5 Discussion

The above series of analyses suggest that at least for British participants, as with Münzer et al.'s German participants, there is more than one way to score highly on the SBSOD, indicating confidence in your large-scale spatial ability. The participants whose sense of good spatial ability rested more on their memory for routes were nevertheless scoring equally well on the SBSOD to those with more affinity to topographic 'survey' representations. Although the third factor (concerning cardinal directions) was not strongly supported, it was impossible to support a single-factor interpretation of the scores in this sample, while still explaining at least around half of the variance in the data. Thus we suggest that a two-factor model may better represent the range of spatial strategies, for at least this British population.

Consequently the SBSOD cannot, for all respondents, be assumed to indicate their degree of survey-like cognitive mapping. Furthermore, where it correlates with a given spatial task, this should not be taken as support that a survey representation is the key to good performance on that task. Arguably, the SBSOD is conflating two sometimes equally valued spatial strategies, broadly characterised as 'route' (compiling and linking egocentric information) or 'survey' (deriving integrated allocentric spatial knowledge) in most of the literature. Indeed, these two have already been specifically teased apart in other psychometric scale developments [20], possibly even with the 'survey' ability being split further into 'allocentric-survey' versus 'egocentric-survey' [27].

Why, then, have previous studies shown the SBSOD to be more closely related to performance on tasks which clearly *do* require an integrated, and at least approximately metrically accurate, 'survey' representation? The answer to this may partly lie in the population sampled for those studies. It is reasonable to assume that participants in the original US West Coast student population are more familiar with environments where a survey representation is relatively easy to acquire, and reliable for drawing inferences (such as alternative routes through a known street grid). As noted by Montello and Xiao [17], where this is very much not the case – as in European and other 'old world' cities, and where people are more used to any highly irregular environments – participants may often obtain good spatial performance by relying on a more topologically-based heuristic.

In addition, the evidence from the present sample suggests that in general, those who have the strongest reliance on a survey representation will, like the group of Ordnance

Survey surveyors examined here, obtain high scores across most aspects of the scale. Thus in a mixed sample which includes (say) geographers or other map-proficient subgroups, the most extremely map-oriented participants will probably tend to perform the best on survey-demanding tasks, as well as scoring the highest on the SBSOD overall. However, across a less extreme and more typical population, such as we like to think is represented by psychology students, the present data suggests that the participants with the highest confidence in their spatial skills will *not* necessarily be more inclined towards survey-based strategies.

In other words, where the present study goes further than previous analyses is in showing that any lack of a fully metric, survey-based representation may not necessarily reduce participants' confidence in their 'sense of direction', because for them a route-based strategy has been performing well, and may actually work better. This may also help to explain why an advantage for highly 'survey'-oriented participants was not found in at all in a German indoor study by Hölscher et al. [12], and why Meilinger's extensive studies on spatial strategies [15] similarly posited a 'network of reference frames' (memorised scenes from individual vista-scale spaces) as apparently the most common cognitive mapping strategy.

Neuroscience evidence also supports this, suggesting that in environmental (as opposed to vista) spaces the processing of landmarks from visual and other sensory data in the parahippocampal place area is particularly focused on decision points, and that the outputs from such processing are then linked in the retrosplenial cortex to place knowledge and head-direction information, to indicate which way to turn ([7, 1]).

Ishikawa and Montello [13] characterised an imperfect, largely qualitative and only metrically approximate, mental representation as "undoubtedly desirable in the face of the limited cognitive capacity of humans" [p. 124]. Indeed, this would seem to be a key point in understanding people's spatial cognition of large, complex spaces. Network topology is undoubtedly more computationally efficient for many tasks, even including relative distance and direction estimates, than a 'survey' perspective drawing on topographically accurate maps.

Modern environments, unlike the open savannah where our ancestors apparently evolved, do not allow either simultaneous viewing nor free roaming over the entire area. Unless the shape and pattern of the space is quite predictable, as in grid cities, distortions in our understanding (when based solely on experience) are inevitable, but not necessarily problematic. Many readers from 'old-world', less regular, environments may well have had the disorienting experience, like the present author, of eventually viewing a cartographic map of an environment which they learned solely through repeated route experience (e.g., the town they grew up in). They may find it very hard to relate their own undoubted local expertise to the projected 2D topography in front of them. Yet their stored metric inaccuracies and simplifications may have caused them no problems over extended periods of time, and have proved repeatedly efficient not only at route-finding but also at giving and receiving directions, and otherwise sharing place knowledge with fellow locals. Perhaps this helps to explain why particularly the younger participants in the present (relatively intelligent and educated) sample were apparently quite reluctant map users, even when claiming a strong 'sense of direction' and a good 'mental map' for themselves.

Overall, then, many survey-demanding tasks are rare and irrelevant to everyday life for many people, so a true survey representation of complex, irregular spaces would be a waste of cognitive resource. The spatial information community may therefore do better to focus on simulating and supporting simplified, more efficient cognitive mapping, both by humans and by robots or simulated agents, rather than attempting to encourage or impose metrically accurate (but in many situations, cognitively inefficient) mental representations. There is a good reason why, in the chaotic geography of cities like London, the classic Tube map's simplicity is greeted with relief.

Acknowledgements. We are grateful to David Peebles for originally collecting the Huddersfield samples used in this paper. Its use and interpretation within the present analysis, however, is entirely the responsibility of the first author. Some of the studies included in the paper were originally funded by Ordnance Survey of Great Britain (now Ordnance Survey Ltd.).

References

1 Andrej Bicanski and Neil Burgess. Environmental anchoring of head direction in a computational model of retrosplenial cortex. *Journal of Neuroscience*, 36(46):11601–11618, 2016.
2 Anke Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. Interactivity improves usability of geographic maps for visually impaired people. *Human-Computer Interaction*, 30(2):156–194, 2015.
3 Heather Burte and Mary Hegarty. Revisiting the relationship between allocentric-heading recall and self-reported sense of direction. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pages 162–167. Cognitive Science Society, 2012.
4 John B. Carroll. *Human Cognitive Abilities: A survey of factor-analytic studies*. Cambridge University Press, 1993.
5 Clare Davies and Eric Pederson. Grid patterns and cultural expectations in urban wayfinding. In Daniel R. Montello, editor, *Spatial Information Theory: COSIT 2001*, pages 400–414. Springer, 2001.
6 Julie R. Dumont and Jeffrey S. Tauber. The neural correlates of navigation beyond the hippocampus. In Shane O'Mara and Marian Tsanov, editors, *Progress in Brain Research: The Connected Hippocampus*, volume 219, pages 83–102. Elsevier, 2015.
7 Russell A. Epstein and Lindsay K. Vass. Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B*, 369, 2014.
8 Amanda M. Ferguson, Erin A. Maloney, Jonathan Fugelsang, and Evan F. Risko. On the relation between math and spatial ability: The case of math anxiety. *Learning and Individual Differences*, 39:1–12, 2015.
9 Mark A. Halko, Erin C. Connors, Jaime Sanchez, and Lotfi B. Merabet. Real world navigation independence in the early blind correlates with differential brain activity associated with virtual navigation. *Human Brain Mapping*, 35(6):2768–2778, 2014.
10 Mary Hegarty, Daniel R. Montello, Anthony E. Richardson, Toru Ishikawa, and Kristin Lovelace. Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34:151–176, 2006.
11 Mary Hegarty, Anthony E. Richardson, Daniel R. Montello, Kristin Lovelace, and Ilavanil Subbiah. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30:425–447, 2002. URL: `http://www.earthgate.ucsb.edu/~montello/pubs/sbsod.pdf`.
12 Christoph Hölscher, Tobias Meilinger, Georg Vrachliotis, Martin Brösamle, and Markus Knauff. Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology*, 26:284–299, 2006.
13 Toru Ishikawa and Daniel R. Montello. Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, 52:93–129, 2006.
14 Kashka Iwanowska and Daniel Voyer. Dimensional transformation in tests of spatial and environmental cognition. *Memory & Cognition*, 41(8):1122–1131, 2013.

15 Tobias Meilinger. *Strategies of Orientation in Environmental Spaces*, volume 22 of *MPI Series in Biological Cybernetics*. Max-Planck-Gesellschaft, 2008.

16 Daniel R. Montello. Scale and multiple psychologies of space. In Andrew U. Frank and I. Campari, editors, *Spatial Information Theory: a theoretical basis for GIS*, pages 312–321. Springer-Verlag, 1993.

17 Daniel R. Montello and Danqing Xiao. Linguistic and cultural universality of the concept of sense-of-direction. In Max Egenhofer, Nicholas Giudice, Reinhard Moratz, and Michael Worboys, editors, *Spatial Information Theory: COSIT 2011*, pages 264–282. Springer-Verlag, 2011.

18 Stefan Münzer, Benedict C. O. F. Fehringer, and Tim Kühl. Validation of a 3-factor structure of spatial strategies and relations to possession and usage of navigational aids. *Journal of Environmental Psychology*, 47:66–78, 2016.

19 Stefan Münzer and Christoph Hölscher. Entwicklung und Validierung eines Fragebogens zu räumlichen Strategien (Development and validation of a self-report measure of environmental spatial strategies). *Diagnostica*, 57:111–125, 2011.

20 Francesca Pazzaglia and Rossana De Beni. Strategies of processing spatial information in survey and landmark-centred individuals. *European Journal of Cognitive Psychology*, 13(4):493–508, 2001.

21 Thomas Pingel and Keith Clarke. Perceptually shaded slope maps for the visualization of digital surface models. *Cartographica*, 49(4):225–240, 2014.

22 Alexander W. Siegel and Sheldon H. White. The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, 10:9–55, 1975.

23 Joost Wegman, Hubert M. Fonteijn, Janneke van Ekert, Anna Tyborowska, Clemens Jansen, and Gabriele Janzen. Gray and white matter correlates of navigational ability in humans. *Human Brain Mapping*, 35:2561–2572, 2014.

24 Steven M. Weisberg, Victor R. Schinazi, Nora S. Newcombe, Thomas F. Shipley, and Russell A. Epstein. Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3):669–682, 2014.

25 Thomas Wolbers and Mary Hegarty. What determines our navigational abilities? *Trends in Cognitive Sciences*, 14(3):138–146, 2010.

26 Thomas Wolbers and Jan M. Wiener. Challenges for identifying the neural mechanisms that support spatial navigation: the impact of spatial scale. *Frontiers in Human Neuroscience*, 8:571, 2014.

27 Jimmy Y. Zhong and Maria Kozhevnikov. Relating allocentric and egocentric survey-based representations to the self-reported use of a navigation strategy of egocentric spatial updating. *Journal of Environmental Psychology*, 46:154–175, 2016.

How Subdimensions of Salience Influence Each Other. Comparing Models Based on Empirical Data*

Markus Kattenbeck

University of Regensburg, Information Science Group, Regensburg, Germany
Markus.Kattenbeck@ur.de

Abstract

Theories about salience of landmarks in GIScience have been evolving for about 15 years. This paper empirically analyses hypotheses about the way different subdimensions (visual, structural, and cognitive aspects, as well as prototypicality and visibility in advance) of salience have an impact on each other. The analysis is based on empirical data acquired by means of an in-situ survey (360 objects, 112 participants). It consists of two parts: First, a theory-based structural model is assessed using variance-based Structural Equation Modeling. The results achieved are, second, corroborated by a data-driven approach, i.e. a tree-augmented naïve Bayesian network is learned. This network is used as a structural model input for further analyses. The results clearly indicate that the subdimensions of salience influence each other.

1998 ACM Subject Classification G.3 Multivariate Statistics

Keywords and phrases Salience models, consistent PLS-SEM Analysis, Bayesian Networks

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.10

1 Introduction

Human navigation is an intrinsically complex task, involving a diverse range of spatial cues, computational mechanisms and spatial representations (cf. [41]). Despite its complexity, humans are able to successfully find their way on a day-to-day basis. The importance of landmarks for human navigation is undoubted across disciplines. Prototypical systems using landmarks have revealed their usefulness in supporting human wayfinding of pedestrians and drivers, alike (cf. [32, p. 83]). Theories about the landmarkness of objects, i.e. about the salience of (geographical) objects have been developed for the last 15 years (cf. Section 2). However, lack of empirically validated models of *salience* was identified to be a major weakness in current research on estimation of salience (cf. [32]). Accordingly, the goal of this paper is to add to state-of-the-art theories by proposing hypotheses about the way subdimensions of salience, i.e. *visual salience*, *cognitive salience*, *structural salience*, *visibility in advance*, and *prototypicality*, are intertwined. It focuses, thereby, on pedestrian navigation scenarios. Using a dataset based on an in-situ study (cf. [24]) the analysis of the predictive capabilities of the model proposed here, in turn, comprises two steps. First, the degree of influence different subdimensions of salience show on each other is assessed using consistent Partial Least Squares Structural Equation Modeling (PLSc). Afterwards, the results of this theoretical model are compared to those based on a prior Bayesian Network analysis (cf. Section 5.2.2) in order to further backup theoretical claims empirically.

* Parts of this paper were taken from an unpublished doctoral thesis (cf. [25]).

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 10; pp. 10:1–10:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2 Related Work – Theories about Salience

While the earliest empirical attempt to gain an insight into the factors which contribute to a building's *salience* date back to [1], *salience* as a concept had been formalized around the turn of the century. Five papers, published between 1999 and 2005 build the nucleus of the work done. In [36] Sorrows and Hirtle distinguish three dimensions contributing to salience: visual, structural and cognitive aspects (encompassing, among others, prototypicality, thereby drawing heavily on [34]). However, they do not develop a formal model to capture these. Raubal and colleagues (cf. [31]) introduce a formal model providing measures for each of the three constructs. However, Raubal et al. refer solely to the façades of buildings. Nothegger et al. (cf. [29]) show that the model introduced in [31] is useful for distinguishing between different buildings. Winter (cf. [39]) adds the notion of visibility in advance as contributing to a landmark's *salience*, i.e. he clearly stresses the importance of the particular route. Finally, Klippel and Winter (cf. [26]) give a very detailed account of structural salience, and, in doing so, change the meaning proposed in [31]: 'Objects are called structurally salient if their location is cognitively or linguistically easy to conceptualize in route directions' [26, p. 347].

This initial work was refined by two publications (cf. [5, 6]). The key idea of this refinement is the fact that no object is salient eo ipso. [6] stresses the importance of context by focusing on the interaction between observer, observed, and surroundings. Based on this understanding Caduff proposes a Bayesian network for computing salience values which is largely based on visual attention research (cf. [5]). He distinguishes between

perceptual salience which reflects exogenous allocation of attention
cognitive salience which mirrors endogenous allocation of attention
contextual salience which acknowledges the current navigational context

Based on these definitions Caduff introduces several auxiliary components, e.g. degree of recognition, idiosyncratic relevance, scene context, and combines these to a Bayesian Network. It is noteworthy, though, that – in opposition to the current study – no relationships among perceptual, cognitive or contextual salience as high-level components were hypothesized.

Based on these studies, the following operational definition of *salience* can be derived.

▶ **Definition 1** (Salience)**.** Given a local environment an observer is in, *salience* is the degree to which an object, persistent enough to be used in route instructions, draws the average pedestrian observer's attention. This degree is evoked by

1. visual features the objects has (*visual salience*),
2. the degree of prototypicality it shows (*prototypicality*),
3. how identifiable it is when approached (*visibility in advance*),
4. the ease with which it may be integrated into a route description (*structural salience*) and
5. the degree as to which it can evoke prior knowledge about the object (*cognitive salience*).

According to [24] several items for each of these dimensions were included in the survey they used for data acquisition. Therefore, instead of repeating the full list of questions, which can be found in [24], Table 1 is used to give an impression of the questions asked.

3 A theory-driven Structural Model

Based on these theoretical explanations it is important to note that none of the studies mentioned hypotheses causal relationships between the different subdimensions of *salience*. Contrastingly, the theoretical model proposed (cf. Figure 1) is based on several hypotheses about the way the subdimensions influence each other. Given these hypotheses, *prototypicality* is the only exogenous latent variable.

Table 1 The number of survey questions per construct. The wording of questions can be found in [24].

Construct	n	Example
Overall Sal.	3	To what extent does this object draw your attention?
Visual Sal.	15	intensity of color / tone / size
Cognitive Sal.	6	To what extent does this object's appearance suggest it to be historic?
Structural Sal.	4	How easy is it for you to refer to this object in a route description?
Visibility in Adv.	4	To what extent can one easily refer to this object from afar?
Prototypicality	3	To what extent does this object represent your impression of such objects?

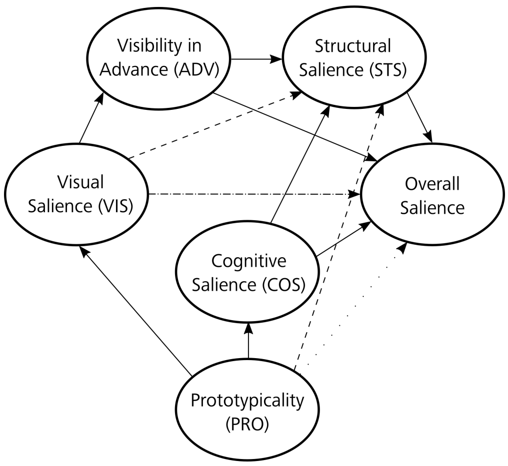


Figure 1 The structural relationships of the theoretical model. The dotted line $PRO \rightarrow Overall$ reflects the full mediation via VIS. The paths $VIS \rightarrow STS$ and $PRO \rightarrow STS$ are added in order not to inflate unexplained variance. Finally, the path $VIS \rightarrow Overall$ was dashed and dotted in order to indicate that a partial mediation of this effect is hypothesized. The figure was drawn using Inkscape [38].

H_1–H_5. Each of the subdimensions contributes positively to *overall salience*.

H_6. The greater an object's *visual salience*, the easier it is to see from advance.

H_7. The greater an object's *visibility in advance*, the more suitable it is to be included in route instructions.

H_8. The greater an object's *prototypicality* ,the larger its *cognitive salience* is.

H_9. The greater an object's *cognitive salience*, the easier it is to be integrated in route instructions.

H_{10}. The effect *prototypicality* has on *overall salience*, is mediated by *visual salience*.

These hypotheses reflect a proposed three-path mediated effect[1] for *visual salience*: Visual aspects become salient at a very early stage of human perception and are consistent across individuals (cf. [20, 5]). Hence, they determine whether or not, as well as to what extent other subdimensions are affected by it. The positive impact *visual salience* has on *overall*

[1] It is important to note that a number of assumptions regarding correctness apply to three-path mediated models, cf. [37, p. 265] for details.

salience is modeled to be partially mediated by *visibility in advance*, which in turn has a positive influence on *structural salience*, which is positively related to *overall salience*, too. A rationale to propose a positive influence of *visibility in advance* on *structural salience* can be based on the understanding of *visibility in advance*. Basically, objects that 'are identifiable early on along a route are more useful than those that can only be spotted at the very last moment' [33, p. 142]. [4] found strong evidence that salient objects in unknown environments must be first and foremost recognizable, a property that relies mostly on the visual features in a given context. Additionally, [27] reports on the strong influence *visual salience* has on object recognition (imagine, e.g. a blue colored house in a neighborhood, where all other houses are painted white). Furthermore, the hypotheses presented indicate a multiple mediation for *prototypicality*. On the one hand, it is mediated by *visual salience*, which is reasonable based on the fact that mental images of objects may well guide our visual attention on the pre-attentive level (cf. [43]). On the other hand, *prototypicality* is supposed to have a positive influence on *cognitive salience* because prototypical objects may eventually be conceptualized more easily. This presumably has, in turn, a positive effect on the value the object has for use in route instructions, i.e. on *structural salience*. As it is common not to model direct paths in mediator analysis [45, pp. 204–205], it must be stressed that this is done purposefully in the hypotheses H_1 to H_5. Based on prior empirical evidence full mediation cannot be assumed. It is important to note, moreover, that these hypotheses are motivated by the aim of establishing a causal chain, which is a major difference to existing models. [31] propose different weights for visual, semantic and structural attraction based on its significance. This means, they do not account for any kind of impact that measures may have on one another. Similarly, the Bayesian network presented in [5] does not include any connections between high-level components such as *visual salience* or *cognitive salience*.

4 Method

As Structural Equation Modeling in general and PLSc in particular are currently not widespread in GIScience research, some general remarks on this method are appropriate. In opposition to that, Bayesian networks (BNs) are much more common and, therefore, only few remarks regarding the algorithm applied to learn the structure of the latent variables network and the steps used to combine BNs and PLSc approaches are given. This section ends with a short description of the in-situ, survey-based data acquisition method according to [24].

4.1 A Rational to use Structural Equation Models

All current models of *salience* share one important aspect: *Salience* is always viewed as having multiple subdimensions. The hypotheses presented (cf. Section 3) lead to a model including multiple relationships between multiple constructs. As a consequence, a statistical method is needed which allows for the use of all available information concurrently. In contrast to factor analysis, multiple regression or MANOVA approaches, Structural Equation Modeling (SEM) has these capabilities. The relations between several latent variables in a so-called structural model can be assessed simultaneously accompanied by the measurement models proposed for each of these constructs (cf. [16]). This means, in contrast to exploratory factor analysis, where no measurement model specification is required at all [17, p. 641], SEM analysis requires a specification of dependencies according to theory. Using latent, i.e. not directly measured, variables to build a model is particularly sensible as the use of multiple indicators to measure a single variable reduces measurement error [17, p. 635]. While covariance-

(commonly referred to as LISREL, cf. [23]) and variance-based methods (commonly referred to as PLS Path Modeling, cf. [42]) to assess models exist, the variance-based approach, i.e. PLS Path Modeling, is used here. There are two reasons for this decision: First, PLS Path Modeling allows for formative measurement and *visual salience* was modeled to be measured formatively[2]. Second, PLS Path Modeling is particularly suitable to assess the degree of influence each subdimension has in terms of predicting both, each other and *overall salience*, thereby virtually making no assumptions about the distribution of the data (cf. [7]). In accordance with recent methodological advancements (cf. [12, 13]) – and, therefore, in contrast to [24], where non-consistent PLS Path Modeling was used – PLS Path Modeling in its consistent version (PLSc) using ADANCO (cf. [8]) is applied. PLSc comprises four steps (cf. [12] for a detailed account):

1. Run the PLS-SEM algorithm, which is alternating the estimation of the measurement model and the structural model estimation until convergency.
2. Calculate ρ_A for all reflective latent variables (i.e. set $\rho_A = 1$ for those modeled formatively).
3. Correct the correlations of latent variables obtained in step one to find consistent correlations.
4. (Re-)Estimate path coefficients using the correlations found in step 3.

4.2 Why combine Bayesian Networks and consistent PLS-SEM – and how

As mentioned above (cf. Section 4.1) the structural model part in SEM must generally be specified prior to a PLSc analysis. It allows hypotheses to be tested with respect to the way latent variables influence each other. However, as these hypotheses are based on theoretical considerations solely it is interesting to investigate whether data driven methods yield similar results. BNs are particularly useful in this context. Their network structure can either be predefined or derived from input data (cf. e.g. [22]). The latter case is particularly useful to establish an empirically based structural model. Following the method of combination suggested in [44], PLSc and BN analyses are linked based on a two-step procedure.

1. Learn a network structure between latent variables from data using Tree-Augmented Naïve Bayes as a search algorithm in WEKA [14].
2. Use the network structure as input for a subsequent PLS-SEM analysis using ADANCO [8].

While WEKA implements several different search algorithms (e.g. K2, C4.2, Naïve Bayes) tree-augmented naïve Bayes (TAN) is particularly suitable for the current research questions. [15] provides evidence that TAN is capable to achieve stable results for correlated attributes while yielding a directed acyclic graph with a singular top level node. It, therefore, allows for an increase in network structure complexity (cf. [44, p. 136]). At the same time, [22] stress that, compared to Naïve Bayes, common measures of classification analyses are significantly increased if TAN is applied (cf. [44, p. 136]).

Found differences or commonalities between the theoretical and the empirical model yield insights into the degree and the way subdimensions of *overall salience* influence each other.

[2] While the ongoing discussion about formative measurement in general (cf. e.g. [2]) cannot be detailed here, a major difference to reflective measurement shall be given: Formative causes must not be mutually interchangeable (cf. [21, p. 203]). From my point of view, the dimensions found to be important to *visual salience* in earlier studies (cf. [24] for a comprehensive list) are not interchangeable, but all of them contribute to *visual salience*. Hence, this subdimension was modeled formatively

4.3 Data acquisition

The data used in this paper are user ratings of a a large-scale, in-situ, survey-based study. The 361 objects to be rated were selected based on randomly chosen geographic coordinates, yielding a variety of objects, two thirds of which comprise buildings and the remaining third a large variety of other urban objects, fences, post boxes and benches among them. Each participant was guided by the first author on one of 55 different routes (routes may have had overlapping segments) which the chosen objects were randomly assigned to. The trials took 60 *min* on average and routes showed a mean length of 1.5 *km*. Participants rated 7 objects by answering 41 German language questions (see [24] for the comprehensive list and Table 1 for examples) on a five-point Likert scale for each object. Participants were required to spot the object presented to them using a photo shown on a 7 inch tablet themselves. Two ratings per object were collected and all calculations were done on the average of both ratings for each variable in order to counterbalance potential bias due to personal preferences. More details about this data can be found in [25].

5 Results

For the sake of readability of tables three letter acronyms for each of the (sub-)dimensions of *salience* are used throughout this section: ADV ::= *visibility in advance*, COS ::= *cognitive salience*, PRO ::= *prototypicality*, $OVSAL$ ::= *overall salience*, STS ::= *structural salience*, VIS ::= *visual salience*. First, a short glance on PLSc measurement model results is provided. Second, the theoretical structural model is assessed. Third, the estimation results of a structural model resulting from a prior BN analysis are presented.

5.1 Measurement Model Results

As the focus of this paper is on ways subdimensions of *salience* influence each other only a short report about the measurement model results is given. It is necessary, though, as [24] reports results based on PLS instead of PLSc. Formative measurement model results, however, are not affected by this shift in estimation methods. Therefore, *visual salience* is not discussed below. Table 2 presents standard measures for the reflectively measured[3] latent variables. The figures indicate well-fitting measurement models except for *cognitive salience*. For this subdimension common thresholds are neither met for Cronbach's α ($\alpha < .0.6$, cf. [17, p. 92]) nor for rho_A ($rho_A < 0.7$, cf. [18, p. 12]) nor for AVE ($AVE < 0.5$, cf. [17, p. 688], i.e. the latent variable explains, on average, less than 50% of the variance present in its measured variables). The figures indicate that *cognitive salience* was revealed to be a latent variable with a meaning, difficult for people to grasp. The HTMT-values[4] (cf. [19]) suggest a good discriminant validity of the reflective latent variables (cf. Table 3).

All HTMT-values achieved are significantly lower than one at a significance level of $\alpha = 0.01$. However, despite the significant difference to one, the HTMT-values for ADV and $OVSAL$, for ADV and STS and for STS and $OVSAL$ are large. This suggests that these

---

[3] Measured variables are considered as effect indicators, i.e. they 'share [. . .] [a] common cause' [10, p. 12] in case of reflective measurement, which is, therefore, often referred to as *common factor model*.

[4] The HTMT is defined in [19, p. 121] 'as the average of the heterotrait-heteromethod correlations (i.e., the correlations of indicators across constructs measuring different phenomena), relative to the average of the monotrait-heteromethod correlations (i.e., the correlations of indicators within the same construct). Since there are two monotrait-heteromethod submatrices, we take the geometric mean of their average correlations'.

Table 2 Cronbach's α, Dijkstra-Henseler's ρ_A and Average Variance Extracted (AVE) for each of the reflectively measured latent variables.

Method	OVSAL	PRO	COS	STS	ADV
Cronbach's α	0.922	0.849	0.589	0.890	0.900
Dijkstra-Henseler ρ_A	0.923	0.875	0.622	0.900	0.916
AVE	0.800	0.753	0.341	0.700	0.684

Table 3 The bootstrapping results for HTMT-values of reflective constructs. *** indicates $p < 0.001$. A significant result means that the HTMT-value is significantly smaller than one.

	COS	PRO	OVSAL	STS
ADV	0.547***	0.373***	0.815***	0.881***
COS		0.293***	0.694***	0.566***
PRO			0.394***	0.346***
OVSAL				0.831***

constructs are interrelated – a fact further examined by means of the mediation analysis reported below. Overall, the measurement models show a good fit and the items, consequently, provide a sound basis for further structural model analyses. In particular the items derived for *overall salience* show desirable properties, which is important, as all other items are used to measure this particular value.

5.2 Structural Model Results

A two step approach is taken in providing structural model results: First, PLSc figures for the theoretical model are presented. Second, a structural model using TAN involving the subdimensions of *salience*, is learned and assessed based on PLSc.

5.2.1 Theory-based Structural Model

Table 4 presents figures about the size of direct, indirect and total effects constructs have on each other according to the theoretical model (cf. Figure 1).

The figures show:

1. that visual dimensions have the largest effect on *overall salience* and that this effect is only partially mediated via ADV and STS, because both, the direct and indirect effect *visual salience* has on *overall salience* are significant;
2. that *visual salience* has a very large effect on *visibility in advance*, which in turn has a very large effect on *structural salience* while the direct effect $VIS \rightarrow STS$ is rendered insignificant, i.e. the more salient the visual features of an object are, the easier it can be recognized from afar and the easier it is to be referred to in route instructions;
3. *prototypicality* has a significant but small effect on *overall salience*, whereas its effect on *cognitive salience* is medium sized;
4. *cognitive salience* does not substantially add to capturing *overall salience*.

The adjusted R^2-values of endogenous constructs ($R^2(OVSAL) = 0.92$, $R^2(STS) = 0.82$, $R^2(ADV) = 0.56$, $R^2(COS) = 0.13$) reveal, that the subdimensions have a very high predictive relevance for *overall salience*. On the other hand, they further support the influence visual attributes and *visibility in advance* have on ease of reference in route instructions. Finally, the small amount of variance explained in *cognitive salience* shows that the degree of prototypicality is not enough to explain as to why an object is seen as historical etc., although prototypicality has a medium sized effect on this construct.

Table 4 Direct, indirect and total effects of the theoretical model. *** indicates $p < 0.001$, ** indicates $p < 0.01$ and * means $p < 0.05$ ($K = 5000$ resamples).

Effect	Direct	Indirect	Total	Cohen's f^2	Hypotheses
ADV → STS	0.773***	n/a	0.773***	1.410	H_7 holds
VIS → ADV	0.750***	n/a	0.750***	1.278	H_8 holds
STS → OVSAL	0.234**	n/a	0.234**	0.123	H_3 holds
PRO → VIS	0.094**	n/a	0.094**	0.025	H_{10} holds partially
VIS → OVSAL	0.634***	0.220***	0.854***	1.527	H_1 holds
ADV → OVSAL	0.090n.s.	0.181**	0.271***	0.018	H_4 holds
PRO → COS	0.368***	n/a	0.368***	0.157	H_8 holds
COS → OVSAL	0.060n.s.	0.027n.s.	0.087n.s.	0.018	H_2 holds not
VIS → STS	0.075n.s.	0.579***	0.654***	0.010	
COS → STS	0.116n.s.	n/a	0.116n.s.	0.031	H_9 holds not
PRO → OVSAL	0.040n.s.	0.110**	0.150***	0.017	H_5 holds
PRO → STS	−0.006n.s.	0.104**	0.097*	0.000	

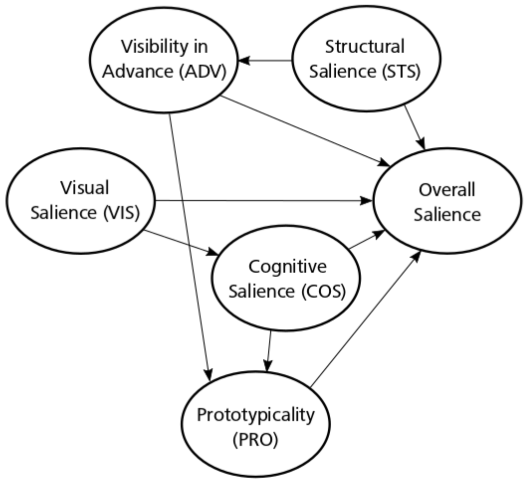


Figure 2 The structural model resulting from a Bayesian Network analysis using TAN as search algorithm.

Overall, the results stress the model's plausibility. However, as stressed by Hair et al. (cf. [17, p. 647]), there are always at least two models, which demonstrate an equally good fit in SEM analyses.

5.2.2 Bayes Net based Structural Model

In order to cross-check the results achieved so far a structural model is divised based on a BN analysis using TAN as a search algorithm. With this goal in mind a multiple regression analysis to calculate the *visual salience* for each of the objects was applied first. This method is reasonable due to the fact that formative measurement was used for *visual salience*. Second, values for all remaining subdimensions were calculated as means of all items associated with a particular dimension – which is in line with the common understanding of reflective measurement as all items reflect the latent variable and their mean provides a most suitable proxy, consequently (cf. e.g. [11]). The structural model resulting from the TAN search based on these figures is shown in Figure 2 while the numerical results are given in Table 5.

Only two direct effects on *overall salience* are rendered significant in this case. *Visual salience* shows a significant, large direct effect on *overall salience* and *structural salience* has

Table 5 Direct, indirect and total effects of the structural model derived by means of a Bayesian Network analysis using TAN as search algorithm. Cohen's f^2 values refer to the direct effects. *** indicates $p < 0.001$, ** indicates $p < 0.01$, * means $p < 0.05$ ($K = 5000$ resamples).

Effect	Direct	Indirect	Total	Cohen's f^2
VIS → OVSAL	0.644***	0.033n.s.	0.677***	1.580
COS → OVSAL	0.036n.s.	0.009n.s.	0.045n.s.	0.007
PRO → OVSAL	0.047n.s.	n/a	0.046n.s.	0.022
ADV → OVSAL	0.090n.s.	0.013n.s.	0.102n.s.	0.017
STS → OVSAL	0.242**	0.092n.s.	0.334***	0.131
COS → PRO	0.183*	n/a	0.183*	0.026
VIS → COS	0.732***	n/a	0.732***	1.152
ADV → PRO	0.267***	n/a	0.267*	0.056
STS → ADV	0.899***	n/a	0.899***	4.210
VIS → PRO	n/a	0.134*	0.134*	n/a
STS → PRO	n/a	0.240***	0.240**	n/a

a medium sized effect. This construct has a very large impact on *visibility in advance*, too. Furthermore, this model reveals a strong impact *visual salience* has on *cognitive salience*. In terms of variance explained ($R^2(OVSAL) = 0.92$, $R^2(COS) = 0.53$, $R^2(PRO) = 0.16$, $R^2(ADV) = 0.81$) the TAN-based model can explain an equal amount of variance in *overall salience* as compared to the theoretical model. *Visual salience* accounts for half of the variance present in *cognitive salience* which stresses its importance.

6 Discussion

From the beginning of *salience* theory, weights for the different subdimensions have been incorporated (cf. [31]). However, studies trying to estimate weights are rarely found nor do they simultaneously take all subdimensions into account. This shortcoming is overcome by the current analysis based on an in-situ dataset (as compared to online studies like [40] or those conducted in virtual reality environments such as [35]). Although the evidence-based structural model and the theoretical model presented show major differences, the total effect of *visual salience* is large in both cases. This finding is in line with other studies in the broader field of research on *salience*. For example, [9] study the importance of *visual salience* for the strategies used to orient oneself in a real-world spatial environment using different kinds of maps. They provide evidence for the high distractive impact visually salient objects have on the orientation of map viewers. Furthermore, the influence *structural salience* and *visibility in advance* have on each other is similar to earlier findings, where objects located at intersections and their resulting *structural salience* have drawn particular interest in recent years. For example, [35, p. 146] finds that participants prefer those 'landmarks that were located in the direction of turn' in case of cross-intersections. However, whether *structural salience* affects *visibility in advance* or vice versa is not evident from the statistical results of both models.

In general, the results provide sound empirical evidence that the subdimensions of *overall salience* are not equally important and highly intertwined. This is in clear contrast to the assumptions of independence made in [5]. Similarly, the results of the analysis presented are in contrast to the findings in [24], where a model with acceptable predictive capabilities is presented in which subdimensions are independent. This shows, first, the importance to assess different models based on the same data. Second, the differences may stem from the fact that non-consistent PLS-SEM was used then which is now outdated. As the models presented here are capable to explain a larger proportion of the variance present in *overall salience*, they are to be preferred. A major difference between models found here, however, is the direct effect *visual salience* has on *cognitive salience* in case of the BN-based model. This effect is particularly reasonable, though: Visual aspects are rendered salient in an early stages of perception (cf. Section 3) whereas *cognitive salience* needs conscious cognitive processing. However, the impact visual dimensions have on *overall salience* is not mediated through *cognitive salience*.

Given these statistical results presented important subdimensions other than those proposed in common theories may be missing. One candidate dimension is *emotional salience*, which has recently gained importance particularly in psychological research. [28, p. 13:1] show that '[e]motional salience can override visual salience and can determine attention allocation in complex scenes.'. By means of a lab-based VR study [3] find evidence wayfinding performance is enhanced by those landmarks with which negative emotions are associated, whereas positive emotions foster route learning. Another dimension worth investigating is *familiarity*. [30] reveal *visual salience*, *structural salience* and *semantic salience* to have an impact on all participants, but those who are familiar with the study area prefer objects which have a meaning for them. *Familiarity*, however, may be hard to distinguish from *emotional salience* or may at least have an impact on it. Imagine the object to be rated is a person's school house. This object is certainly familiar to her/him, but it is also likely to evoke emotional affect due to this familiarity. Further analysis of the dimensions of *emotional salience*, however, is necessary to substantiate this claim.

7 What Do Found Differences Mean – Conclusion and Future Work

This study uses state-of-the-art theories about *salience* to investigate the way commonly accepted subdimensions of *salience* influence each other. In doing so, the nature of the study is, at the same time, both theoretical and empirical in nature. It proposes hypotheses about causal relationships between *overall salience*, *visual salience*, *visibility in advance*, *prototypicality*, *structural salience*, and *cognitive salience*. Then, survey-based ratings of 361 different objects collected in-situ (cf. [24]) are used to assess the predictive capabilities of the model. The structural relationships between the subdimensions are double checked by combining Baysian networks and consistent PLS-SEM. Using TAN as a search algorithm, an empirically based structural model is created by means of a Bayes Network analysis and estimated using consistent PLS-SEM. The results of both, the theoretical and the data-driven model, are not contradictory in terms of effect size and amount of variance explained. Indeed, an important effect of visual dimensions is found, which is in line with results of earlier studies. However, some differences with respect to paths and their causal direction are found. As a consequence, future work will be guided along three lines of research. First, we are currently working on data acquisition in a city environment different to the one described in [24] in order to further evaluate the stability of sizes and directions of effects. Second, lab-based, controlled studies are planned in order to further investigate the direction of influence between *structural salience* and *visibility in advance* dimensions. Third, several experiments will be divised to find ways of capturing *emotional salience* (and other personal factors) and to understand its impact.

Acknowledgements. I'm grateful to Bernd Ludwig who pointed me to the Baysian network analysis and to David Elsweiler for his valuable comments on a draft version. Furthermore, thanks to four anonymous reviewers whose comments helped to improve this paper. Finally, my sincere gratitude to 120 persons willing to participate during winter time.

References

1 Donald Appleyard. Why Buildings Are Known: A Predictive Tool for Architects and Planners. *Environment and Behavior*, 1(1):131–156, 1969.

2 Sierra A. Bainter and Kenneth A. Bollen. Interpretational Confounding or Confounded Interpretations of Causal Indicators? *Measurement: Interdisciplinary Research and Perspectives*, 12(4):125–140, 2014. doi:10.1080/15366367.2014.968503.

3 Ceylan Z. Balaban, Florian Röser, and Kai Hamburger. The effect of emotions and emotionally laden landmarks on wayfinding. In P. Bello, M. Guarani, M. McShane, and B. Scasselati, editors, *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Austin, Texas*, pages 3315–3320. Cognitive Science Society, 2014.

4 Nicola J. Bidwell, Christopher Lueg, and Jeff Axup. The territory is the map: designing navigational aids. In *Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-human Interaction: Making CHI Natural*, CHINZ'05, pages 91–100, New York, NY, USA, 2005. ACM.

5 David Caduff. *Assessing Landmark Salience for Human Navigation.* PhD thesis, Mathematisch-naturwissenschaftliche Fakultät der Universität Zürich, Zürich, 2007.

6 David Caduff and Sabine Timpf. On the assessment of landmark salience for human navigation. *Cognitive Processing*, 9(4):249–267, 2008. doi:10.1007/s10339-007-0199-2.

7 C. Cassel, P. Hackl, and A. H. Westlund. Robustness of partial least squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, 26(4):435–446, 1999.

8 Composite Modeling GmbH & Co. KG. ADANCO 2.0, 2015. last access on May 25th, 2017. URL: http://www.composite-modeling.com/.

9 Clare Davies and Davies Peebles. Spaces or Scenes: Map-based Orientation in Urban Environments. *Spatial Cognition & Computation*, 10(2-3):135–156, 2010.

10 Robert F. DeVellis. *Scale Development. Theory and Applications.* SAGE Publications, Thousand Oaks, CA et al., 3rd edition, 2012.

11 A. Diamantopoulos and J. A. Siguaw. Formative Versus Reflective Indicators in Organizational Measure Development. *British Journal of Management*, 17(4):263–282, 2006.

12 Theo K. Dijkstra and Jörg Henseler. Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, 81:10–23, 2015.

13 Theo K. Dijkstra and Jörg Henseler. Consistent Partial Least Squares Path Modeling. *Management Information Systems Quarterly*, 39(2):297–316, 2015.

14 Eibe Frank, Mark A. Hall, and Ian H. Witten. *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*. Morgan Kaufmann, 4th edition, 2016.

15 Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29:131–163, 1997.

16 David Gefen, Detmar Straub, and Marie-Claude Boudreau. Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, 4:Article 7, 2000.

17 Joseph F. Hair, William C. Black, Barry J. Babin, and Rolph E. Anderson. *Multivariate Data Analysis. A Global Perspective.* Global Edition. Person Education, Upper Saddle River, NJ, 7th edition, 2010.

18 Jörg Henseler, Geoffrey Hubona, and Pauline Ash Ray. Using PLS path modeling in new technology research: updated guidelines. *Industrial Managment & Data Systems*, 116(1):2–20, 2016.

19 Jörg Henseler, Christian M. Ringle, and Marko Sarstedt. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1):115–135, 2015.

20 L. Itti. Visual salience [rev 72776]. *Scholarpedia*, 2(9):3327, 2007.

21 Cheryl B. Jarvis, Scott B. MacKenzie, and Phil M. Podsakoff. A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research*, 30:199–218, 2003.

22 Liangxiao Jiang, Harry Zhang, Zhihua Cai, and Jiang Su. Learning tree augmented naive bayes for ranking. In *Database Systems for Advanced Applications*, pages 688–698. Springer Nature, 2005.

23 Karl Jöreskog. Simultaneous Factor Analysis in Several Populations. *Psychometrika*, 36:409–426, 1971.

24 Markus Kattenbeck. Empirically Measuring Salience of Objects for Use in Pedestrian Navigation. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS'15, pages 3:1–3:10, New York, NY, USA, 2015. ACM.

25 Markus Kattenbeck. *Empirically Measuring Salience of Objects for Use in Pedestrian Navigation.* Dissertation, Lehrstuhl für Informationswissenschaft, Universität Regensburg, 2016. URL: `http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:bvb:355-epub-341450`.

26 Alexander Klippel and Stephan Winter. Structural salience of landmarks for route directions. In *Proceedings of the 2005 International Conference on Spatial Information Theory*, COSIT'05, pages 347–362. Springer-Verlag Berlin / Heidelberg, 2005.

27 Jared Miller and Laura Carlson. Selecting landmarks in novel environments. *Psychonomic Bulletin & Review*, 18:184–191, 2011.

28 Yaqing Niu, Rebecca M. Todd, Matthew Kyan, and Adam K. Anderson. Visual and Emotional Salience Influence Eye Movements. *ACM Trans. Appl. Percept.*, 9(3):13:1–13:18, 2012.

29 Clemens Nothegger, Stephan Winter, and Martin Raubal. Selection of Salient Features for Route Directions. *Spatial Cognition & Computation*, 4(2):113–136, 2004.

30 Teriitutea Quesnot and Stéphane Roche. Quantifying the Significance of Semantic Landmarks in Familiar and Unfamiliar Environments. In SaraIrina Fabrikant, Martin Raubal, Michela Bertolotto, Clare Davies, Scott Freundschuh, and Scott Bell, editors, *Spatial Information Theory*, volume 9368 of *Lecture Notes in Computer Science*, pages 468–489. Springer International Publishing, 2015.

31 Martin Raubal and Stephan Winter. Enriching Wayfinding Instructions with Local Landmarks. In Max Egenhofer and David Mark, editors, *Geographic Information Science*, Lecture Notes in Computer Science, pages 243–259. Springer, Berlin / Heidelberg, 2002.

32 Kai-Florian Richter. Prospects and Challenges of Landmarks in Navigation Services. In Martin Raubal, David M. Mark, and Andrew U. Frank, editors, *Cognitive and Linguistic Aspects of Geographic Space. New Perspectives on Geographic Information Research*, Lecture Notes in Geoinformation and Cartography, pages 83–97. Springer, Heidelberg et al., 2013.

33 Kai-Florian Richter and Stephan Winter. *Landmarks. GIScience for Intelligent Services.* Springer International Publishing, 2014.

34 E. Rosch, C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem. Basic Objects in Natural Categories. *Cognitive Psychology*, pages 382–439, 1976.

35 Florian Röser. *The cognitive observer-based landmark-preference model – What is the ideal landmark position at an intersection?* Fachbereich 06: Psychologie und Sportwissenschaften, Justus-Liebig-Universität Giessen, 2015. last access May 2nd, 2016.

URL: http://geb.uni-giessen.de/geb/volltexte/2015/11640/pdf/RoeserFlorian_2015_07_15.pdf.

36 Molly Sorrows and Stephen Hirtle. The Nature of Landmarks for Real and Electronic Spaces. In Christian Freksa and David Mark, editors, *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, Lecture Notes in Computer Science, pages 37–50. Springer, Berlin / Heidelberg, 1999.

37 Aaron B. Taylor, David P. MacKinnon, and Jenn-Yun Tein. Tests of the Three-Path Mediated Effect. *Organizational Research Methods*, 11(2):241–269, 2008.

38 The Inkscape Team. Inkscape 0.91, 2016. last access May 25th, 2017. URL: https://inkscape.org/de/.

39 Stephan Winter. Route Adaptive Selection of Salient Features. In Walter Kuhn, Michael Worboys, and Sabine Timpf, editors, *Spatial Information Theory. Foundations of Geographic Information Science*, Lecture Notes in Computer Science, pages 349–361. Springer, Berlin / Heidelberg, 2003.

40 Stephan Winter, Martin Raubal, and Clemens Nothegger. Focalizing Measures of Salience for Wayfinding. In L. Meng, Zipf A., and T. Reichenbacher, editors, *Map-based Mobile Services: Theories, Methods, and Design Implementations*, pages 125–139. Springer Geosciences, 2005.

41 Thomas Wolbers and Mary Hegarty. What determines our navigational abilities? *Trends in Cognitive Sciences*, 14(3):138–146, 2010.

42 Herman Ole Andreas Wold. Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, and V. Capecchi, editors, *Quantitative sociology: International perspectives on mathematical and statistical modeling*, pages 307–357. Academic Press, New York, 1975.

43 J. M. Wolfe and J. M. Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5:1–7, 2004.

44 Wei Wen Wu. Linking Bayesian networks and PLS path modeling for causal analysis. *Expert Systems with Applications: An International Journal*, 37(1):134–139, 2010.

45 Xinshu Zhao, John G. Lynch Jr., and Qimei Chen. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *The Journal of Consumer Research*, 37(2):197–206, 2010.

Evidence-Based Parametric Design: Computationally Generated Spatial Morphologies Satisfying Behavioural-Based Design Constraints

Vasiliki Kondyli[1], Carl Schultz[2], and Mehul Bhatt[3]

1 Human-Centred Cognitive Assistance Research Group, University of Bremen, Bremen, Germany; and
The DesignSpace Group, Germany
2 University of Münster, Münster, Germany; and
The DesignSpace Group, Germany
3 Human-Centred Cognitive Assistance Research Group, University of Bremen, Bremen, Germany; and
The DesignSpace Group, Germany

Abstract

Parametric design is an established method in engineering and architecture facilitating the rapid generation and evaluation of a large number of configurations and shapes of complex physical structures according to *constraints* specified by the designer. However, the emphasis of parametric design systems, particularly in the context of architectural design of large-scale spaces, is on *numerical* aspects (e.g., maximising areas, specifying dimensions of walls) and does not address human-centred design criteria, for example, as developed from behavioural evidence-based studies. This paper aims at providing an evidence-based human-centred approach for defining design constraints for parametric modelling systems. We determine design rules that address wayfinding issues through behavioural multi-modal data analysis of a wayfinding case study in two health-care environments of the Parkland hospital (Dallas). Our rules are related to the environmental factors of visibility and positioning of manifest cues along the navigation route. We implement our rules in FreeCAD, an open-source parametric system.

1998 ACM Subject Classification J.6 Computer-Aided Engineering

Keywords and phrases parametric modelling, behavioural studies, Evidence-Based Design, design computing, wayfinding, spatial cognition

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.11

1 Introduction

Behavioural-based parametric design systems. Parametric modelling is a popular paradigm in the design industry, particularly in the domains of architecture, engineering, and construction: objects are modelled with parameters, constraints are defined between parameters. By "designing by constraints", the designer is specifying a family of designs that satisfy the given set of constraints, and parametric design tools assist designers by providing adaptability and flexibility in the design procedure [13], and enabling them to explore the resulting *design space* in various ways. Two common parametric system tools are *intelligent sketch* and *evolutionary design*.[1]

[1] In intelligent sketch (also called dynamic geometry), a user is able to modify a design e.g. by clicking and dragging objects in a visual representation of their design, and the system automatically adjusts

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 11; pp. 11:1–11:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

While parametric design systems lend themselves well to the manipulation of numerical, geometric features and relationships between object parameters (typically support points, lines, circles, and incidence and orientation constraints), and they have thus far failed in integrating the dimension of human behaviour as a variable of morphological formulation. Currently, all prominent parametric systems are restricted to constraints that are rather geometric in nature, e.g. maximising a numerical volume, fixing the numerical dimensions of walls, and so on. In [21] we extend industry-standard parametric systems to support a range of qualitative and visuo-locomotive spatial constraints: *incidence* (points interior or exterior to regions), *topology* (i.e. Region Connection Calculus), *size* (smaller, larger), *visibility* and *movement*. In Section 5 we use this extended language to formalise evidence-based design rules.

Evidence-based parametric design for large-scale buildings. Designing for large-scale built-up spaces, the architect needs to take into consideration the visuo-locomotive experience of representative groups of people (e.g. children, seniors, individual with physical disabilities) in various circumstances according to the building's functional program. For instance, designing a health-care environment, the architect sets navigation requirements such as *"the moment the user enters the lobby/corridor of a hospital, they should immediately detect the related signage and be confident to proceed in the correct direction"*. In practice, a user's ability to detect signage varies and the spatial structure of the environment plays a major role (Fig. 1). Consequently, we aim at embedding behavioural evidence into the design procedure and simultaneously support the designer in exploring a wide range of morphological possibilities. These objectives lie at the intersection of evidence-based design and parametric computational design. Our perspective on *evidence-based parametric design systems* is rooted in evidence-based design, and aims to ensure that human-centred design objectives are fulfilled (e.g. people should (not) get lost, the environment should satisfy inclusive design criteria) through a computational generative system. This agenda encompasses research in environmental psychology and cognitive-assistive technologies [21, 18].

2 Behavioural evidence from empirical wayfinding studies

Evidence-Based Design for wayfinding. In this paper we investigate the case of wayfinding experience in large-scale built-up spaces, as an example of using bahavioural evidence from a cognitive process to establish design constraints. Design for successful wayfinding performance in large-scale buildings (e.g. hospitals, airports, museums) includes plan configuration and manifest cues, technology, and user characteristics [10, 14]. The significant variables for wayfinding performance that designers can manipulate include spatial characteristics such as visible lines-of-sight, the position of manifest cues, and the geometry of the layout, colors and lighting, visibility connections etc. [23, 20]. For instance, empirical studies in real and virtual space suggest that people tend to move towards the direction of the stated area with the longest line of sight [25], views to the external environment can enhance the legibility of the interiors [11], and that wayfinding includes both attention to the building structure and to manifest cues (landmarks, signage) [3].

the remaining parts of the design in order to maintain the constraints. In evolutionary design, the system automatically generates designs according to the given constraints, i.e. the designer guides this generative process through constraints.

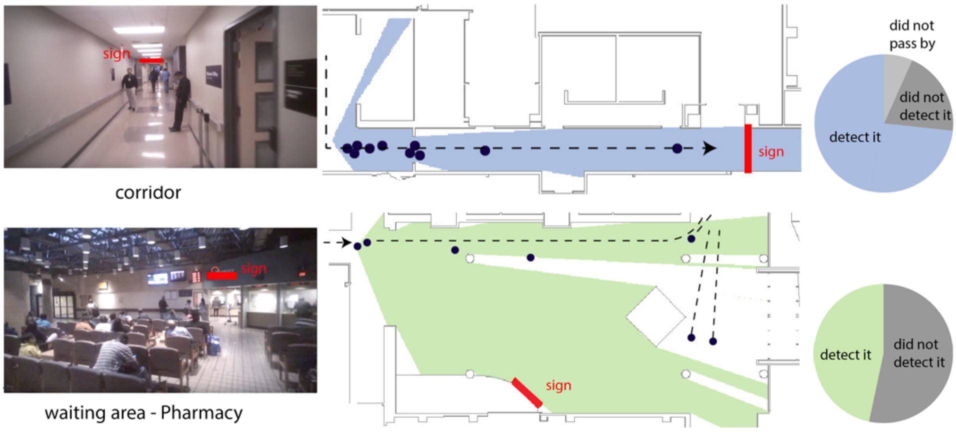


Figure 1 A comparison in behavioural data during wayfinding in a corridor (1.8–2m width) and the pharmacy waiting area (7–8.5m width) in the old Parkland hospital, indicates that the sign above the passage of the corridor was detected by 72% of the participants, while the pharmacy sign (destination point), was detected by 55% of the participants. The isovist analysis reveals visibility differences and justify the behavioural analysis.

Visuo-locomotive experience in a wayfinding case study at the Parkland Hospital. We conducted a wayfinding case study in two health-care environments: the old and the new building of the Parkland hospital in Dallas (Texas). Our study consisted of 25 participants, between 18-83 years old, from the local community that were unfamiliar with the buildings. They were fitted with eye-tracking glasses[2], and were asked to pursue a complex wayfinding task for approximately 15 minutes. With the exception of the vocal instructions given at the beginning of the task, the participants were not allowed to use maps but only the manifest cues (landmarks, signage) which are available in each building. During the experimental procedure we employed a range of sensors for measuring the embodied visuo-locomotive experience of users (mobile eye-tracking, GPS, egocentric and allocentric video recording, questionnaires, manual observations) [7]. Our approach is driven by cognitive vision theory and the high-level semantic analysis of multi-modal perceptual data currently encompassing visual perception analysis, people-movement trajectories based on locomotive path taken by subjects, including other events as well as 3D morphological analysis (e.g., topology, routes, isovists) [8, 17].

3 Integrating empirical and analytical methods to reveal wayfinding issues (I1–I4)

Behavioural analysis of the multi-modal data from our Parkland hospital case study in combination with morphological analysis of the architectural space (a) demonstrates the interaction between users and the Parkland environments, (b) highlights a number of navigation difficulties and uncomfortable situations that participants experienced[3], and (c) reveals environmental features that reduce navigation performance. Many of the outcomes

[2] Wearable eye tracking devices designed to record a person's natural gaze in real-time and capture natural viewing behaviour in real-world environment.

[3] Situations or events that seem to reflect *discomfort* are: time delays, hesitations, detours, or the need to ask for help as well as extensive visual search of the surrounding environment.

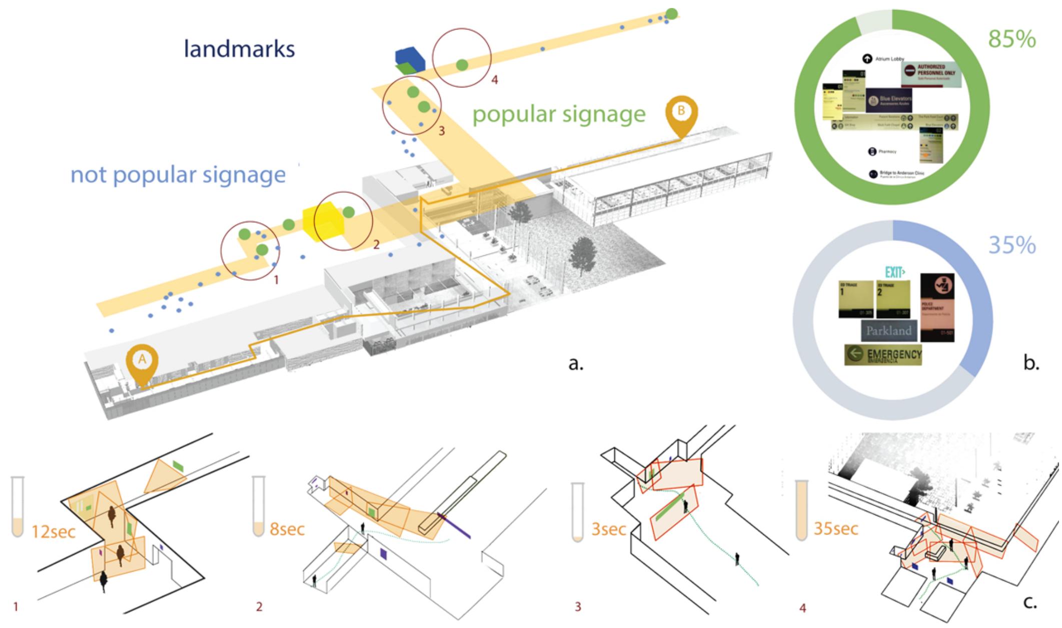


Figure 2 (a) The position of landmarks, manifest cues and decision points along the wayfinding route, (b) signage detection rate, (c) the position of popular (green) and not popular (blue) signage in every decision points and the average time participants spend in each one of them.

confirm prior experimental results about the effect of environmental features on the wayfinding performance concerning, for example, signalization detection, visual connectivity, affordances and manifest cues (landmarks and signage) [24, 2, 12]. In this process, we examine the most, and the least, noticeable signage and landmarks, the time delays at the decision points, gaze patterns in threshold positions[4], visibility connections, the geometry and layout of the scene. Our approach for multi-modal behavioural analysis is founded in Spatial Reasoning, Cognitive Vision and Environmental Psychology [8, 6]. The morphological analysis is based on cognitive design computing foundations resulting in a novel ontology of the *shape of empty space* [5]. As a result, this systematic analysis in the Parkland hospital case study, leads us to highlight four major wayfinding issues (I1–I4).

ISSUE I1 – Signage detection problem at threshold positions. Eye-tracking analysis indicates that out of a total of 60 signs placed along the experimental route in the new Parkland hospital (NPH), only 9 of them have been detected by 85% of the participants, and 6 by less than 35% (Fig. 2b). These results can be interpreted in relation to the morphological analysis of the scene and the layout of the built environment. In particular, the detected signs in NPH, were the ones directly related to the destination and the vocal information given to participants, or they were positioned on decision points vertically along the participant's route (Fig. 2a). Missing signage at a decision point can cause delays, confusion and stress [9]. In the case of NPH, the average time that participants spend at each decision point is directly related to the signage detection rate and the time of the first fixation from the threshold position (Fig. 2c).

[4] Threshold position considers a transitional point between two places in the building, this could signify the entrance to a room of the passage from a corridor to a lobby etc.

To understand how this issue is related to the morphology of the environment and the placing of the signs, we examine the different positions from where participants detect the pharmacy sign while entering the waiting area of the old Parkland hospital, in combination with the line of sight at the moment of detection (Fig. 3a). The distance between the position and the signage varies between 8.7 and 13.5 meters and the viewing angle (formulated by the line of sight and the sign's surface) varies between 10° and 90°. However, the majority of participants detect the sign from an average angle of 78°. Based on DIN-1450 regulations[5] the pharmacy sign (approximately type size 300mm) is readable from 28m distance and visible for an angle between 15° and 90° (Fig. 3b). Even though the distance between the threshold position the sign is less than the suggested by DIN maximum one (15.7 m), the users tent to have difficulty to detect the sign mainly because of the angle formulated between the line of sight of the user in the threshold position and the line representing sign's surface (137°) (Fig. 2a).

ISSUE I2 – Landmarks are not efficient for wayfinding if their position is not related to spatial geometry. The detection of a landmark is based on its position, its size or its differentiation from the environment [16]. Landmarks are important for basic development of spatial knowledge and they enable users to connect fragments of spatial memory in a cognitive map [19]. The results of the behavioural data analysis in the new Parkland hospital indicate that outdoor landmarks in combination with established visual connectivity along the route serve to explain the success of the orientation pointing task that took place after the users changed floors. Additionally, by analysing the visual patterns of participants we observe that they tend to fixate on the outdoor landmarks when these appear in participants' "comfortable" visual range during locomotion. However, in the case of a landmark positioned at a crocked corridor in the old Parkland hospital, 30% of the participants hesitated, slowed their pace, or detoured and asked for help despite the instructions about the landmark in the beginning of the task. These observations show that landmarks are not always helpful in navigation, and that spatial structure must also be considered.

ISSUE I3 – Important manifest cues are not included in the fixation zone of participants. The analysis of participants' gaze directions and fixation patterns reveals a zone of visual search that changes dynamically according to locomotion. In the second decision point of the new Parkland hospital (Fig. 2) the integrated fixation map, for the group of participants, demonstrates that the fixation zone is formulated according to the average comfortable visual range (60° arc) in the moment when participants pass the threshold position (Fig. 4a). As a result, a major signalization text on the right which is not included in the zone was not detected by a large number of participants or it was detected with delay. Additionally, we observe that the zones that map visual attention are changing respectively to the geometric changes of the environment along the route. Specifically, the fixation zones formulated in the corridors of the new Parkland hospital in comparison to the ones from the atrium lobby (Fig. 4b) are narrower in the vertical axis. the comparison between the fixation zones - generated on average by the participants - in two corridors with same dimensions, made of different materials (walls and transparent surfaces) (Fig. 4c) demonstrates a difference on the horizontal axis. As a result, we conclude that the fixation zones created by the participants

[5] DIN (Deutsches Institut für Normung) is the German standards body, and specifically DIN 1450 refers to legibility of texts.

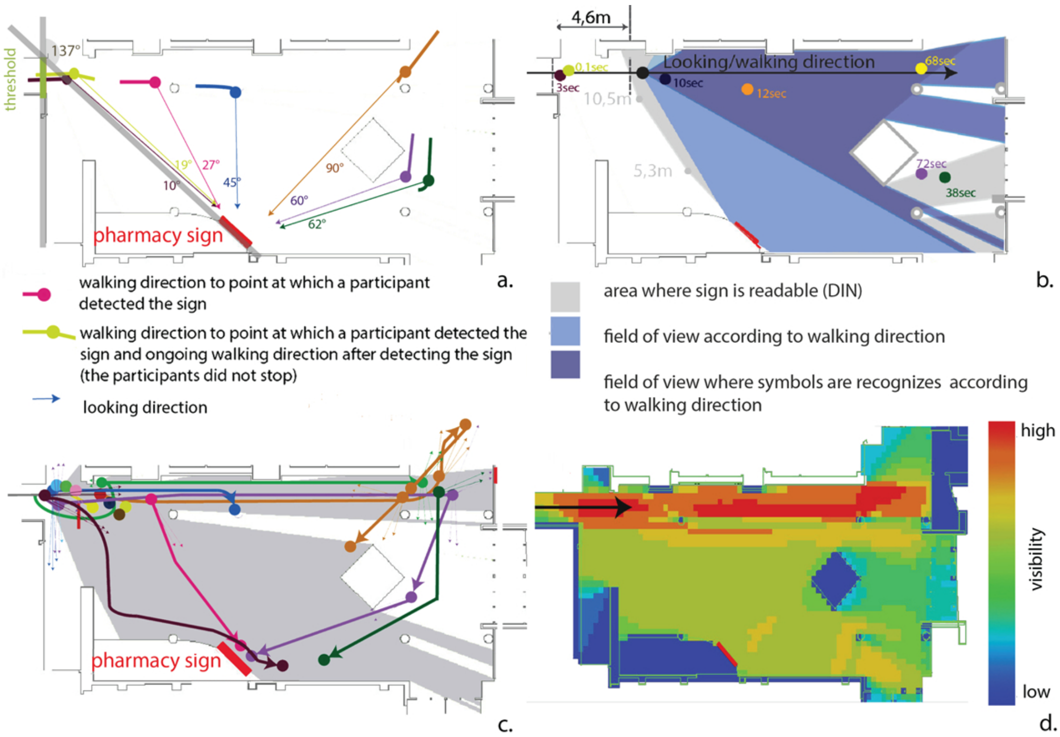


Figure 3 (a) The position of participants, the line of sight and the angle of their view, when they detect the sign, (b) the range of visibility and readability according to DIN regulation in the particular layout and the dimensions of the sign, (c) Isovist graph from the threshold position, (d) Visibility graph from the threshold position.

during the first moments of a visual search are related to the geometry of the space and constitute an important factor for signage and landmarks placement in space.

ISSUE I4 – Participants unconsciously move towards the direction with the longer line of sight. Our observations confirm the argument of Wiener [25] that people tend to move towards the direction with the longer line of sight. In the forth decision point (Fig. 2) of the new Parkland hospital, 70% of the participants did not detect the sign at the threshold position. The eye-tracking analysis shows that participants' visual attention was placed on the open corridor on their right, towards the end of the available field of view (Fig. 5a,c) before they also decide to move towards this direction (Fig. 5b). Moreover, the behavioural analysis for the old Parkland hospital suggests that many people walking on the narrow corridors of the hospital, tend to first observe the farther visual cues immediately after entering a new space, and they also tend to get distracted by several openings along the route (doors, crossroads, windows, glass walls). These outcomes indicate that user's visual attention and decision making could be unconsciously guided by the visual cues under specific circumstances such as distraction or confusion.

4 Evidence-based design rules (R1–R4)

The results of the multi-modal analysis of the wayfinding case study led us to extract some of the major issues that degrade users' navigation performance. Based on these observations

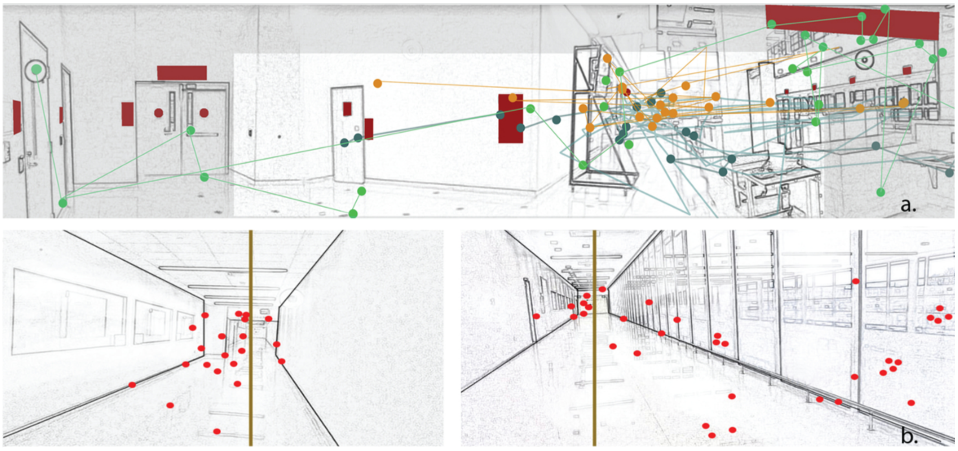


Figure 4 (a)The highlighted zone concentrates the average fixations of the participants from the threshold position at the 2nd decision point in new Parkland hospital. The major sign on the top right is not included in this zone, (b) fixations zones in a corridor and in a corridor with a glass wall and view towards the restaurant of the hospital.

we define design requirements that address navigation issues (I1–I4). We present these requirements in a form of design rules (R1–R4) with the scope to transform them into geometric constraints that can be formulated within parametric design systems.

RULE R1 – The manifest cues should be detected from the threshold positions. The term 'visual field' refers to human's visual abilities concerning the degrees of visual angle during a stable fixation [1, 22]. Humans have an sightly over 180° forward-facing horizontal diameter of their visual field. Their binocular vision covers 114° and concequently the zone where human fixates (fixation is directly related to perception and cognition) [26, 15] is 60° in the horizontal axis and 55° in the vertical axis (Fig. 6a) from which the central 5° represent the normal line of sight each moment. These dimensions create a cone of view, the area where humans are able receive visual information from the surrounding space (Fig. 6b). According to the DIN-1450 concerning signalization text in a public building, the seeing angle[6] is considered different than the viewing angle. The regulations indicate that the legibility depends on the size the signalation text in relation to the distance of viewing and the angle in the horizontal and the vertical axis[7] (Fig. 6c).

As a result, to reassure visibility or readability of a sign, based on humans' visual perception and DIN regulations, the necessary variables to consider are the distance, the viewing angle and the size of the signage. Moreover, based on the behavioural observations, threshold positions are significant for wayfinding. So, the rule (R1) suggests that the manifest cues should be included in the visual range of the user or on the limit of the viewing arc, as this is developed in a threshold position. Considering that this range is defined by the angle of 60° (with central line identical to the route vector), and the radius of this arc is the max distance (based on DIN regulation) so that the size of the particular size is visible.

[6] This is the angle with vertex at the eye and the sides surround the object to see, it is measured in arc minutes ($1' = (\frac{1}{60})°$), minimum for the seeing angle with which the middle length can be perceived: $9'$

[7] The minimum viewing angle with which the middle length can be perceived is 9'.

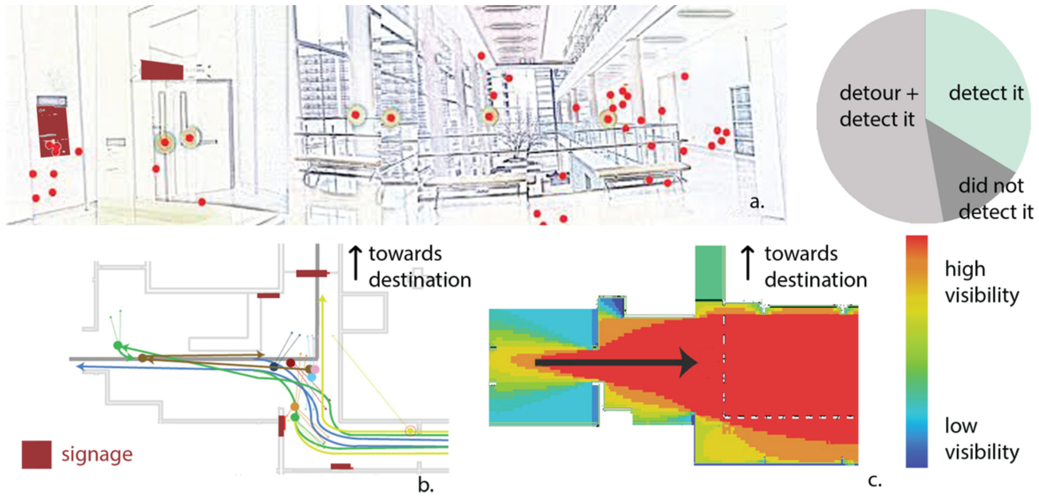
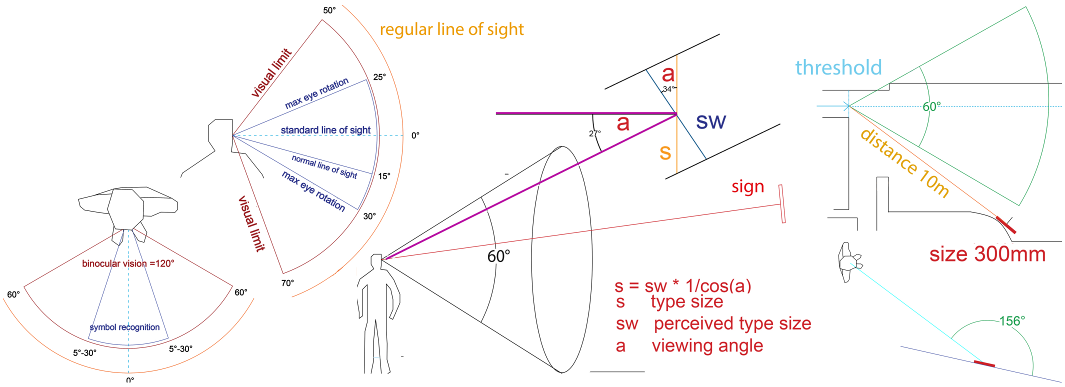


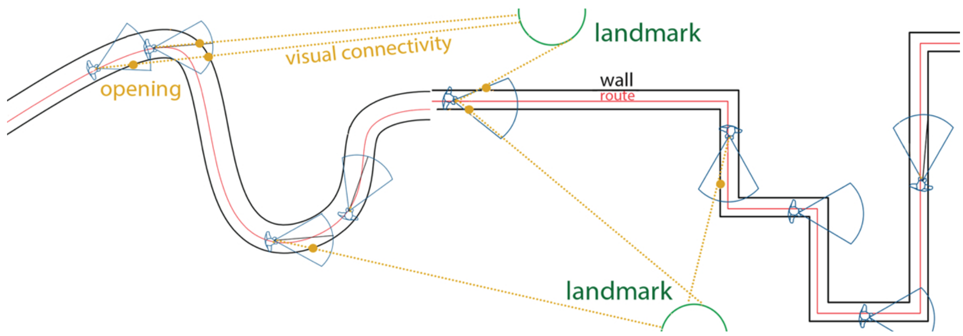
Figure 5 (a) Eye-tracking patterns of participants (on an average level) from the threshold position (decision point 4 in new Parkland hospital (Fig. 2); the route analysis (b) in combination with the visibility analysis (c) confirm that participants tend to move towards the direction with the longer line of sight from the threshold position.

RULE R2 – Ensure visual access to landmarks at the key points of the route where the probability of user's visual fixation is increased. The behavioural analysis of the case study suggests that people tend to detect the manifest cues when they are positioned vertically to the route or within a deviation of 30° towards each direction. As the variable of visual range is dependent on the route line, consequently the dynamic visual field is also shaped according to the limitations of the environmental geometry. This rule suggests that in the design process we should consider possible openings or gaps on the building's volume, based on the intersection between the physical boundaries and the dynamic visual connection between the user with the landmark (Fig. 7). In practice this will provide multiple possibilities that ensure visual connectivity with the landmarks and at the same time it will give the opportunity to the designer to choose the optimal design solution.

RULE R3 – The manifest cues should be positioned such that they are included in the anticipated fixation zone from a threshold position. Eye-tracking data analysis from our case study reveals that the average fixation zone is related to the geometry of the scene as a consequence of the spatial geometry (Fig. 8). Having as an input the three-dimensional space and the route, we are able to estimate the dimensions of the fixation zone, based on the geometrical characteristics and the user's position. For instance, from a threshold position, we draw lines towards the edges of the space that demarcate the horizontal lines of the floor and the ceiling, based on the egocentric perspective of the user. This provides the height of the fixation zone and its position on the vertical axis. Concerning the horizontal axis, the width of the zone is identical to the borders of the physical space with the exception of transparent boundaries or gaps, where we should consider a second boundary available on the scene or the arc defined by human's visual abilities (60° arc) (Fig. 8a). This fixation zone can be a useful design tool, because it can indirectly indicate where the manifest cues should be placed (in the three-dimensional space) in order to be visible by the user from a particular threshold position.



■ **Figure 6** (RULE 1) Taking into consideration human's visual abilities and the design standards for sign legibility, design should ensure that the manifest cues should be detected from the threshold positions.



■ **Figure 7** (RULE 2) The design should ensure visual access to landmarks at the key points of the route where the probability of user's visual fixation is increased. The geometry of the building, the available routes together with the dynamic visual range of the user, constitute the necessary variables in the process of defining openings' position.

RULE R4 – The main route provides in every decision point the longer line of sight. Based on our behavioural observations that confirm the results of previous empirical studies, people tend to follow the route that provides the longer line of sight in the decision points. To address the problem of people's disorientation, and detour during wayfinding, we suggest an adaptive system that prioritises the main against the secondary route and modifies the spatial geometry in every decision point according to this principle. As soon as one route is defined as the main one, this route should follow a new geometric pattern (Fig. 9). We expect that this tool could provide suggestions to the designer concerning the modification of significant decision points in a functional diagram of routes of a large-scale building.[8]

[8] A main route is defined as a path for the general public that visit a large-scale public building. It should be distinguishable from the routes used by the staff, or specialised emergency transfer routes. A main route in a hospital is considered to be a route from the atrium lobby to the restaurant, or in a airport from the entrance to the main boarding gates. In large scale public buildings, where multiple paths are involved, defining one route as "main", depends on the critical opinion of the designer.

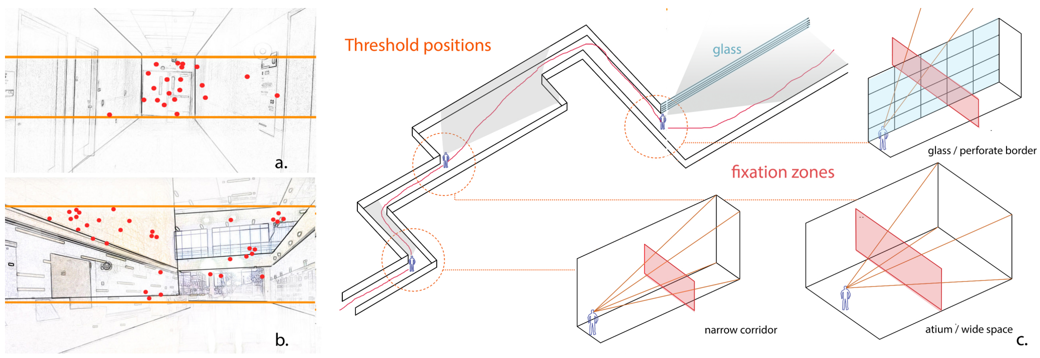


Figure 8 (RULE 3) (a–b) The dimensions of the fixation zones depend on the geometrical characteristics and the materials of the built environment, the height of the zone in a corridor and an atrium lobby differs significantly; (c) The manifest cues should be positioned such that they are included to the anticipated fixation zone from a threshold position.

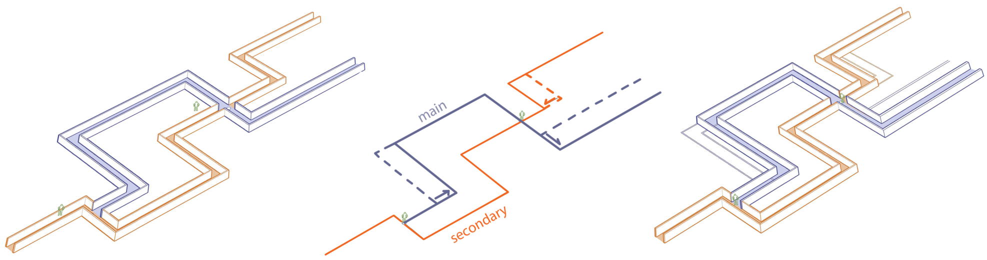


Figure 9 (RULE 4) If you define one of the route as the main, then the geometry of the space can be adjusted so that a user will experience longer view towards the main route in several decision points during his path.

5 Translating design rules to parametric design constraints

We now use our extended parametric constraint language [21] to define constraints that express evidence-based Rules R1–R4. We have implemented all rules in the constraint system FreeCAD. A two-dimensional point $p_i = (x_i, y_i)$ is defined by two real coordinates x_i, y_i. A two-dimensional line from point p_i to point p_j is denoted $[p_i, p_j]$. A vector from point p_i to p_j is $(p_j - p_i)$. An oriented point $o = (p, v)$ is a point p and a vector v. A $\mathsf{triangle}(p_1, p_2, p_3)$ is a polygonal region defined by points (vertices) p_1, p_2, p_3. Let $\theta(v_1, v_2)$ be the angle between the vectors v_1, v_2. Let $d(p_i, p_j)$ be the distance between points p_i, p_j. All distance units are in *metres*.

RULE R1: This rule requires that a sign (represented by oriented point $o_2 = (p_2, v_2)$) be placed within a certain distance and angle of a viewer ($o_1 = (p_1, v_1)$). Let p_1 be the point from which a sign must be visible (e.g. the entry point of a room), and let v_1 be a vector representing the facing direction from which the sign must be viewable. Let p_2 be the location of the sign, and let vector v_2 be the orientation of the sign (i.e. the direction that the sign is "facing").

Constraint: $\mathsf{visible\_sign}(o_1, o_2) \equiv_{\mathrm{DEF}}$

$$\begin{aligned} \theta(v_1, (p_2 - p_1)) &\leq 30°, && \text{(sign location is within user's field of view)} \\ d(p_1, p_2) &\leq 10, && \text{(sign location within viewing distance)} \\ \theta((p_1 - p_2), v_2) &\leq 10°. && \text{(sign must face viewer)} \end{aligned}$$

RULE R2: This rule restricts the location of a window opening so that a landmark can be viewed from the user path without requiring the user to turn their head beyond 30° along the direction of the path. Let p_L represent the point location of the landmark. Let p_1, p_2 be the start and end points of the user path along a corridor from which the landmark is intended to be visible. We define a point p_V that represents the last point along the path from which p_L is visible from the required viewing angle.

Constraint: $\mathsf{visible\_landmark}(p_L, p_1, p_2) \equiv_{\text{DEF}}$

$$\begin{aligned} &\exists p_V, && \text{(introduce point representing last viewing position)} \\ &p_V \in [p_1, p_2], && \text{(last viewing point lies on user's path)} \\ &\theta((p_2 - p_1), (p_L - p_V)) = 30°, && \text{(last viable line-of-sight is 30°from user's path)} \\ &p_W \in \mathsf{triangle}(p_1, p_v, p_L). && \text{(window lies within viewable region)} \end{aligned}$$

RULE R3: This rule constructs a 3D "viewing" volume that determines where signs should be placed to be noticeable. In the simplest case the viewing volume V is a polyhedron defined by six vertices based on a given oriented point (p_A, v_A) representing the observer. We construct this volume using isovists [4]. A 2D isovist is a polygon defining the set of points visible from a given point (top-down perspective). Let 3D point $p_1 = (x_1, y_1, z_1)$ be defined by horizontal coordinates x, z and vertical axis y. Consider Figure 11 with the viewing polyhedron V defined be vertices $p_1, \ldots, p_6$:

- generate the 2D isovist from a top-down perspective
- rotate v_A 90^o anticlockwise and clockwise (horizontal plane) to construct vectors v_B, v_C
- extend v_B, v_C until they hit the isovist boundary to get $(x_1, z_1), (x_2, z_2)$ (resp.)
- extend v_A to find surface w; select isovist vertices on w to define $(x_3, y_3), \ldots, (x_6, y_6)$
- the vertical position of p_1, p_2 equals the vertical position of p_A: $z_1 = z_2 = z_A$
- the vertical positions of vertices $v_3, \ldots, v_6$ are determined by the base and height of w.

A sign represented by a 3D point must remain within volume V. If the position or direction (p_A, v_A) is modified then V is reconstructed. This procedure for generating volume V also applies in more complex environments where the end surface w consists of more vertices.

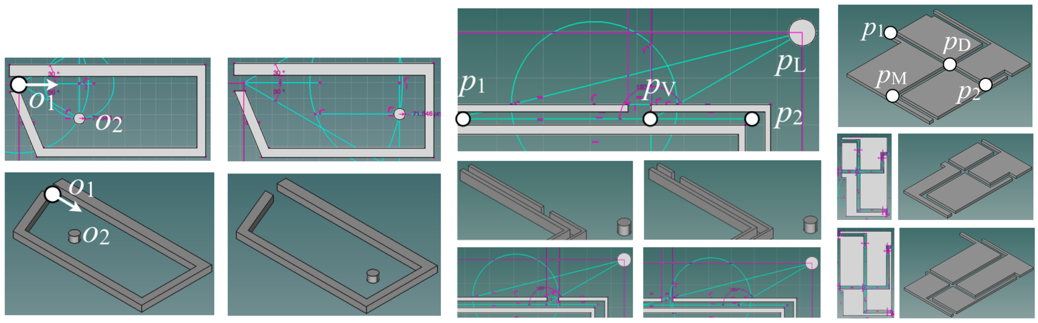
RULE R4: This rule requires that main corridors have a longer line-of-sight from a given decision point than the lower priority corridors. Let point p_D be the decision location where the user will stand, with n corridors to choose from. Let points $p_1, \ldots, p_n$ be the farthest viewable point from p_D down each corridor, i.e. the line $[p_D, p_i]$ is an unobstructed line-of-sight for each $i = 1 \ldots n$. Let M represent the main corridor, $1 \leq M \leq n$. The length of the line-of-sight from p_D to p_M must be longer than all other lines-of-sight from p_D:

Constraint: $\mathsf{priority\_corridor}(p_D, M, p_1, \ldots, p_n) \equiv_{\text{DEF}}$

$$d(p_D, p_M) > d(p_D, p_i), \text{ for } i = 1 \ldots n \wedge i \neq M.$$

6 Summary and Outlook

Parametric design shifts the designer from the position of the *author* to the position of the *coordinator*. The designer defines variables and rules to create and modify a structure through an adaptable and flexible procedure. However, a significant gap exists between the parametric tools, developed for designing human space, and the human experience inside



(a) R1: position of sign o_2 restricted by distance and orientation to viewer o_1. **(b)** R2: window to landmark p_L must be within triangle p_1, p_L, p_V. **(c)** R4: $[p_D, p_M]$ is longest corridor.

Figure 10 Constraints R1,R2,R4 implemented in the parametric system FreeCAD.

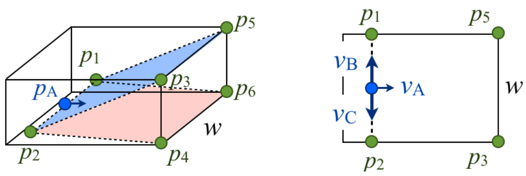


Figure 11 Constraint R3: (left) viewing volume polyhedron (vertices $p_1, \ldots, p_6$) from observer at p_A; (right) defining horizontal coordinates using isovists from a top-down perspective.

the generated design. Everyday human experiences, such as a wayfinding task in a public building, should be directly addressed in such design processes.

We propose to bridge this gap by introducing a human-centred parametric design approach coordinated by evidence of empirical studies. Parametric synthesis seeks the specification of the properties of the elements present in the encountered topology. For this reason to embed people-centred variables into parametric design systems, we establish design constraints based on human embodied visuo-locomotive experience in space. In the case of a cognitive process such as wayfinding, these constraints should be fulfilled with respect to the environmental aspects that influence the wayfinding performance (e.g. visibility, positioning of manifest cues). In this study we present examples of how to define design rules based on behavioural evidence derived by a wayfinding study conducted at the old and the new Parkland hospital in Dallas, and how to translate them into design constraints that can be utilised in parametric design modelling systems.

Acknowledgements. We acknowledge collaborations with Corgan Associates and the site management at the New Parkland Hospital(s) in Dallas, and support by Mitchell Associates (Delaware). We acknowledge the support by Tobii Pro (Sweden). We thank several student assistants, particularly Inga Lehne for technical support.

References

1 Patricia L. Alfano and George F. Michel. Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, 70(1):35–45, 1990.

2 P. Arthur and R. Passini. *Wayfinding: people, signs, and architecture*. McGraw-Hill Book Co., 1992.

3 Aysu Baskaya, Christopher Wilson, and Yusuf Ziya Özcan. Wayfinding in an unfamiliar environment. different spatial settings of two polyclinics. *Environment and Behaviour*, 36, 2004.

4 M. L. Benedikt. To take hold of space: isovists and isovist fields. *Environment and Planning B*, 6:47–65, 1979.

5 Mehul Bhatt, Carl Schultz, and Minqian Huang. The shape of empty space: Human-centred cognitive foundations in computing for spatial design. In *IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC 2012, Austria*, pages 33–40, 2012.

6 Mehul Bhatt, Carl Schultz, and Madhura Thosar. Computing narratives of cognitive user experience for building design analysis: KR for industry scale computer-aided architecture design. In *KR 2014, Vienna, Austria*. AAAI Press, 2014.

7 Mehul Bhatt, Jakob Suchan, Vasiliki Kondyli, and Carl Schultz. Embodied visuo-locomotive experience analysis: Immersive reality based summarisation of experiments in environment-behaviour studies. In *Symposium on Applied Perception (SAP 2016)*, Anaheim, USA, 2016.

8 Mehul Bhatt, Jakob Suchan, Carl P. L. Schultz, Vasiliki Kondyli, and Saurabh Goyal. Artificial intelligence for predictive and evidence based architecture design: Integrating spatial reasoning, cognitive vision, and eye-tracking for the analysis of embodied visuo-locomotive experience in the built environment. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 4349–4350. AAAI Press, 2016.

9 J. R. Carpman, M. A. Grant, and D. A. Simmons. Hospital design and wayfinding: A video simulation study. *Environment and Behavior*, 17:296–314, 1985.

10 Ann Sloan Devlin. Wayfinding in healthcare facilities: Contributions from environmental psychology. *Behavioral Sciences*, 4:423–436, 2014.

11 T. Gärling, A.Böök, and E. Lindberg. Spatial orientation and wayfinding in the designed environment: A conceptual analysis and some suggestions for postoccupancy evaluation. *Journal of Architectural and Planning Research*, 3:55–64, 1986.

12 J. J. Gibson. *The ecological approach to visual perception.* Boston: Houghton Miain, 1979.

13 OYuce Gun, R. Woodbury, B. Peters, and M. Sheikholeslami. *Elements of Parametric Design.* Routledge, 2010.

14 D. K. Hamilton and D. H. Watkins. *Evidence-Based Design for Multiple Building Types.* Wiley, 2009.

15 John M. Henderson and Andrew Hollingworth. High-level scene perception. *Annual review of psychology*, 50(1):243–271, 1999.

16 Petra Jansen-Osmann and Petra Fuchs. Wayfinding behavior and spatial knowledge of adults and children in a virtual environment: The role of landmarks. *Experimental psychology*, 53(3):171–181, 2006.

17 V. Kondyli, M. Bhatt, and T. Hartmann. Towards people-centred precedents for parametric design: The case of wayfinding in large-scale buildings. In *Lean & Computing in Construction Congress (LC3), Vol. 2 (CONVR)*, Heraklion, Greece, July 2017. (To appear.)

18 V. Kondyli, M. Bhatt, and T. Hartmann. Towards precedent based design foundations for parametric design systems. In *24th International Workshop on Intelligent Computing in Engineering*, Nottingham, July 2017. (To appear.)

19 K. Lynch. *The Image of the City.* Harvard-MIT Joint Center for Urban Studies Series. M.I.T. Press, 1960.

20 M. J. O'Neill. Effects of signage and floor plan configuration on wayfinding accuracy. *Environment and Behavior*, 23:553–574, 1991.

21 Carl Schultz, Mehul Bhatt, and André Borrmann. Bridging qualitative spatial constraints and feature-based parametric modelling: Expressing visibility and movement constraints. *Advanced Engineering Informatics*, 31:2–17, 2017. Towards a new generation of the smart built environment.

22 Hans Strasburger, Ingo Rentschler, and Martin J¸ttner. Peripheral vision and pattern recognition: A review. *Journal of Vision*, 11(5):13, 2011.

23 R. S. Ulrich, L. L. Berry, X. Quan, and J. T. Parish. A conceptual framework for the domain of evidence-based design. *Health Environ. Res. Des. J.*, 4:95–114, 2010.

24 Jerry Weisman. Evaluating architectural legibility: Way-finding and the built environment. *Environment and Behavior*, 13(2):189–20, 1981.

25 J. M. Wiener, C. Holscher, S. Buchner, and L. Konieczny. Gaze behaviour during space perception and spatial decision making. *Psychological Research*, 76:713–729, 2012.

26 A. L. Yarbus. *Eye movements and vision.* Plenum Press, New York, 1967.

Targeted Cognitive Training of Spatial Skills: Perspective Taking in Robot Teleoperation

Liel Luko[1] and Avi Parush[2]

1 **Industrial Engineering and Management, Israel Institute of Technology, Technion City, Haifa, Israel**
liluko@campus.technion.ac.il

2 **Industrial Engineering and Management, Israel Institute of Technology, Technion City, Haifa, Israel**
aparush@ie.technion.ac.il

Abstract

Spatial skills are critical for robot teleoperation. For example, in order to make a judgment of relative direction when operating a robot remotely, one must take different perspectives and make decisions based on available spatial information. Training spatial skills is thus critical for robot teleoperation, yet, current training programs focus primarily on psycho-motoric skills of the task, and less on the essential cognitive aspects of spatial skills. This work addresses this need by considering previous findings on relative direction judgments in training robot teleoperation. We developed and tested a basic training paradigm of perspective taking skill targeting the cognitive skill rather than psycho-motoric skill. An experiment tested a basic training paradigm using a stationary robot, with a training group receiving perspective taking training and a control group without training, and both tested on a transfer test with the robot. The results show that participants who went through a targeted cognitive skill training reached mastery level during the training, and performed better than the control group in an analogue transfer of learning test. Moreover, results reveal that the training facilitated participants with initial poor perspective taking skills reach the level of the high-skilled participants in transfer test performance. The study validates the possibility to target only cognitive aspects of spatial skills and result in better robot teleoperation.

1998 ACM Subject Classification I.2.9 Robotics, I.2.10 Vision and Scene Understanding

Keywords and phrases Cognitive Training, Spatial Information Processing, Targeted Training, Teleoperation, Training Spatial Skills, Visual Perspective Taking

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.12

1 Introduction

Perspective taking, an egocentric perspective transformation [36], is the ability to "transfer" mentally oneself to another location in a 3D space in order to get a different point of view of the same object [1], [16]. Such skill is required in order to judge a relative direction where one must obtain the appropriate perspective, and determine the correct direction accordingly. Studies such as [12] and [11] continue to investigate the ability to take a different perspective in a 3D space, yet, the use of the perspective taking skill manifests itself differently when used for self-navigation vs. teleoperation. The ability to judge directions in a 3D space is compromised under suboptimal conditions. For example, there may be insufficient spatial information due to a camera's single perspective in robot teleoperation or a static perspective that constraint the operator to acquire partial spatial information and make the relative direction decision relying on this partial information.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 12; pp. 12:1–12:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The present study focuses on training perspective taking skill during robot teleoperation. It is based on the theory and practice of cognitive training [28, 17, 3, 26, 6, 10, 2, 20] whereby repeated practice of cognitive skills under specific conditions will result in efficient and effective completion of tasks requiring those skills.

2 Previous Work

2.1 Spatial Skills in Teleoperation

In domains such as space teleoperation and Unmanned Ground Vehicle (UGV), effective operation requires skills such as mental rotation, visualization and perspective taking [7, 8, 9, 22, 23, 25, 34, 24]. The importance of visual information is evident in these domains of teleoperation. For example, according to [23], the three main categories which astronauts are evaluated on are: "(1) General Situation Awareness – based on the selection of appropriate camera views for the task, recognition of unexpected arm movements, and avoiding arm self-collisions. (2) Clearance – evaluated on maintaining proper clearance from structure and proper camera selection for clearance monitoring and (3) Maneuvers – evaluated on the astronaut operator's ability to make correct hand controller inputs, selecting the correct control frame for the task and planning a safe but efficient arm trajectory."

Exploring UGV teleoperation, [8, 7] used a video game/simulator to create the environment for the experiment. [8] used one monitor for the teleoperation of the mobile robot, with a single and dynamic perspective from a camera. [7] used one monitor mounted on the UGV and the other image was delivered from a unmanned aerial vehicle (UAV)

In the set of spatial skills for robot teleoperation, perspective taking is an important one. In order to specify the desired movement of the robot or the robotic arm, the operator must take the camera's perspective and use it in order to make a correct decision regarding the direction and path that the robotic arm will take. Such a skill must be trained and practiced.

2.2 Training and Transfer of Spatial Skills

The fundamental questions of training and transfer have been addressed before (for example see: [32, 4, 35]). Yet, it is unclear to what degree the training is transferred to a real task: "Estimates suggest that only 10 per cent of training expenditures transfer to the job" [13] as quoted by [15].

Different approaches to the training and transfer of skills are evident in the literature. [30, 33, 31] investigated the issue of simulation-based training in laparoscopic surgery and came to the following conclusion: "Skills acquired by simulation-based training seem to be transferable to the operative setting". Another approach was used by [14]. They concluded that training in a computer game improves the performance of pilots during real flight. Another kind of studies such as, [35], investigated the transfer of skills from one task to another. These studies used pen and pencil tests or computerized traditional spatial ability tests in order to train the subject's spatial skills.

Using targeted training ensures that the specific targeted skills are enhanced and gives insight into the core skills that need the attention and practice. For example, in their study, [19] used paper-based exercises, hands-on block construction, and two computer-based activities. Each focused specifically on the training of the spatial visualization of the student. Furthermore, [27] used a computerized training program which focused on the student's spatial skills. They concluded that targeted training had improved the spatial skills of undergraduate students as was measured by standardized spatial ability tests. Moreover, studies show that low-skilled students benefit more from spatial training than higher-skilled students as measured both with standardized tests and tasks with dominant spatial elements.

2.3 Analogic vs. Adaptive Transfer

An important distinction is made between analogic and adaptive transfer of training. In a broad sense, analogic transfer is training on one task followed by testing on an analogue task ([5]). On the other hand, adaptive transfer is training on one task followed by testing on a novel task in which there are no similarities to the training task ([29]). An example of adaptive transfer is learning moves and ball passing between players in a professional soccer game which involve geometry elements followed by taking a written geometry test. An example for an analogic transfer task is learning geometry in a class followed by being tested in the classroom on the same subject of geometry. The current study will use an analogic transfer task to evaluate transfer of acquired spatial skills.

3 Goals and Questions

The main goal of this study was to test a new paradigm for training and acquiring perspective taking skill as an essential part of robot teleoperation training. The main objective is to study perspective taking training in a stationary robot environment where people remotely operate robots and the transfer of the acquired skill to an analogical teleoperation task. The main questions are:

1. Will targeted training of perspective taking, in a robotic environment, improve performance in a spatially analogical task with a stationary robot?
2. Can we quantify and predict the improvement of performance?

In order to answer the questions presented above, a basic paradigm for perspective taking training was established and transfer of perspective taking skills in a stationary robotic environment was investigated. The following hypotheses were tested:

H1. Targeted training of perspective taking with a stationary robot will facilitate the acquisition of perspective taking skill as measured by performance during training and standardized tests.

H2. Performance in a teleoperation task of participants who receive targeted training of perspective taking will be better than participants who do not receive such training.

H3. Improvement of low-skilled participants will be greater than the improvement of high-skilled participants.

4 Method

4.1 Experiment Design

The study was a between-participant experimental design with two conditions: 1. Receiving perspective taking training and 2. No training (Figure 1). A given training trial consisted of a sequence of eight robot arm movements requiring the participant to make a decision before the subsequent movement. The training tasks requirements were to determine the relative direction of a location of a graphic element on a single plane. The direction is with respect to a figure held by the robotic arm, facing a given direction. For example: “if the figure is facing the direction of the blue star, in which direction is the yellow circle?”. An example for an answer is: “Front-Right direction with regards to the front of the figure”.

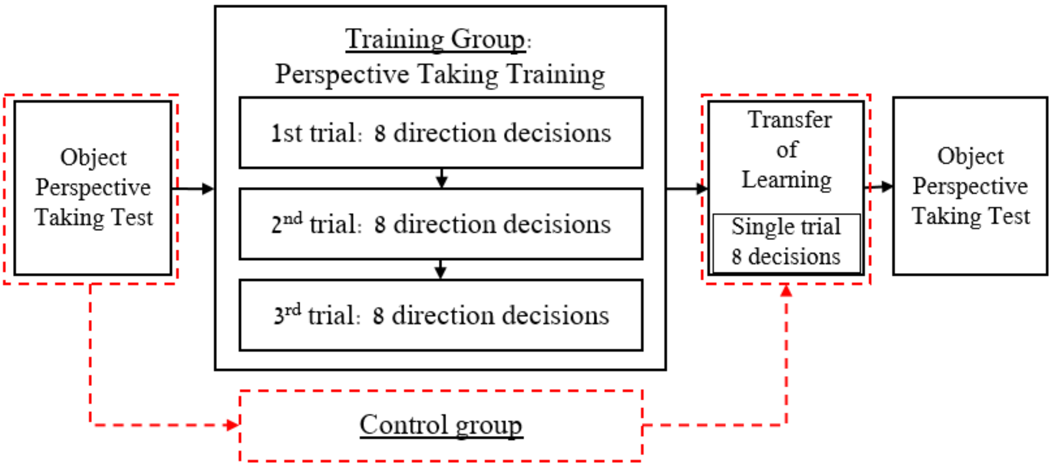


Figure 1 Experiment design: (1) Solid lines and arrows – the experiment flow of the training group. (2) Dotted red lines and arrows – the experiment flow of the control group.

Table 1 Descriptive Statistics of Participants.

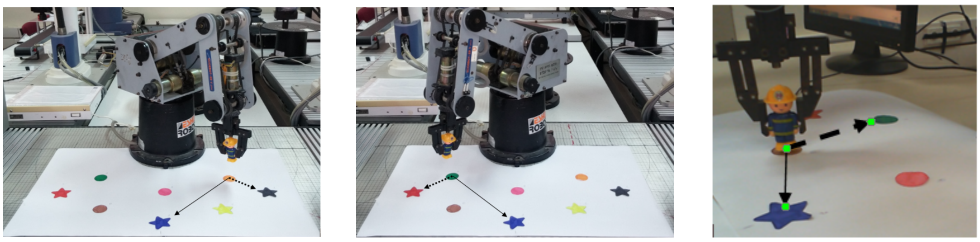
Descriptive Statistics of Participants			Number of Participants	
age			Control group	Training group
	M	SD		
Male	24.84	3.25	15	13
Female	24.83	3.21	7	8

The transfer of learning test phase (Figure 1) included a different sequence with different locations of graphic elements and relative directions from the sequence in the training phase. In addition, a pencil & paper test of perspective taking skills ([21]) of each participant in both groups was administered before the tasks to create a baseline. Participants in the training group also did the test after the training to assess if there were any changes in the skill due to training.

4.2 Participants

Forty-three participants took part in the experiment, all from STEM (science, technoloy, engineering and mathematics) fields . Although there are known gender differences in spatial skills and performance, we could not address this factor due to large differences in the sample size between male and female participants. The assignment to the experimental conditions along with age and sex parameters are presented in Table 1.

Thirty-nine of the participants were undergraduate students, and four graduate students from various faculties at the Technion. Twenty-nine participants had no previous experience with controlling robots. Fourteen either took an undergraduate course at the faculty of industrial engineering and management (Engineering of Production Systems) or had some kind of experience with robots. The experienced participants were randomly assigned to the training and the control groups. Participants received forty New Israeli Shekels for participation.



(a) The relative direction is Left.

(b) The relative direction is Right.

(c) The relative direction is Back-Left.

Figure 2 Perspective Taking training and transfer of learning objects. The robotic arm is grasping the policeman figure. Solid line – The direction the figure is facing. Dotted line – The relative direction.

4.3 Apparatus

The experiment was conducted in the Computer Integrated Manufacturing & Robotics laboratory in the Industrial Engineering and Management Faculty, Technion – Israel Institute of Technology.

The robot: An industrial stationary robot, SCORBOT ER-V plus was used for the training and test of the transfer of learning.

The program execution: For each training session, an algorithm was written using Advanced Control Language (ACL) to execute the rotating command. During the training sessions, the participant was required to use only certain keys on the keyboard in order to send desired rotation commands to the robotic arm.

Training objects: For perspective taking training, a policeman figure (Figure 2) and a map with eight graphic locations indicated using different colors were used (Figure 2).

Technological apparatus: For each training session: one desktop computer, one Microsoft HD-3000 web camera and two computer screens were used. During the tasks, participants had no direct line of sight, and received a streaming video of the working area along with the robotic arm through one camera on the right screen (Figure 3). The participants used only the right hand keys on the keyboard to insert their numeric answers.

Spatial skills standardized tests: The perspective taking ability was evaluated by the Pen & Paper Perspective Taking Test [21].

Teleoperation environment setup: A camera was placed in the front-right corner of the working surface (45 degrees deviation of the robotics' arm "Front"), 30 degrees above the working plane, capturing both the map and the robotic arm as depict in Figure 3. This specific setting allowed the flexibility to create multiple situations that require different levels of mental transformation in order to take the perspective of the robotic arm and judge a relative direction to a location of a graphic element on the map. There were no situations in which any of the colored markers were occluded.

A partition was placed between the robot and the desk with the two monitors. This allowed only the information received from the camera on site, with no direct line of sight. The proposed setting: 1. did not allow situations in which some of the graphic elements were occluded and, 2. elevated the spatial complexity of the task. Figure 3 presents the perspective taking task environment.

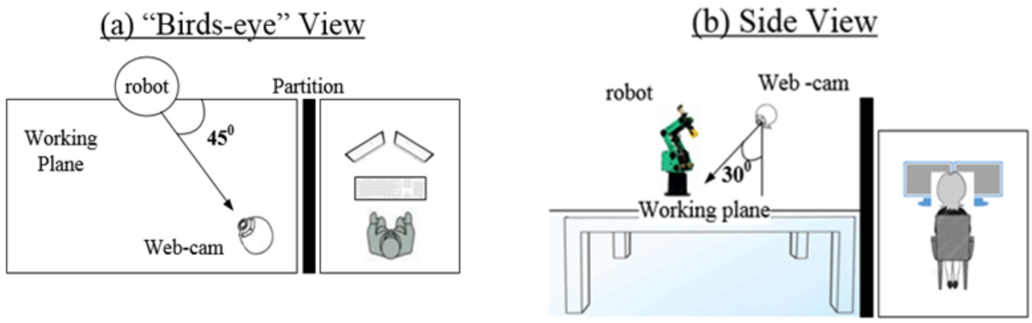


Figure 3 Perspective taking task environment. (a) "Birds-eye" View: a stand with a web-camera in the front right corner (45 degrees deviation of the front of the robotic arm). (b) Side View: a stand with a web-camera located 30 degrees above the working plane, capturing both the map and the robotic arm in the camera's perspective.

4.4 Measurements

Performance of perspective taking skill was measured by the following:

- Mean Time to Decision (TTD) – time from "ENTER" keying until a correct decision is made.
- Mean Number of Mistakes – Mean number of mistakes for each of the trials and each of the decisions in every trial. (Any incorrect direction to target location of a graphic element was considered a mistake)
- Perspective Taking Test Score – The sum of correct answers on the Perspective Taking Test.

4.5 Procedure

First, each participant signed an informed consent form and was administrated the object perspective taking test to establish a baseline. Next, the participant received a demonstration of the operation of the robotic arm and instructions as to how to operate it using the keyboard.

After receiving the instructions, the participant sat behind a desk with two monitors. On the left, the monitor displayed the instructions during training, and on the right a display of the robotic arm and the working area as it seen from the camera's perspective. Next, the training group participants engaged in a training session, according to the original assignment, consisting of three trials. Each participant took a test with general knowledge questions at the end of each trial in order to clear the working memory and minimize the probability of memorizing the answers.

After the training session, the training group participant engaged in an analogical task to test the transfer of learning followed by a spatial ability test. The control group received no training and engaged in the same analogical task.

Perspective Taking Training: Each trial began with the robotic arm holding a figure in a given direction above a specific location of a graphic element on the map. The participant set behind a desk with two monitors and a keyboard. Once the participant started the session, a question would appear on the left screen, asking for a direction to another location of a graphic element on the map (Figure 2). The participant was required to determine the relative direction of the location of the graphic element with respect to the figure held by the

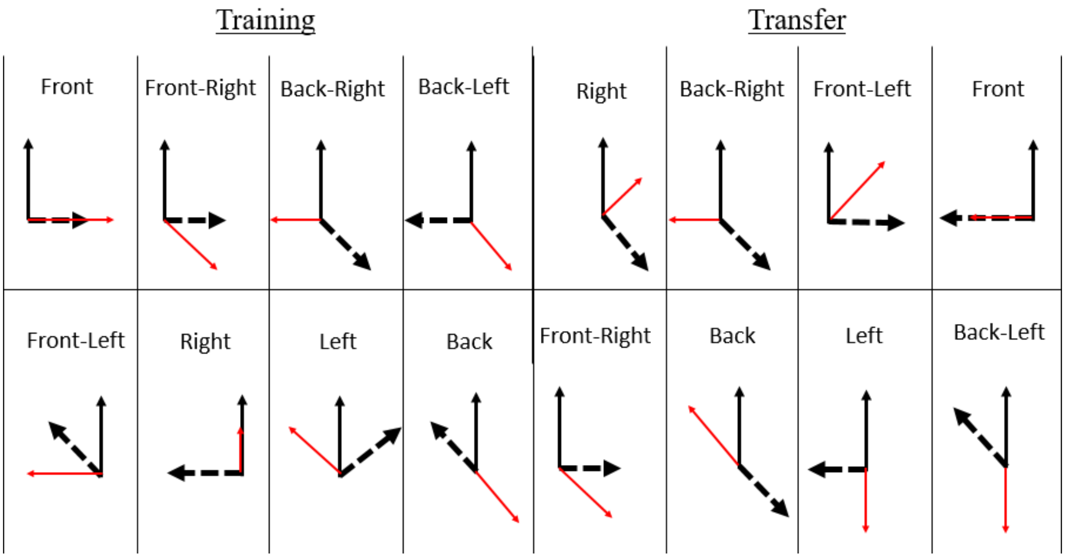


Figure 4 Judgment of relative direction: the solid black up arrows represents the camera's perspective. The dotted black arrows represent the direction of the figure. The thin red arrows represent the relative direction to the target location of a graphic element (relative to the figure's perspective).

robotic arm. The participant entered a code direction (for example, Front=482) and pressed the "ENTER" key on the keyboard. Figure 4 presents the perspective of the participant as given by the camera, the imagined heading (dotted line-the direction of the figure in the gripper) and the relative direction of the location of the graphic element to the imagined heading (red line).

Test of Transfer of Learning: The analogical transfer test resembled the training task and included one trial with eight decisions. In the perspective taking task, the order of the required directions were changed and so were the starting and target locations of the figure held by the robotic arm.

5 Results

5.1 Training Perspective Taking Skill

Twelve participants of the training group who completed all tasks successfully were included in the final analysis of the learning phase. Two of the participants had at least one unsuccessful attempt in at least one trial (9.5%). Additional seven participants (36%) were excluded (out of the successful) due to TTD greater than three standard deviations above the mean of the TTD.

Time To Decision (TTD). Analysis of variance showed there was a significant trial effect, $F(1.23, 13.57) = 31.51, p < 0.01, \eta = 0.741$ and observed power of 1, indicating that the mean TTD was significantly different in the three trials. Overall, an improvement in performance as measured by TTD is evident in Figure 5.

There was a significant relative direction effect, $F(7, 77) = 17.67, p < 0.01, \eta = 0.616$ and observed power of 1, indicating that the TTD differs significantly between the relative

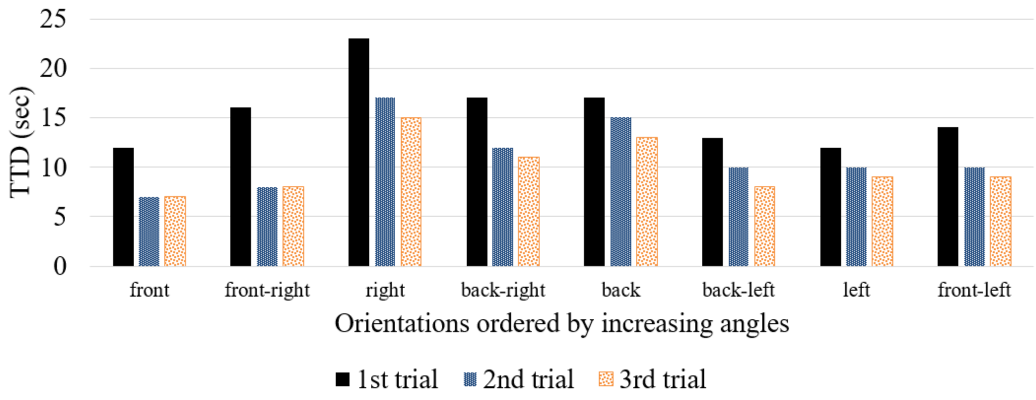


Figure 5 Practice effect of each relative direction. Presented with rotations ordered by increasing angles.

directions (Figure 5). The interaction effect between trial and relative direction was not significant, $F(14, 154) = 1.044, p = 0.4$.

Mean Number of Mistakes. There was a significant trial effect, $F(2, 22) = 4.827, p = 0.018, \eta = 0.305$ and observed power of 0.74, indicating that the mean number of mistakes was significantly lower in each subsequent trial. There was a significant relative direction effect, $F(2.369, 26.064) = 4.822, p = 0.013, \eta = 0.305$ and observed power of 0.795, indicating that the mean number of mistakes was significantly different throughout the different directions. The interaction between trial and direction was not significant: $F(14, 154) = 0.702, p = 0.43$.

5.2 Test of Transfer of Learning

Performance was measured by TTD and Number of Mistakes of successful trials. Due to the nature of the repeated measure design, only seventeen participants from the training group and seventeen from the control group were included in the final analysis. In the training group, four participants (19%) were excluded due to TTD greater than three standard deviations above the mean of the TTD in the relative direction performance. In the control group, one participant (4.5%) was excluded because of one unsuccessful attempt. Four participants (19%) were excluded due to TTD greater than three standard deviations above the mean of the TTD in the relative direction performance. Nothing particular was observed in the skill test scores of the excluded participants. Descriptive statistics is presented in the following results.

Time To Decision (TTD). A two-way analysis of variance with repeated measures and group as a between subject effect resulted in a significant group effect, ($F(1, 32) = 8.012, p = 0.008, \eta = 0.2$ and observed power of 0.784), indicating that the TTD of the training group (M=14, SD=7) is significantly better than the control group TTD (M=27, SD=33) (Figure 6). The relative direction effect was significant, $F(2.22, 71.16) = 6.057, p = 0.003, \eta = 0.159$ and observed power of 0.896. The interaction effect between group and relative direction was not significant.

Mean Number of Mistakes. The group effect was not significant with SL of 0.05, $F(1, 37) = 0.249, p = 0.621$. The relative direction effect was significant: $F(7, 259) = 3.864, p =$

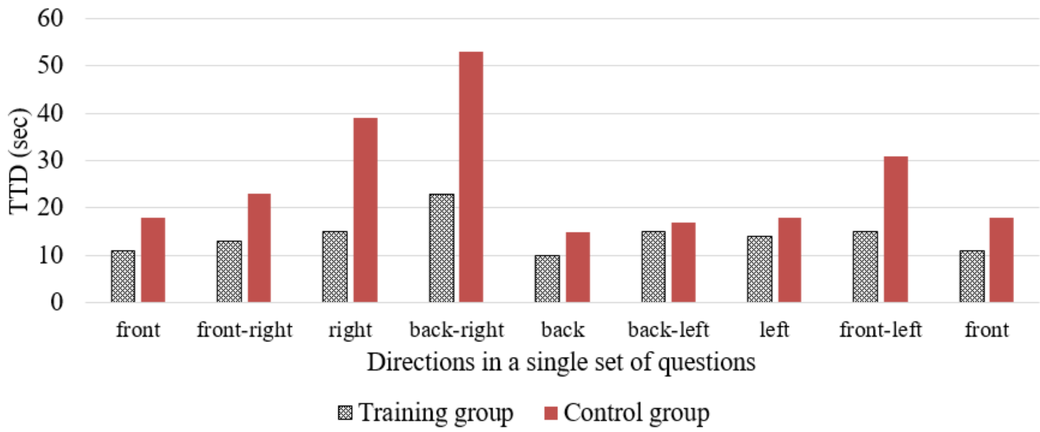


Figure 6 Perspective taking transfer of learning test performance. Presented with directions ordered by increasing angles.

$0.001, \eta = 0.095$ and observed power of 0.981. The interaction effect between group and relative direction was not significant in S.L of 0.05.

Descriptive statistics of the excluded participants from the training analysis reveals that four of the seven participant succeed in the transfer task with mean TTD of 15.4 seconds (similar to the mean of the training group).

5.3 Perspective Taking Standardized Test

Twenty-one participants from the control group were included in the test analysis, one participant did not follow the instructions and therefore was excluded from the analysis.

Control Group: A linear regression was performed to test the relationship between required spatial skills in the test of transfer of learning and standardized tests. Results show that pre-task perspective taking test score significantly predicted the single set of robotic task performance in the control group, where no confounding variables are present, $\beta = -0.756, t(19) = -4.9, p < 0.01$, with $R^2 = 0.572, F(1, 19) = 24.013, p < 0.01$.

Training Group: A linear regression was performed to test the relationship between required spatial skills in the training task and standardized tests. Results show that pre-training perspective taking test score significantly predicted the first set of training task performance in the training group, $\beta = -0.633, t(19) = -3.563, p = 0.002$, with $R^2 = 0.401, F(1, 19) = 12.693, p = 0.002$.

Multiple regression analysis was used to test if the standardized perspective taking test score and training performance significantly predicted participants' performance on the transfer task as measured by TTD. The results of the regression indicated the two predictors explained 83% of the variance ($R^2 = .83$, $F(3, 30) = 48.88$, $p < 0.000$). It was found that training performance significantly predicted transfer performance (B = -44.818, $t(1) = -4.452$, $p < .000$), as did standardized test score (B = -4.503, $t(33) = -10.543$, $p < .000$), and their interaction (B = 4.203, $t(33) = 4.182$, $p < .000$). Participants' predicted transfer task performance is equal to 62.354-4.5(spatial skill level)-44.818(training)+4.203(level of skill*training) where training is coded as 1=trained, 0=control, and level of skill measured by standardized test score. In the control group, participant's TTD decreased 4.5 seconds for each point in the standardized test. In the training group, the more skilled the participant, the less improvement in TTD is evident. On the other hand, for less skilled participants, the improvement was greater. Descriptive statistics of the excluded participants from the training analysis reveals that four of the seven participant who succeeded in the transfer task had also improved their initial perspective taking skill level as measured by standardized test score from 8 to 10.75.

6 Summary

Recapping the findings, the effectiveness of the targeted cognitive training of the perspective taking spatial skill was particularly evident for participants with poorer initial perspective taking skills. The training helped them reach the performance level of participants with high initial perspective taking skills.

The Time To Decision (TTD) measure of the training group suggests that the most difficult directions to judge during the training sessions were Back-Right and Front-Left. Nevertheless, even these directions kept improving during the third trial. This may imply that one more session would have reduced all differences between directions. From the TTD performance of the training group during the transfer of learning test, it is evident that there were no substantial differences between TTD of the relative directions, including the more difficult directions from the training phase, other than the Back-Right direction. This implies that a learning process took place during training, which had leveled all differences between TTD of different directions. The difference between the training and the control group in the transfer of learning test was significant, indicating the training was effective and improved the performance of the training group.

Results of both the training and control group, in the transfer of learning test, resemble previous findings of relative direction judgment [21, 18]. Back and front directions are relatively easy, and directions with angles greater than 90 degrees are more difficult to judge. However, in the present work, not all relative directions follow the performance pattern found in previous literature ([18]). An example of an exception of that rule is Back-Left, which was relatively easy to judge as opposed to the literature ([18]) that suggests that Back-directions should be the hardest to judge. A possible explanation is presented in the discussion section.

7 Discussion and Theoretical Implications

A model of information processing can be adapted to teleoperation tasks with a focus on the cognitive spatial aspects. Figure 7 presents a model, which is composed of links that were empirically studied here (Solid lines) and hypothetical links (Dotted lines).

The model depicts the information-processing-action flow while teleoperating a robot, and the influence of initial spatial skills and training of spatial skills. The flow starts from the point where the remote robot position or movement (at the left of the diagram in Figure 7) is perceived through a technological aid, such as a camera. The flow ends with the control of the remote robot's movements through a technological aid, such as a remote control (at the right of the diagram in Figure 7).

The model is based on the premise that spatial skill level influence the process of acquisition of spatial skills; Initial spatial skills influence the acquisition of spatial information, the spatial cognitive processes such as perspective taking, and the decision how to proceed with the robot operation. Specifically, lower-skilled participants will benefit more and are more effected by the process of training. The model also suggests that training spatial skills such

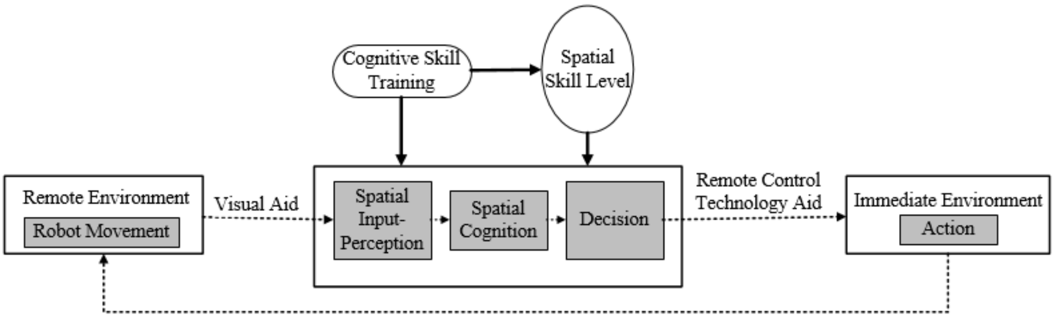


Figure 7 Cognitive task flow. The technology aid tool that effects the perception can be a visual aid or other feedback from the environment. The technological aid tool that effects the relationship between decision and action is a remote control such as joystick or a keyboard.

as perspective taking will influence those cognitive processes, but can also facilitate the acquisition of further spatial skills. Finally, the model suggests that the technological aids, such as the camera, either at the perception stage or at the action stage, can also influence the cognitive processes.

The pattern of reaction times is different from previous findings on relative direction judgments. An example is the TTD of left directions with respect to right directions. In the current setting, the camera's perspective was fixed to the left of the participant's perspective. This implies that relative direction judgments might be influenced by external technological aids such as camera's perspective during teleoperation, which attenuates the available spatial information. Specifically, it seems that the teleoperation environment: specific perspective during the task, limited visibility and the usage of egocentric frame of reference, might have had an effect on the ability of the operator to judge directions in space. The notion of the impact of limited visual information attained through technologic aids on performance is consistent with current results found in literature.

The findings here suggest that training perspective taking skill using the proposed paradigm had different benefits for different initial perspective taking skills as the model implies. In light of these results, we propose revisiting the approach to training and acquisition of spatial skills, both on the theoretical and practical levels. Specifically, future studies, should explore the theoretical aspect of teleoperation performance in terms of: 1. the cognitive processes that underlie training and acquisition of spatial skills; and, 2. the technological factors present in teleoperation that may moderate our ability to perceive, analyze, and execute spatial strategies.

The effectiveness of the paradigm should be explored further with regards to its length, for example, a different design to test a single training trial and its effect on perspective taking skills acquisition. Moreover, due to unsuccessful trials of participants, the sample size was smaller than predicted, this had an influence on the effect size and observed power. Additional studies should also consider the effect of the technological aids on the process of skill acquisition and the transfer of learning in various teleoperation settings. Specifically, exploring the process of perception, analysis and decision during robot teleoperation with various control methods and visual aids systems. For example, teleoperation using virtual reality with a head mounted display. In such 3D environment, the process of spatial perception, analysis and decision may be effect by the issue of telepresence and present different results. The effect of technology aids on performance in spatial tasks should be investigated, and the suggested model would help future studies generate and explore hypotheses regarding the acquisition of spatial skills.

Acknowledgments. This study was partially supported by the Bernard M. Gordon Center for Systems Engineering at the Technion.

References

1. Edith Ackermann. Perspective-taking and object construction: Two keys to learning. *Constructionism in practice: designing, thinking, and learning in a digital world, Lawrence Erlbaum, Mahwah, NJ*, pages 25–35, 1996.
2. Michel Jose Anzanello and Flavio Sanson Fogliatto. Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*, 41(5):573–583, 2011.
3. Yohanan Arzi and Avraham Shtub. Learning and forgetting in mental and mechanical tasks: a comparative study. *IIE transactions*, 29(9):759–768, 1997.
4. Maryann Baenninger and Nora Newcombe. The role of experience in spatial test performance: A meta-analysis. *Sex roles*, 20(5-6):327–344, 1989.
5. Susan M. Barnett and Stephen J. Ceci. When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, 128(4):612, 2002.
6. Joseph Bukchin, Ruth Luquer, and Avraham Shtub. Learning in tele-operations. *Iie Transactions*, 34(3):245–252, 2002.
7. Jessie Y. C. Chen. Uav-guided navigation for ground robot tele-operation in a military reconnaissance environment. *Ergonomics*, 53(8):940–950, 2010.
8. Jessie Y. C. Chen and Carla T. Joyner. Concurrent performance of gunner's and robotics operator's tasks in a multitasking environment. *Military Psychology*, 21(1):98, 2009.
9. C. Christopher et al. Spatial perception and robot operation: Selection, training, gaming, and design. *Naval Engineers Journal*, 124(1):131–140, 2012.
10. L. Davidovitch, A. Parush, and A. Shtub. Simulation-based learning in engineering education: Performance and transfer in learning project management. *Journal of Engineering Education*, 95(4):289–299, 2006.
11. Pio Alfredo Di Tore. Perception of space, empathy and cognitive processes: Design of a video game for the measurement of perspective taking skills. *iJET*, 9(7):23–29, 2014.
12. Andrea Frick, Wenke Möhring, and Nora S. Newcombe. Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds. *Frontiers in psychology*, 5:386, 2014.
13. David L. Georgenson. The problem of transfer calls for partnership. *Training & Development Journal*, 1982.
14. Daniel Gopher, Maya Well, and Tal Bareket. Transfer of skill from a computer game trainer to flight. *Human Factors*, 36(3):387–405, 1994.
15. Rebecca Grossman and Eduardo Salas. The transfer of training: what really matters. *International Journal of Training and Development*, 15(2):103–120, 2011.
16. Mary Hegarty and David Waller. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2):175–191, 2004.
17. Darlene Hewitt, Kelli Sprague, Robert Yearout, Donald Lisnerski, and Catherine Sparks. The effects of unequal relearning rates on estimating forgetting parameters associated with performance curves. *International Journal of Industrial Ergonomics*, 10(3):217–224, 1992.
18. Douglas L. Hintzman, Carla S. O'Dell, and David R. Arndt. Orientation in cognitive maps. *Cognitive Psychology*, 13(2):149–206, 1981.
19. Sherry Hsi, Marcia C Linn, and John E Bell. The role of spatial reasoning in engineering and the design of spatial instruction. *journal of engineering education-washington-*, 86:151–158, 1997.
20. Mohamad Y. Jaber and Christoph H. Glock. A learning curve for tasks with cognitive and motor elements. *Computers & Industrial Engineering*, 64(3):866–871, 2013.

21 Maria Kozhevnikov and Mary Hegarty. A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5):745–756, 2001.

22 Philip Lamb and Dean Owen. Human performance in space telerobotic manipulation. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 31–37. ACM, 2005.

23 Andrew M. Liu, Charles M. Oman, Raquel Galvan, and Alan Natapoff. Predicting space telerobotic operator training performance from human spatial ability assessment. *Acta Astronautica*, 92(1):38–47, 2013.

24 Lindsay O. Long, Joshua A. Gomer, Jessica T. Wong, and Christopher C. Pagano. Visual spatial abilities in uninhabited ground vehicle task performance during teleoperation and direct line of sight. *Presence: Teleoperators and Virtual Environments*, 20(5):466–479, 2011.

25 M. Alejandra Menchaca-Brandan, Andrew M. Liu, Charles M. Oman, and Alan Natapoff. Influence of perspective-taking and mental rotation abilities in space teleoperation. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*, pages 271–278. IEEE, 2007.

26 David A. Nembhard and Mustafa V. Uzumeri. Experiential learning and forgetting for manual and cognitive tasks. *International journal of industrial ergonomics*, 25(4):315–326, 2000.

27 Richard M. Onyancha, Matthew Derov, and Brad L. Kinsey. Improvements in spatial ability as a result of targeted training and computer-aided design software use: Analyses of object geometries and rotation types. *Journal of Engineering Education*, 98(2):157, 2009.

28 Catherine Sparks and Robert Yearout. The impacts of visual display units used for highly cognitive tasks on learning curve models. *Computers & Industrial Engineering*, 19(1-4):351–355, 1990.

29 Harold D. Stolovitch and Achi Yapi. Use of case study method to increase near and far transfer of learning. *Performance improvement quarterly*, 10(2):64–82, 1997.

30 Debi Stransky, Laurie M. Wilcox, and Adam Dubrowski. Mental rotation: cross-task training and generalization. *Journal of Experimental Psychology: Applied*, 16(4):349, 2010.

31 Lana P. Sturm, John A. Windsor, Peter H. Cosman, Patrick Cregan, Peter J. Hewett, and Guy J. Maddern. A systematic review of skills transfer after surgical simulation training. *Annals of Surgery*, 248(2):166–179, 2008.

32 Melissa S. Terlecki, Nora S. Newcombe, and Michelle Little. Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied cognitive psychology*, 22(7):996–1013, 2008.

33 J. Torkington, S. G. T. Smith, B. I. Rees, and A. Darzi. Skill transfer from virtual reality to a real laparoscopic task. *Surgical endoscopy*, 15(10):1076–1079, 2001.

34 Markus Wilde, Jan Harder, and Ralf Purschke. Operator learning effects in teleoperated rendezvous & docking. In *Aerospace Conference, 2013 IEEE*, pages 1–11. IEEE, 2013.

35 Rebecca Wright, William L. Thompson, Giorgio Ganis, Nora S. Newcombe, and Stephen M. Kosslyn. Training generalized spatial skills. *Psychonomic Bulletin & Review*, 15(4):763–771, 2008.

36 Jeffrey M. Zacks, Jon Mires, Barbara Tversky, and Eliot Hazeltine. Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, 2(4):315–332, 2000.

What Makes the Difference When Learning Spatial Information Using Language? The Contribution of Visuo-Spatial Individual Factors

Chiara Meneghetti[1] **and Veronica Muffato**[2]

1 **Department of General Psychology, University of Padova, Padova, Italy**
chiara.meneghetti@unipd.it
2 **Department of General Psychology, University of Padova, Padova, Italy**

Abstract

Within the spatial cognition domain, increasing interest is being paid to identifying the factors able to support good-quality environment learning. The present study examined the role of several individual visuo-spatial factors in supporting representations derived from spatial language, using descriptions. A group of undergraduates performed visuo-spatial and verbal cognitive tasks and completed visuo-spatial questionnaires, then listened to descriptions of fictitious large-scale environments presented from survey (map-based) and route (person-based) views, and to non-spatial descriptions for control purposes. Their recall was assessed using a verification test and a graphical representation task. The results showed that: (i) verbal abilities support accuracy in recall tasks of spatial and non-spatial descriptions; (ii) visuo-spatial abilities, preferences (such as pleasure in exploring), and visuo-spatial strategies specifically support accuracy in recall tasks of spatial descriptions. The contribution of individual visuo-spatial factors varies, however, as a function of the type of description and the type of recall task: preference for the survey strategy seems more associated with performance in survey description recall and graphical representation. The results are discussed in the light of spatial learning models and in terms of their implications.

1998 ACM Subject Classification J.4 Social and Behavioral Sciences

Keywords and phrases Spatial language, survey description, route description, visuo-spatial abilities, self-reported visuo-spatial factors

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.13

1 Introduction

Spatial information can be acquired directly from sensorimotor experience, or indirectly from maps or virtual displays ([14, 26] for a review), or from spatial descriptions ([12] for a review). The last of these is commonly used, and involves reading from a device or hearing from a speaker the description of a path, or of the location of a landmark in an environment. The use of language to convey spatial information is attracting increasing attention in disciplines that deal with spatial information, such as engineering and geography. The interest lies in devising systems capable of handling spatial language in order to transfer knowledge of a route indications from a user to a robotic system [32], for instance, or systems capable of deriving a sketch-map from a speaker's spatial instructions [19]. Psychology studies such as ours can suggest ways for other disciplines to approach the spatial language issue. It has been clearly demonstrated that the processing of a verbally-conveyed spatial description leads to the formation of a mental model, i.e. an abstraction that resembles the structure of the corresponding state of affairs in the outside world [18], in which spatial relations between objects (landmarks) are mentally represented [2, 38, 13]. Mental models derived from the

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 13; pp. 13:1–13:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

processing of spatial descriptions have been shown to have spatial features, though they may not perfectly resemble the mental representations acquired from visual input [31, 33]. A relevant question in this research domain concerns how to identify which individual features support a person's ability to produce mental representations. Among several others, individual visuo-spatial factors can have a major role, especially in predicting environment learning performance. It has been demonstrated [1, 14] that individual visuo-spatial factors (typically tested with paper and pencil tasks) represent small-scale abilities; the latter predict the ability to move in and represent the environment, which is an expression of large-scale abilities [35, 40, 41]. When individual verbal abilities were examined, on the other hand, they did not predict environment learning performance [14, 40]. Examining individual visuo-spatial factors (small-scale abilities) therefore enables us to predict environment learning accuracy (a large-scale ability), and this represent a relevant research question in the spatial cognition domain.

There are several aspects to take into account when considering the literature on how people mentally represent verbally-conveyed spatial information. For a start, there is the type of individual visuo-spatial factor, i.e. the various competences, including both cognitive abilities and self-reported preferences and strategies. Then there is the modality used to convey spatial information, i.e. from a route or survey perspective [38]. Route descriptions present landmarks and their relative positions from an egocentric perspective (or path view) and use an intrinsic frame of reference (e.g. "to your left", "behind you"). Survey descriptions present them from an allocentric perspective (or bird's-eye view) and use an extrinsic frame of reference, such as compass points (north, south, east, west). The literature review presented in the following paragraphs illustrates findings on spatial description learning considering: the type of visuo-spatial factor examined (objectively-tested cognitive abilities vs. self-reported attitudes and behaviors) in relation to type of description considered (survey vs. route).

Visuo-spatial abilities and spatial descriptions. Visuo-spatial abilities are needed to generate, retain and transform abstract visual images [20]. They comprise distinct aspects [16, 39], such as mental rotation, which is the ability to mentally rotate an object or oneself when imagining different views of a set of objects [15]. Another aspect responsible for individual differences concerns working memory, and particularly visuo-spatial working memory (VSWM), which is needed to process and retain visuo-spatial information. VSWM is generally tested on the recall of increasingly long series of elements, as in the Corsi blocks task [6]. Studies – mostly considering route descriptions – have shown that both mental rotation and VSWM abilities support the accuracy of mental representations derived from spatial descriptions [34, 25, 22, 23]. When people's recall of survey and route descriptions is compared, their final representations may differ [36], and this may at least partly relate to the cognitive abilities required. Learning a route description demands more VSWM resources than learning a survey description [2, 30, 10]. It is noteworthy, however, that the involvement of cognitive abilities also differs in relation to the type of recall task: performing graphical recall tasks after listening to a spatial description (e.g., asking participants to reproduce a map of the environment described) is more demanding on an individual's visuo-spatial cognitive resources than performing verbal tasks (e.g., answering questions about spatial relations) [22].

Self-reported visuo-spatial factors and spatial descriptions. By self-reported visuo-spatial factors, we mean a number of preferences, attitudes and strategies used when dealing with spatial information. People's visuo-spatial preferences consist in their inclination to orient themselves in an environment based on a mental map (survey/allocentric view) or from a personal view (route/egocentric view). These preferences influence their spatial description recall [23, 29]. Differences between survey and route description recall emerge in relation to the type of task used to test what a participant remembers. For instance, individuals with a stronger preference for the survey view performed better in a map drawing task after learning from a survey description [29]. Accuracy in performing spatial recall tasks is also influenced by self-reported strategy use, i.e. the type of procedure adopted to deal with certain recall demands [4, 11]. Concerning spatial descriptions, individuals report using more visuo-spatial strategies, mentally visualizing a path (route strategy) or forming a mental map (survey strategy), than verbal strategies based on repetition [23]. Comparisons between different types of description found survey description learning more associated with the use of survey strategies, while route description learning was associated with the use of both survey and route strategies [24].

The above-cited findings demonstrate the important influence of spatial (mental rotation) ability, VSWM, and self-reported (survey and route) strategy use on people's approach to spatial information, and their different modulatory effects as a function of the perspective learnt and the recall task performed. It should be noted, however, that the individual visuo-spatial factors were, in most cases, taken into account separately, and route descriptions were usually considered. Indeed, few studies examined the simultaneous role of several visuo-spatial factors in spatial description learning, and showed that both mental rotation and VSWM abilities, together with self-reported preferences and visuo-spatial strategies, play a part in supporting the recall of spatial (route) descriptions [25].

Visuo-spatial and verbal factors in spatial descriptions. When spatial information is conveyed verbally, people's verbal abilities naturally have a role too. In fact, when verbal working memory (VWM), i.e. the ability to process and maintain verbal information, was analyzed, it was found involved in the processing both of non-spatial and spatial (route) descriptions (though the latter specifically involved VSWM too) [30]. Reading comprehension, i.e. the ability to identify the meaning of a text, was also found to support performance in the recall of both non-spatial and spatial (route) descriptions, and the latter was additionally sustained by people's visuo-spatial abilities. This indicates that processing spatial descriptions requires the involvement of different verbal and visuo-spatial cognitive abilities, depending on the descriptions' format and type of content [12]. While the contribution of verbal abilities, such as VWM and reading comprehension, to the formation of a mental model has been demonstrated [7, 42], we do not know for sure how different visuo-spatial competences (both cognitive abilities and self-reported strategies) work in supporting the learning of descriptions with a spatial content, from a survey or route perspective, and how they emerge in different recall measures.

The novel aim of the present study was therefore to explore the role of visuo-spatial factors (in term of both cognitive ability and self-reported strategies) in supporting the learning of spatial descriptions, and the possibly different modulation effects of visuo-spatial factors as a function of the perspective learnt and the type of recall task administered. Given that gender is a source of variability in spatial task performance, and in spatial description learning [23], a large group consisting entirely of females was selected to participate in this study in order to avoid any confounding influence of gender. Participants were first assessed on their individual small-scale abilities by means of visuo-spatial tasks (testing their mental rotation and VSWM abilities), and verbal tasks (testing their reading comprehension and VWM abilities), and they completed a number of visuo-spatial questionnaires assessing their preferences in approaching the environment and pleasure in exploring (given the evidence of this positively influencing spatial learning [27]. Then they were assessed on their ability to represent spatial information by means of spatial descriptions: they listened to descriptions of fictitious large-scale environments in survey and route views, and to non-spatial descriptions for control purposes. The effect of perspective relates to the type of recall task administered [33, 12], so spatial recall was assessed using tasks both in a verbal format, by asking participants to judge the truthfulness of some relations (i.e. a verification test), and in a visuo-spatial format, by asking them to reproduce the arrangement of landmarks in a layout (i.e. a graphical representation).

We explored the different modulation of a set of individual visuo-spatial differences in environment representation. In particular, we expected accuracy in the recall of all descriptions to be supported by verbal abilities (as suggested in [7] due to the verbal format of the input used. After controlling for verbal abilities, we expected visuo-spatial cognitive abilities to specifically support spatial description learning (as suggested in [2, 25, 5]. Individual visuo-spatial preferences and strategies should also support the learning of spatial descriptions [2, 25]. Their contribution could differ as a function of the perspective learnt and/or the type of recall task administered. In particular, we expected the contribution of visuo-spatial factors to be stronger for active recall tasks (i.e. graphical representation) than in the recognition of the truthfulness of spatial relations (i.e. verification test) [21]. We also examined whether the effect of perspective related to the strategy used, such as the use of a survey strategy to memorize a survey description [24], or to complete a map-view task [25, 23].

2 Method

2.1 Participants

The study involved 173 female undergraduates (M age = 20.99, SD = 3.73), all native Italian speakers, in exchange for course credits. The study was approved by the local ethical committee for psychology studies.

2.2 Materials and procedure

Participants were tested individually in two sessions lasting an hour each. In the first session, they completed the verbal and visuo-spatial individual difference measures in a balanced order. The tasks and questionnaires are described below.

2.2.1 Individual differences in verbal and visuo-spatial measures

Verbal/Visuo-Spatial Working Memory tasks. The Backward digit span task [8] and backward Corsi blocks task [6] involve repeating in reverse order increasingly long sequences of numbers and blocks, respectively (from 2 to 9), that are presented by the experimenter. The final score is the longest correctly-repeated sequence.

Reading Comprehension Task (RCT [5]). The task consists in reading an argumentative text “the Rio conference” about climate change and pollution, and answering 10 multiple-choice questions on its content (maximum score: 10).

Perspective-Taking Task (PTT [9], adapted from [15]. The task consists in looking at a picture showing a configuration of 7 objects (on a piece of paper) and having to imagine standing at one object, facing towards another, and pointing in the direction of a third (always misaligned with respect to the respondent's view). The answer is given by drawing an arrow from the center towards the perimeter of a circle drawn on the paper, below the configuration of objects. The answer is scored in terms of absolute degrees of error (six items; time limit: 5 minutes).

Sense of Direction and Spatial Representation questionnaire (SDSR [28]). This comprises 11 items measuring 3 factors: (i) Sense of Direction – preference for survey mode (e.g., "Do you think you have a good sense of direction?"), 4 items; (ii) knowledge and use of cardinal points (e.g., "When you are outside, do you naturally identify cardinal directions, i.e., which way is North, South, East and West?"), 3 items; and (iii) preference for landmark and route mode (e.g., "Think about how you orient yourself in different surroundings. Would you describe yourself as a person who orients him/herself by remembering routes?"), 4 items.

Attitudes to Orientation Tasks scale (AtOT [9]). This comprises 10 items assessing pleasure in exploring (e.g., "I like to find new ways to reach familiar places"), with 5 positive and 5 negative items. For scoring purposes, the reverse score of the negative items was considered. Responses in the SDSR and AtOT were given on Likert scales ranging from 1 (not at all) to 5 (very much).

The internal consistency of all tasks and factors in the questionnaires were shown to be good (Cronbach's alpha from .71 to .86).

2.2.2 Descriptions, strategy use measures and recall tasks

In the second sessions, participants listened twice (for 6 minutes in all) to a non-spatial description, or to route or survey spatial descriptions (balanced across participants). After hearing each description participants scored their self-reported strategy use and completed the verification test and the graphical representation task. The descriptions, the strategy scale and the recall tasks are described below.

Descriptions

Non-spatial descriptions. Two descriptions were used ("grape harvest" and "olive oil", adapted from [28]). The descriptions describe the phases of wine production (from the grape harvest to bottling, and the differences between red and white wine), or olive oil production (from refining to bottling, and the different types of oil).

Spatial descriptions. Four descriptions of two fictitious outdoor environments were used ("tourist center" and "holiday farm", adapted from [23], two presented from a route and two from a survey perspective. In the survey version, the description first outlined the layout of the environment, then defined the relationship between landmarks using canonical terms (e.g. "north", "south-east"); in the route version, the description was given as if a person were walking along a route and the positions of the landmarks were presented as seen by the person using egocentric terms (e.g. "on the left", "turning right").

All descriptions were of similar difficulty (as tested in previous studies). They contained 14 units of information (in the non-spatial descriptions) or 14 positions of landmarks (in the spatial descriptions), and were all of similar length (between 288 and 309 words). Examples of the descriptions are given in Table 1. The descriptions were presented using .mp4 files (each presentation taking 3 minutes).

Table 1 Examples of non-spatial, route and survey descriptions; and examples of sentences in the verification test.

Non-spatial text (grape harvest)	Route description (tourist center)	Survey description (tourist center)
"[...] There are two types of vinification process, i.e. two different ways to make wine, for red and white wine. [...] Before bottling, the wine undergoes a crystallization process, when it cooled to sub-zero temperatures of around −5°C. This procedure lasts 2 days and enables the excess tartar to deposit so that it can be eliminated later."	"[...] Go straight ahead and you will soon see the tennis courts, which are used for a number of local competitions; they are on your left, at the end of the oak wood. Keep going as the road bends slightly to the right and, beyond the bend, on your left, you will see the hills that surround the whole area."	"[...] a dense oak wood, famous for its many centuries-old trees, stretches from north to south. This dense oak wood extends to the south as far as the tennis courts. At the southernmost tip of the lake there are hills stretching from east to west across the whole area of the tourist center."
Verification test		
During fermentation the new wine is stored at sub-zero temperatures. (False)	As you go towards the hills, you will find the oak wood on your right. (False)	The tennis courts are to the south of the hills. (False)

Strategy use scale. Three strategies were considered (as in [25]): survey ("I form a mental map"), route ("I imagine the path to cover"), and verbal ("I mentally repeat the information"). Participants were asked to judge their strategy use on a Likert scale, ranging from 1 (not at all) to 5 (very much).

Verification test. For each description, twenty true/false sentences were used, half of them true, the other half false (adapted from [23]). The sentences assess inferential information drawn from the non-spatial, route and survey texts (examples are given in Table 1). One point was awarded for each correct answer (maximum score: 20).

Graphical representation. For the non-spatial text, participants were asked to produce a diagram or a list containing the core units of information. For the survey and route texts, they were asked to draw a map of the environment described. In both cases, participants freely reproduced the information on a sheet of paper. They scored one point for each unit of information (in the non-spatial texts) or landmark (in the spatial texts) correctly reported (maximum score: 14).

3 Results

3.1 Correlations between variables

Concerning the correlations between the strategies used and the recall tasks (considering as significant the values $\geq$ .26, corresponding to $ps \leq$.001 according to Bonferroni's correction), there was a significant correlation between the route and survey strategies and both the recall tasks on the route description (verification test-route strategy: $r = .30$; verification test-survey strategy $r = .28$; map drawing-survey strategy: $r = .33$; with $ps \leq .01$). There were also significant correlations for the survey strategy with survey description recall performance (verification test-survey strategy: $r = .31$; map drawing-survey strategy: $r =$

Table 2 Descriptive statistics and correlations for verbal and visuo-spatial individual difference measures and description recall tasks.

	1	2	3	4	5	6	7	8
1. Backward digit span (Verbal WM task)	–							
2. Backward Corsi (Visuo-spatial WM task)	.20	–						
3. Reading comprehension task	.04	.06	–					
4. Perspective-Taking Task	−.03	−.18	−.15	–				
5. SoD – preference for survey mode (SDSR)	.01	.18	−.01	−.14	–			
6. Knowledge and use of cardinal points (SDSR)	.05	.18	.06	−.07	**.40**	–		
7. Preference for landmark and route mode (SDSR)	.21	.03	.07	−.19	**.35**	.19	–	
8. Pleasure in exploring (AtOT)	−.06	.10	.10	−.16	**.70**	**.38**	**.36**	–
Non-spatial descriptions – Verification test	.14	.08	.10	−.12	.04	.11	.05	.09
Non-spatial descriptions – Diagram	.08	−.04	**.28**	−.14	−.05	−.08	.09	.07
Route descriptions – Verification test	.19	.19	.25	**−.26**	.16	.18	.24	.25
Route descriptions – Map drawing	.25	**.29**	.22	**−.35**	.19	.23	**.26**	**.27**
Survey descriptions – Verification test	**.27**	.21	.16	**−.26**	.10	.21	.17	.20
Survey descriptions – Map drawing	.22	.10	.22	−.18	−.02	.16	.10	.10
M	5.32	5.39	6.87	29.87	17.2	5.06	14.49	29.43
SD	1.29	1.27	1.78	21.51	4.86	2.23	2.29	8.83

Note. $N = 173$. The values of the correlations considered significant are shown in bold type, with $p \leq .001$. SDSR = Sense of Direction and Spatial Representation scale; AtOT = Attitudes to Orientation Tasks scale. For the Perspective-Taking Task we report the degrees of error.

.28, $p \leq .01$), but not with route strategy and survey description recall performance. No significant correlations emerged between verbal strategy use and survey or route description recall performance.

For the correlations between the individual differences in the objective of verbal or visuo-spatial measures and in the recall of the descriptions (see Table 2), we found that – for the non-spatial descriptions – only accuracy in the diagrams of the non-spatial text correlated with reading comprehension task performance (no other significant correlations involving recall accuracy were found); for the route descriptions, performance in both the verification test and the map drawing task correlated with PTT, and only the map drawing task correlated with the backward Corsi task, a preference for landmark and route modes (SDSR), and pleasure in exploring (AtOT) (with $ps \leq .001$). For the survey description, there were correlations between the backward digit span and PTT, but only with the verification test.

3.2 Regression analyses

Regression analyses were run to analyze the predictive value of verbal and visuo-spatial abilities and self-reported preferences and strategies on recall performance (in the verification test and graphical representation task) for all types of description (non-spatial, route and survey). Two independent judges scored performance in the graphical representation task and their scores correlated closely ($rs \geq .93$, $p \leq .001$), so the analyses were run on the scores awarded by the first judge. The order in which the variables were entered in the models was based on theoretical grounds. Given the verbal format used to present the environmental information [7] the contribution of visuo-spatial factors was analyzed after controlling for verbal abilities. Therefore, after controlling for verbal abilities (step 1), it was then we examined the contribution of visuo-spatial cognitive abilities (step 2), self-reported visuo-spatial preferences (step 3), and visuo-spatial strategies (step 4). The verbal strategy was not taken into account because it revealed no correlation with any type of description.

Table 3 Regression analyses for the verification tests and the diagram/list or map drawing tasks, by type of description (non-spatial, route and survey).

Predictors	Non spatial description					Route description					Survey description				
	ΔR2	Evidence ratio based on AIC	ANOVA on steps	β	p	ΔR2	Evidence ratio based on AIC	ANOVA on steps	β	p	ΔR2	Evidence ratio based on AIC	ANOVA on steps	β	p
Verification test															
Step 0															
Step 1: Verbal abilities	.03	1.41	.10			**.09**	**563**	**<.001**			**.09**	**743**	**<.001**		
Backward digit span (VWM)				.13	.08				**.18**	**.02**				**.26**	**<.001**
RCT				.09	.23				**.24**	**.001**				**.15**	**.04**
Step 2: Visuo-spatial abilities	.01	0.37	.38			**.06**	**53**	**.002**			**.07**	**83**	**.001**		
Backward Corsi (VSWM)				.03	.67				.11	.16				.12	.11
PTT				−.10	.21				**−.21**	**.005**				**−.21**	**.004**
Step 3: Self-reported visuo-spatial factors	.00	0.58	.35			**.04**	**.25**	**.004**			**.02**	**5**	**.02**		
Pleasure in exploring (AtOT)				.07	.35				**.21**	**.004**				**.16**	**.03**
Step 4: Strategies	.01	0.32	.44			**.05**	**15**	**.01**			**.05**	**25**	**.007**		
Route				−.12	.20				.16	.06				−.08	.30
Survey				.06	.49				.08	.34				**.25**	**.002**
Total R2	*.05*					*.24*					*.23*				
Graphical representation															
Step 0															
Step 1: Verbal abilities	**.08**	**219**	**<.001**			**.11**	**2993**	**<.001**			**.10**	**900**	**<.001**		
Backward digit span (VWM)				.07	.35				**.25**	**<.001**				**.22**	**.004**
RCT				**.28**	**<.001**				**.21**	**.004**				**.22**	**.004**
Step 2: Visuo-spatial abilities	.02	0.62	.23			**.12**	**1*10^6**	**<.001**			.02	1.04	.13		
Backward Corsi (VSWM)				−.09	.24				**.19**	**.008**				.02	.78
PTT				−.11	.15				**−.29**	**<.001**				−.16	.06
Step 3: Self-reported visuo-spatial factors	.00	0.42	.63			**.05**	**53**	**.002**			.01	0.77	.22		
Pleasure in exploring (AtOT)				.04	.63				**.21**	**.002**				.09	.23
Step 4: Strategies	.01	0.23	.61			**.04**	**7**	**.02**			**.04**	**14**	**.01**		
Route				.09	.35				.02	.99				.08	.28
Survey				−.02	.84				**.27**	**.02**				**.18**	**.03**
Total R2	*.11*					*.32*					*.17*				

Note. $N = 173$; VWM = Verbal Working Memory; VSWM = Visuo-Spatial Working Memory; PTT = Perspective-Taking Test; AtOT = Attitudes to Orientation Tasks scale. Evidence ratio is based on the AIC of the various steps (each step is a model); the “ANOVA” column shows the comparison between one step and its predecessor. Significant values in bold type.

Predictors were entered at each step, and were only considered relevant if they contributed to reducing the model’s Akaike Information Criterion (AIC). This index enables the relative quality of alternative models to be compared for a given dataset: the better the model, the lower its AIC [3]. Thus, the evidence ratio based on the AIC of the models and the F-test were used to confirm an improvement of the model from one step to the next. The R2 was also reported to account for the variance explained. All the models were checked for outliers (Cook’s distance <1). First of all, in step 3 (self-reported measures) we added the SDSR factors (SoD – preference for survey mode, Knowledge and use of cardinal points, Preference for landmark and route mode), and Pleasure in exploring (AtOT). These SDSR factors were never found significant and in the final analyses only Pleasure in exploring was considered in step 3. The results are summarized in Table 3 and presented in Figure 1, which includes – for each dependent variable and for each step – the $\Delta R2$, the evidence ratio based on the AIC with respect to the previous step, the ANOVA comparing one step with its predecessor, and standardized β and p values.

Non-spatial descriptions. The predictors explained 5% of the overall variance in the verification test and 11% in the diagram/list task. No relevant predictors were found for the verification test, and the RCT was the only relevant predictor for the diagram/list task.

Route descriptions. The predictors explained 24% of the overall variance in the verification test, and 32% in the map drawing task. For the verification test, the relevant predictors were: backward digit span and RCT (step 1), PTT (step 2), Pleasure in exploring (step 3). The effect of using a route strategy tended to be significant. For map drawing task, the relevant

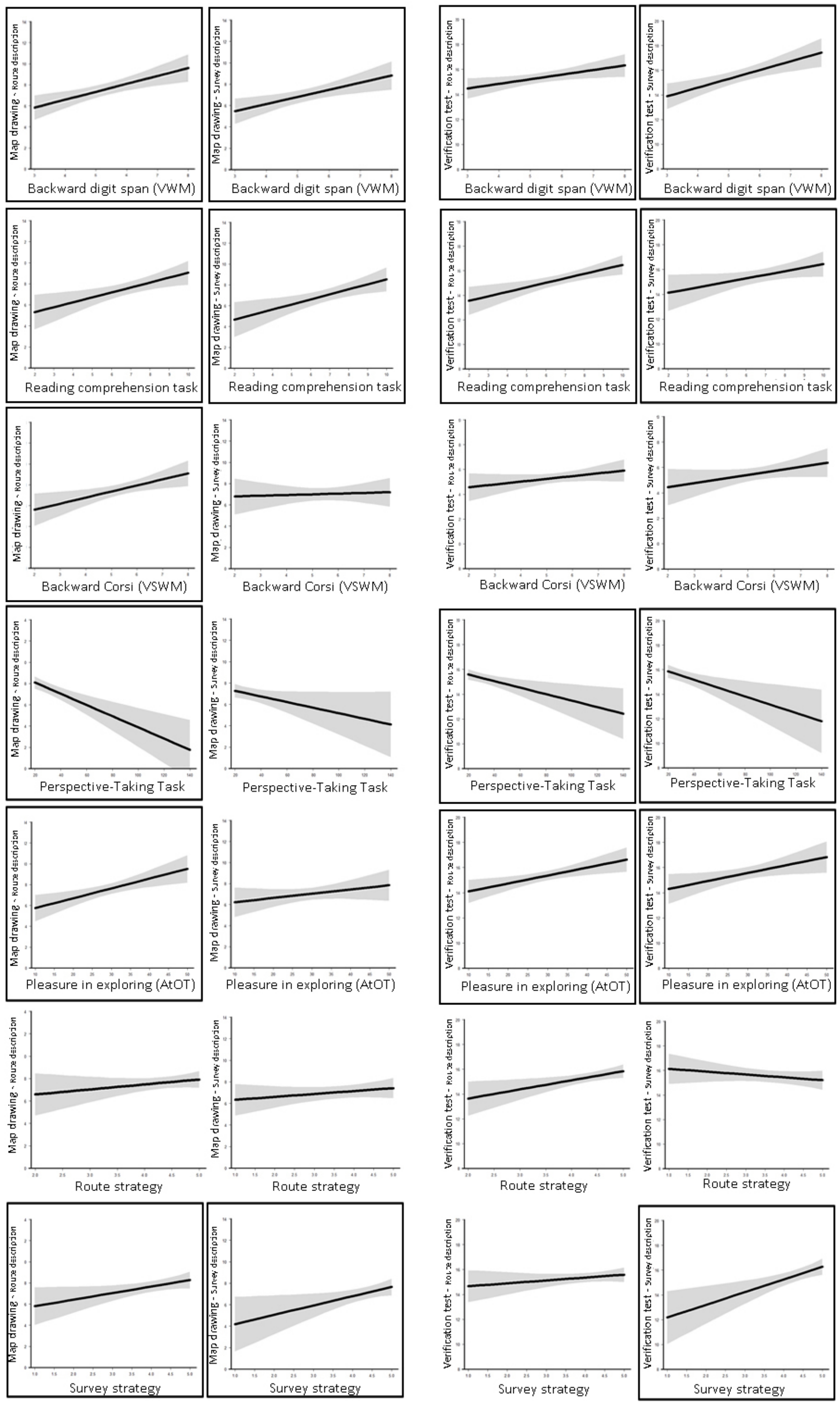


Figure 1 Effects of relevant predictors of route and survey descripitons accuracy in map drawing (first two coloumns) and verification test (second two coloumns). The figures with border indicate significant predictors $p \leq .05$).

predictors were: backward digit span and RCT (step 1), backward Corsi and PTT (step 2), Pleasure in exploring (step 3), and use of a survey strategy (step 4).

Survey descriptions. The predictors explained 23% of the overall variance in the verification test and 17% in the map drawing task. For the verification test, the relevant predictors were: backward digit span and RCT (step 1), PTT (step 2), Pleasure in exploring (step 3) and use of a survey strategy (step 4). For the map drawing task, the relevant predictors were: backward digit span and RCT (step 1), and use of a survey strategy (step 4).
It is worth noting that all steps were significant for the route descriptions (in all tasks) and for the survey descriptions (in the verification test), suggesting that adding the predictors improved the models (as shown by the evidence ratio based on the AIC). In the case of the map drawing task after presenting a survey description, the significant steps were step 1 (backward digit span and RCT) and step 4 (survey strategy).

4 Discussion and Conclusions

The present study was based on the following premises: (i) spatial language is commonly used to convey environmental information with different functions and aims [32, 19]; (ii) people's visuo-spatial competences influence the quality of their visually-acquired environment knowledge [14, 40]; and (iii) most of the contribution of individual visuo-spatial factors in supporting the acquisition of spatial description (especially from a route perspective) derives from the consideration of certain factors (such as cognitive abilities or self-reported preferences). There is therefore a shortage of evidence of the simultaneous contribution of cognitive abilities and self-reported preferences and strategies in supporting the recall of spatial descriptions from survey and route perspectives, as measured with different recall tasks. In particular, we explored whether it is possible to detect – beyond the contribution of verbal abilities – the specific role of visuo-spatial (cognitive and self-assessed) abilities, and possibly also their different role in predicting accuracy in recall performance, in relation to the perspective learnt and the modality used to assess it.

First, regression models showed that the learning of both visuo-spatial and verbal descriptions was supported by verbal abilities. In particular, reading comprehension ability (measured with the RCT) supported non-spatial description accuracy only when recalling information in a schematic form (not in the verification test). Ability in the RCT and the VWM task (backward digit span) supported route and survey description recall (in both the verification test and the map drawing task). This result shows that verbal abilities support the learning and recall of descriptions – as expected, given that a description is verbal per se, irrespective of the content [7, 42].

Second, for spatial descriptions there is a role for visuo-spatial abilities too, as well as for verbal abilities. The contribution of visuo-spatial cognitive abilities and self-reported preferences and strategies clearly emerged for the survey and route descriptions. In particular, spatial (rotation) ability predicted performance in the recall of route descriptions (in both the verification test and the map drawing task) and survey descriptions (in the verification test, while only a trend was found for the map drawing task), while only VSWM predicted map drawing performance after learning route descriptions. Judging from these results, learning route descriptions seems more demanding on WM (in both its visuo-spatial and its verbal aspects) than learning survey descriptions, especially when map drawing is used to test recall [2, 30, 10]. This supports the hypothesis that route description in association with an active reproduction is cognitively more demanding [2, 30, 10].

Further regression models showed that the role of visuo-spatial factors changes in relation to perspective and how recall is assessed, especially for visuo-spatial preferences and strategies. Concerning visuo-spatial preferences, the results revealed the predictive role of pleasure in exploring for route descriptions (in both the verification test and the map drawing task) and for survey descriptions (in the verification test). This result suggests that pleasure in exploring represents a positive personal attitude to approaching (moving in, and guiding others in) environments. The contribution of pleasure in exploring to how environment learning is approached seems to be relevant not only when an environment is conveyed visually [27]: having a positive general attitude to exploring an environment (by moving around in it) was newly related to the ability to represent verbally-conveyed spatial information. This attitude appears to be part of a particular spatial profile (since it also relates to sense of direction and a preference for using a survey mode) [9, 27] as also shown by the correlations in Table 2.

Concerning the visuo-spatial strategies that participants reported having used to understand and recall the descriptions they had heard, it is worth emphasizing that accuracy in both survey and route description recall were significantly associated with the participants' rating of their use of visuo-spatial strategies, but not with their use of verbal strategies (as found previously [25]. To be more specific, survey descriptions were associated with the use of a survey strategy in both the verification test (when participants were asked to judge the truthfulness of spatial relations between landmarks) and the map drawing (when they had to arrange the landmarks on a map). Route descriptions tended, on the other hand, to be associated with the use of a route strategy in the verification test (route view) and with the use of a survey strategy in the map drawing task (survey view). In other words, survey descriptions seem to be more associated with the use of a survey strategy, while route descriptions seem to be associated with the use of both survey and route strategies (as also shown by the correlations and previously suggested [24]). These results show the relation between self-reported visuo-spatial strategy use and spatial description recall accuracy (albeit with some differences depending on the perspective learnt). Therefore, it is not only when the use of visuo-spatial strategies is recommended that their use influences recall accuracy [37], but also when they are used spontaneously: learning a spatial description elicits the spontaneous use of strategies, and the survey strategy in particular.

The route descriptions warrant a few specific considerations. Our results indicate that learning from route descriptions is supported largely (and more than when learning survey descriptions) by visuo-spatial cognitive abilities and self-reported preferences and strategies. This was especially evident when recall was tested on graphical reproduction (in the map drawing task it explained a larger share of the variance, 32%, than the other models run) [2, 30, 10]. The route descriptions were also associated with the use of both route and survey strategies (in line with [25, 24]), suggesting that they prompt a greater degree of flexibility in people's approach to learning from this type of input. On the other hand, survey description learning, as assessed with a map drawing task, would be less demanding in terms of visuo-spatial cognitive abilities (since the step in the regression for visuo-spatial abilities and self-reported preferences did not improve the models).

We wish to acknowledge some of the limitations of the present study. One concerns the all-female sample considered. While this choice restricted the variability, our results are only applicable to young females (all university students in our case). Certainly, males will need to be considered in further studies before our findings can be generalized to the population as a whole. Another issue concerns our spatial descriptions, which were created ad hoc and balanced for length and quantity of information, but were fictitious, not representing real paths (like those shown on the Google Maps website, for instance). It would therefore be interesting to analyze to what extent our results can be generalized to descriptions of real paths or maps. It will also be interesting to explore to what extent spatial descriptions represent the "large scale": even though the passages present large-scale spatial information, we cannot say for sure that participants represent it in terms of large-scale exploration. Moreover, given the interesting role of strategy use in supporting graphical representation accuracy, further studies should more carefully consider criterion scores capable of detecting strategy use in mentally representing survey and route information. Finally, even if our results show the similarities and differences in the contributions of a set of cognitive abilities and self-reported preferences and strategies, it is important to bear in mind that cognitive abilities (both verbal and visuo-spatial) could share processes, and be part of the human intelligence construct (e.g. [17]), so more studies are needed to investigate the relationship between these predictors of environment recall performance.

Overall, these results can be considered consistent with spatial cognition models showing the relationship between small-scale abilities (i.e. individual visuo-spatial features) and large-scale abilities (environment learning [14]), considered here in terms of spatial descriptions. The novelty of our findings lies in that, beyond the contribution of verbal factors, multiple individual visuo-spatial aspects (both cognitive abilities and self-reported factors) need to be considered, and their influence varies as a function of the perspective learnt and the task used to assess recall. Certain learning conditions are more demanding than others (such as map drawing after learning from a route description, as opposed to a survey description), and show the role of certain preferences (such as pleasure in exploring) and strategy use (such as a survey strategy). The present study thus expands our theoretical understanding of how individual visuo-spatial factors influence mental representations of environmental information (in female undergraduates at least), and may have relevant implications in other related disciplines. For instance, for the software implemented by computer scientists to be capable of handling spatial language [32], it should – to some extent, at least – take the user's or speaker's individual differences into account. Our results indicate that the formation of a representation in map view after learning from a route description is more demanding, so such software should present descriptions or information using a survey view. It will be interesting to improve on this line of research by cooperating with other disciplines interested in spatial language.

To conclude, the present study points to the importance of analyzing individual factors (which include several relevant visuo-spatial competences, preferences and strategies) when examining the quality of mental representations of environments derived from spatial descriptions.

References

1 Gary L. Allen, K. C. Kirasic, Shannon H. Dobson, Richard G. Long, and Sharon Beck. Predicting environmental learning from spatial abilities: An indirect route. *Intelligence*, 22:327–355, 1996. `doi:10.1016/S0160-2896(96)90026-4`.

2 Tad T. Brunyé and Holly A. Taylor. Working memory in developing and applying mental models from spatial descriptions. *Journal of Memory and Language*, 58(3):701–729, 2008. `doi:10.1016/j.jml.2007.08.003`.

3 Kenneth P. Burnham and David R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer, 2nd edition, 2002. `arXiv:arXiv:1011.1669v3`.

4 Heather Burte and Mary Hegarty. Individual and strategy differences in an allocentric-heading recall task. In *Proceedings of the 35th annual conference of the Cognitive Science Society*, pages 1958–1963. Austin TX: Cognitive Science Society, 2013.

5 Cornoldi Cesare, Candela Mariangela, Fenzi Virginia, and Esperia Lavis. *Prove di lettura e scrittura MT-16-19 : batteria per la verifica degli apprendimenti e la diagnosi di dislessia e disortografia : classi terza, quarta, quinta della scuola secondaria di 2° grado*. Centro studi Erickson, 2014.

6 Philip Michael Corsi. *Human memory and the medial temporal region of the brain*. PhD thesis, McGill University, 1972.

7 Meredyth Daneman and Philip M. Merikle. Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4):422–433, dec 1996. `doi:10.3758/BF03214546`.

8 Rossana De Beni, Erika Borella, Barbara Carretti, Cinzia Marigo, and Lucia A. Nava. *Portfolio per la valutazione del benessere e delle abilità cognitive nell'età adulta e avanzata [The assesment of well-being and cognitive abilities in adulthood and aging]*. Firenze, Italy: Giunti OS, 2008.

9 Rossana De Beni, Chiara Meneghetti, Feliciana Fiore, Lucia Gava, and Erika Borella. *Batteria Visuo-spaziale. Strumenti per la valutazione delle abilità visuo-spaziali nell'arco di vita adulta [Visuo-spatial battery: Instrument for assessing visuo-spatial abilities across adult life span]*. Firenze, Italy: Hogrefe, 2014.

10 Emilie Deyzac, Robert H. Logie, and Michel Denis. Visuospatial working memory and the processing of spatial descriptions. *British Journal of Psychology*, 97(2):217–243, 2006. `doi:10.1348/000712605X67484`.

11 Valerie Gyselinck, Rossana Beni, Francesca Pazzaglia, Chiara Meneghetti, and Amandine Mondoloni. Working memory components and imagery instructions in the elaboration of a spatial mental model. *Psychological Research*, 71(3):373–382, 2007. `doi:10.1007/s00426-006-0091-1`.

12 Valerie Gyselinck and Chiara Meneghetti. The role of spatial working memory in understanding verbal descriptions: A window onto the interaction between verbal and spatial processing. In A. Vandienrendonck and Szmalec A., editors, *Spatial working memory*, pages 159–180. Hove: Press., Psychology, 2011.

13 Adamantini Hatzipanayioti, Alexia Galati, and Marios N. Avraamides. The protagonist's first perspective influences the encoding of spatial information in narratives. *Quarterly Journal of Experimental Psychology*, 69(3):506–520, 2016. `doi:10.1080/17470218.2015.1056194`.

14 Mary Hegarty, Daniel R. Montello, Anthony E. Richardson, Toru Ishikawa, and Kristin Lovelace. Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34:151–176, 2006. `doi:10.1016/j.intell.2005.09.005`.

15 Mary Hegarty and David Waller. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32:175–191, 2004. `arXiv:arXiv:1011.1669v3`, `doi:10.1016/j.intell.2003.12.001`.

16 Mary Hegarty and David Waller. Individual Differences in Spatial Abilities. In P. Shah and A Miyake, editors, *The Cambridge handbook of visuo-spatial thinking*, pages 122–169. Cambridge University Press, 2005.

17 Wendy Johnson and Thomas J. Bouchard. The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33(4):393–416, 2005. `doi:10.1016/j.intell.2004.12.002`.

18 Philip N. Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness.* Harvard University Press, 1983.

19 Junchul Kim, Maria Vasardani, and Stephan Winter. From descriptions to depictions: A dynamic sketch map drawing strategy. *Spatial Cognition & Computation*, 5868(March):29–53, 2015. doi:10.1080/13875868.2015.1084509.

20 David F. Lohman. Spatial abilities as traits, processes and knowledge. In R. J. Sternberg, editor, *Advances in the psychology of human intelligence*, pages 181–248. Hillsdale, NJ: Erlbaum, 1988.

21 Chiara Meneghetti, Erika Borella, Valerie Gyselinck, and Rossana De Beni. Age-differences in environment route learning: The role of input and recall-test modalities in young and older adults. *Learning and Individual Differences*, 22:884–890, 2012. doi:10.1016/j.lindif.2012.04.006.

22 Chiara Meneghetti, Erika Borella, Veronica Muffato, Francesca Pazzaglia, and Rossana De Beni. Environment learning from spatial descriptions: The role of perspective and spatial abilities in young and older adults. In *Spatial Cognition IX*, pages 30–45. Springer International Publishing, 2014. doi:10.1007/978-3-319-11215-2_3.

23 Chiara Meneghetti, Rossana De Beni, Valérie Gyselinck, and Francesca Pazzaglia. Working memory involvement in spatial text processing: What advantages are gained from extended learning and visuo-spatial strategies? *British Journal of Psychology*, 102(3):499–518, 2011. doi:10.1111/j.2044-8295.2010.02007.x.

24 Chiara Meneghetti, Enia Labate, Francesca Pazzaglia, Colin Hamilton, and Valérie Gyselinck. The role of visual and spatial working memory in forming mental models derived from survey and route descriptions. *British Journal of Psychology*, pages 225–243, 2016. doi:10.1111/bjop.12193.

25 Chiara Meneghetti, Lucia Ronconi, Francesca Pazzaglia, and Rossana De Beni. Spatial mental representations derived from spatial descriptions: The predicting and mediating roles of spatial preferences, strategies, and abilities. *British Journal of Psychology*, 105:295–315, 2014. doi:10.1111/bjop.12038.

26 Daniel R. Montello, David Waller, Mary Hegarty, and Gary L. Allen. Spatial memory of real environments, virtual environments, and maps. In G.L. Allen, editor, *Human Spatial Memory*, pages 251–285. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. doi:10.4324/9781410609984.

27 Veronica Muffato, Chiara Meneghetti, Virginia Di Ruocco, and Rossana De Beni. When young and older adults learn a map: The influence of individual visuo-spatial factors. *Learning and Individual Differences*, 53:114–121, 2017. doi:10.1016/j.lindif.2016.12.002.

28 Francesca Pazzaglia, Cesare Cornoldi, and Rossana De Beni. Differenze individuali nella rappresentazione dello spazio e nell'abilità di orientamento: Presentazione di un questionario autovalutativo. *Giornale italiano di psicologia*, 27:627–650, 2000.

29 Francesca Pazzaglia and Chiara Meneghetti. Spatial text processing in relation to spatial abilities and spatial styles. *Journal of Cognitive Psychology*, 24(8):972–980, 2012. doi:10.1080/20445911.2012.725716.

30 Francesca Pazzaglia, Chiara Meneghetti, Rossana De Beni, and Valerie Gyselinck. Working memory components in survey and route spatial text processing. *Cognitive Processing*, 11(4):359–369, 2010. doi:10.1007/s10339-009-0353-0.

31 Patrick Péruch, Vanessa Chabanne, Marie-Pascale Nesa, Catherine Thinus-Blanc, and Michel Denis. Comparing distances in mental images constructed from visual experience or verbal descriptions: The impact of survey versus route perspective. *The Quarterly Journal of Experimental Psychology*, 59(11):1950–1967, 2006. doi:10.1080/17470210500539408.

32 M. V. Pestun and V. A. Galaktionov. Algorithms for the construction and recognition of navigational route descriptions for cartographic computer systems. *Programming and Computer Software*, 42(6):341–346, 2016. doi:10.1134/S0361768816060037.

33 Luciana Picucci, Valérie Gyselinck, Pascale Piolino, Serge Nicolas, and Andrea Bosco. Spatial mental models: The interaction of presentation format, task requirements and availability of working memory components. *Applied Cognitive Psychology*, 27(3):314–327, 2013. doi:10.1002/acp.2909.

34 Christopher A. Sanchez and Jennifer Wiley. The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction*, 31:33–45, 2014. doi:10.1016/j.learninstruc.2013.12.007.

35 Amy L. Shelton, Steven A. Marchette, and Andrew J. Furman. A mechanistic approach to individual differences in spatial learning, memory, and navigation. In B.H. Ross, editor, *Psychology of Learning and Motivation*, volume 59, chapter VI, pages 223–259. Elsevier, 2013. doi:10.1016/B978-0-12-407187-2.00006-X.

36 Masashi Sugimoto and Takashi Kusumi. The effect of text continuity on spatial representation: Route versus survey perspective. *Cognitive Processing*, 15(1):65–75, 2014. doi:10.1007/s10339-013-0582-0.

37 H. A. Taylor, S. J. Naylor, and N. A. Chechile. Goal-specific influences on the representation of spatial perspective. *Memory & cognition*, 27(2):309–19, 1999. URL: http://www.ncbi.nlm.nih.gov/pubmed/10226440.

38 Holly A. Taylor and Barbara Tversky. Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31(2):261–292, apr 1992. doi:10.1016/0749-596X(92)90014-O.

39 David H. Uttal, Nathaniel G. Meadow, Elizabeth Tipton, Linda L. Hand, Alison R. Alden, Christopher Warren, and Nora S. Newcombe. The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139:352–402, 2013. doi:10.1037/a0028446.

40 Steven M. Weisberg, Victor R. Schinazi, Nora S. Newcombe, Thomas F. Shipley, and Russell A. Epstein. Variations in cognitive maps: understanding individual differences in navigation. *Journal of experimental psychology. Learning, memory, and cognition*, 40:669–82, 2014. doi:10.1037/a0035261.

41 Thomas Wolbers and Mary Hegarty. What determines our navigational abilities? *Trends in Cognitive Sciences*, 14:138–146, 2010. doi:10.1016/j.tics.2010.01.001.

42 Rolf A. Zwaan and Gabriel A. Radvansky. Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2):162–185, 1998. doi:10.1037/0033-2909.123.2.162.

Cities Untangled: Uncovering Order in Arterial Skeletons of Road Maps

Robert Firth

Informing Design, Inc., Pittsburgh, PA, USA
bob@informingdesign.com

Abstract

Survey knowledge, as embodied in the road map, has been seen as too slow a navigational aid to function effectively at the speed of life in the smartphone/GPS-app era, capturing as it does details of the highway network that are seen to present too heavy a cognitive load to the user. Yet this very richness offers the promise of enabling the user to navigate with understanding, providing for exible and resilient trip planning. But what if the map's heavy cognitive load was not because of the difficulty in dealing with its heavy load of information, but because that information was unnecessarily disordered? We suggest a comprehensible ordering has always existed within complex-appearing road maps. We propose a model for making this ordering explicit, highlighting a "skeleton" of arterials so as to appear visually untangled. The concept of the Use-Access Island (UAI), a bounded area with a coordinate axis-like array of spanning arteries, is introduced. As ever-finer meshes of these areas are highlighted across a street map, a hierarchy of visually untangled arteries can be rendered. Locations and routings can then be visualized in terms of nested sequences of "untangled" routings. When married to geographical designations, this iterative UAI schematization is designed to embody routing spatial knowledge. Is such an untangled map fast enough? We invite researchers to test the model.

1998 ACM Subject Classification H.5.2 User Interfaces

Keywords and phrases map schematization, arterial skeleton, untangled map

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.14

1 Introduction

In the face of ubiquitous adoption of navigational apps and devices, maps have been getting a bad rap. In 2010, Hirtle delivered a talk [9] to the 20th anniversary meeting of Cognitive and Linguistic Aspects of Space entitled, "Dinosaurs, Slide Rules and Maps," suggesting (slightly tongue-in-cheek) the eventual demise of maps. Yet, as Hirtle points out, the loss of spatial awareness from the use of these apps and devices is a serious problem [12, 22, 23].

People are drawn to these new devices, as road maps *can* be problematic. As Klippel et al. [16] pointed out, "employing a map to create spatial awareness or to provide wayfinding support requires the ability of the map user to establish element-to-element correspondence within and between maps and entities in the real world. The perceptual and cognitive costs of recognizing such correspondences are potentially very high." This can be especially true in cities with complex road networks. As Kuipers [17] speculated, "in an area which is not even topologically close to a grid, finding novel routes or relative positions will be characterized by high error rates, low confidence, and conservative strategies." Montello [21] reinforces that idea with respect to route angularity.

Compromise solutions have been proposed, schematizations designed to restore some spatial awareness to the directions and route maps which these devices and apps generate without the drawback of the heavy cognitive load of a full map. For example, Schmid et al. [25]

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 14; pp. 14:1–14:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

proposed “Route Aware Maps” that added limited road map context at key points along strip maps of routes. Zipf et al. [33] proposed “Focus Maps” that highlight important elements of a full map and fade out the less important.

But what if the cognitive load of full road maps was *not* intrinsic to the density of highway, landmark and area information that they embody? What if instead it was the result of the spatial information being *unnecessarily* disordered in its presentation? What solutions to the problem of restoring spatial awareness would be possible then?

This paper proposes a schematization that does not eliminate any spatial information from a given scale of map but instead highlights an arterial “skeleton” of the highway network so as to reveal a remarkably coherent ordering of highway elements even within complex cities and across a wide range of granularity, up to the level of continent-wide. Our shorthand term for such a schematization is “untangled map.”

We found Freksa’s paper [6] on strong spatial cognition inspiring: the insight to look at spatial problem-solving in physical (bodily) terms and not solely on an “abstract information level,” as in having a robot solve a shortest route problem using string and physical manipulation and not calculation, would seem to hold tremendous promise for AI. There is a sense in which untangled maps can be thought of as “strong,” not to suggest applications for AI or robots, but rather to suggest the “role of the body” – in particular our perceptual machinery – in spatial problem-solving using this schematization. Can road maps overlaid with an organizing skeleton of colored strings, as it were, enable users to simply see their way to destinations on maps, painlessly restoring spatial awareness?

2 What Skeletons Have Been Hiding

Many researchers have noted that local driving experts in major cities typically exhibit a common approach to wayfinding [2, 8, 29]. They rely on what has been termed a “skeleton” of major highways [19], a finding we independently discovered in our work developing maps for wayfinding traffic signage. We called this skeleton “intermediate wayfinding paths” [3] or “tourism areas and corridors” [5]. The experts’ wayfinding strategy consists of finding the most efficient path from the starting point to the skeleton, and then to traverse the skeleton to the turn-off for the most efficient path to the end point. This remarkable result – a small subset of streets being independently discovered for optimizing wayfinding to all destinations in a given area – raises the question: what network properties distinguish these streets from the others?

Benjamin Kuipers [18] proposed a hypothesis: that the skeleton was composed of the streets rich in boundary relations – that is, a street that serves as a boundary between regions containing destinations, and thus rich in turn-offs for those destinations. Tomko et al. [31] proposed a similar approach based on space syntax: suggesting that ranking streets by between-ness centrality values plausibly corresponds to “experiential hierarchies of streets.”

In our two decades of field studies for many of the largest traffic wayfinding sign installations in North America, our mapping of wayfinding skeletons largely agree with the above hypotheses; however, our experience leads us to propose that using such purely computational analyses to determine “the skeleton” misses both cultural and topographical considerations that can affect which highways actually make the cut as the commonly traveled set of core wayfinding corridors. For example, there may be districts of a city with high crime rates that lead drivers to avoid certain arteries. And in a city like Pittsburgh, with its crazy topography of hills and hollows and rivers, there can be both cultural biases rooted in geography (“going south of the Mon River scares me”) and topographical biases (“I avoid tunnels” or “I’m taking Bigelow Blvd, it’s simpler”).

In fact, the cultural/geographical aspects of route selection were central to our sign-planning process. Our skeletal corridors would always be framed in terms of encompassing areas which would be notable for two things: they each comprise a sense of place at a distinct level of granularity and each exhibit a distinct wayfinding skeleton for travel at that level of granularity. Our areas with a sense of place correspond to Lynch's "districts," or at a finer granularity, "nodes." As he put it, districts are areas which the "observer mentally enters 'inside of,' and which are recognizable as having some common identifying character." [19]

In our sign systems, we would coin names for these areas, based on this common identifying character, and then point to the arterial skeleton inside using these names. The purpose was to create a thematic hierarchy of signage, such that an area's signed name would be predictive of the destinations to be found within. Led to a wayfinding corridor component of the skeleton inside the area (the "tourism corridor"), there would then be signs at the turn-offs for individual destinations of tourist or civic interest.

Yes, creating a hierarchy of signing is a common strategy for "wayfinding sign system" designers [7, 24]. However, we found that districts in such systems were often created for promotional or aesthetic purposes and not with regard for how such districts cohere for navigational purposes. That is, paths to districts would be signed that were actually the wrong way to go for some of the destinations within them. In the next section, we introduce the concept of the "Use-Access Island" (UAI), which we use to formalize the idea of the "well-formed" district, one whose arterial skeleton provides for shorter routings to all the destinations within the area when compared with out-of-area routings (starting from area boundary crossings).

While developing our first such sign system for the City of Pittsburgh (1994–1996), we noticed that for all of the city's notorious complexity, these arterial wayfinding skeletons could be rendered with coordinate axis-like simplicity, and at distinct levels of granularity, with one level's coordinate axes serving to organize the next lower level of areas. That is, there was an untangled map of skeletons revealing an inner logic and clarity as to how the city was organized, one level at a time. In 1997, we first published untangled maps in our atlas, "Finding Yourself in Pittsburgh," with UAIs ranging from continent-wide to local neighborhood in size [4]. Several editions of atlases later, we are now developing the *citytunr* app platform based on simultaneously syncing untangled maps for driving, transit and biking. In the next sections, we outline our untangled mapping approach.

3 Constructing the Untangled Road Map

The Use-Access Island shall be defined as a bounded area:

1. that exhibits a distinct sense of place as compared to surrounding areas of comparable geographical extent (e.g., countries, metro regions, cities, city districts, neighborhoods);
2. that encompasses a coordinate axis-like array of the wayfinding skeleton of arterials meeting the Untangling Conditions (as specified on the following pages),
3. and such that, from a given boundary crossing, this skeleton provides for shorter routings to any destination within the area as compared to routings that include out-of-area roads.

The UAI shall be our unit of map untangling: a well-formed "there," as it were. We can point to such a "there" with confidence – by definition, its skeletal elements will be able to lead optimally to turn-offs for all the destinations within the area. At a given level of geographical extent, the collection of UAIs in a region presents readily perceptible local orderings. These UAIs are in turn ordered by coordinate axis-like arterials of the next higher level (in terms of geographical extent) UAIs. And within each UAI, there may be lower order UAIs. The

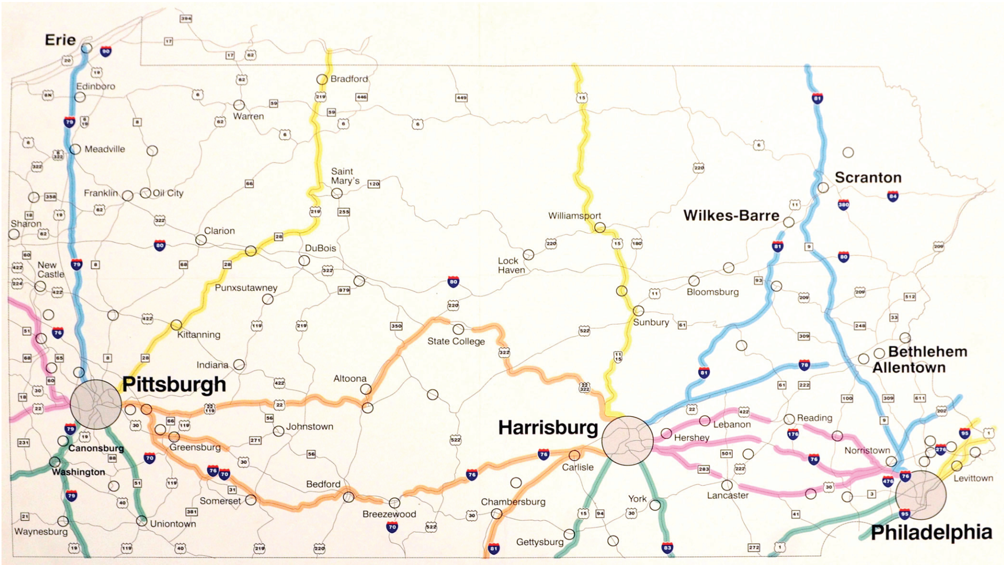


Figure 1 Untangled map of Pennsylvania metro region highways (by the author).

full mechanics of UAI skeleton construction are beyond the scope of this paper (though we do sketch out some of the method in the following), but we begin the mapping out of a UAI by locating a candidate for wayfinding corridor, an artery (or sequence of arteries) that roughly spans the area and that will typically have a high between-ness centrality value and high traffic volume. This corridor then sets the template for determining the given UAI's coordinate axis-like arterial skeleton.

Note that this wayfinding corridor may well correspond to Jiang's [14, 13] "natural road" in the sense that it may consist of "joined road segments based on the Gestalt principle of good continuity." However, unlike Jiang's natural roads that self-organize with respect to predicting traffic flow, the UAI arterial skeletons are *deliberately* organized for the purpose of displaying a ready coherence as coordinate axis-like arrays, with one spanning artery after another highlighted according to the Untangling Conditions below. Thomson [28] did examine the use of natural roads (which he termed strokes) to automatically produce generalizations of road networks that reduce their complexity by eliminating less important streets based on length and road quality. This method, however, did not relate to the generation of perceptible patterns within complex networks.

When studying a new city, we typically start by looking at the metro region as our first UAI. Unfettered by terrain (or absent grid "planning"), there is typically a city center from which emanate radiating major highways. It is interesting to picture a state like Pennsylvania (Figure 1) as being comprised of three regional UAIs (Pittsburgh, Harrisburg and Philadelphia). Those city centers can be thought of as like great seas to which the rivers of radiating highways are finding their way. These radiating arterials are color-coded so as to provide a consistency from one metro region to the next.

When uncovering coordinate axis-like patterns within city centers, we would come upon bothersome regional highways cris-crossing the local pattern and entangling what otherwise would be a simple rectangular grid. Often such "misbehaving" highways were operating at a coarser level of granularity. By dramatically thickening the lines for such highways, the local visual pattern would become clear, as would the regional one, as distinct sets of coordinate axes.

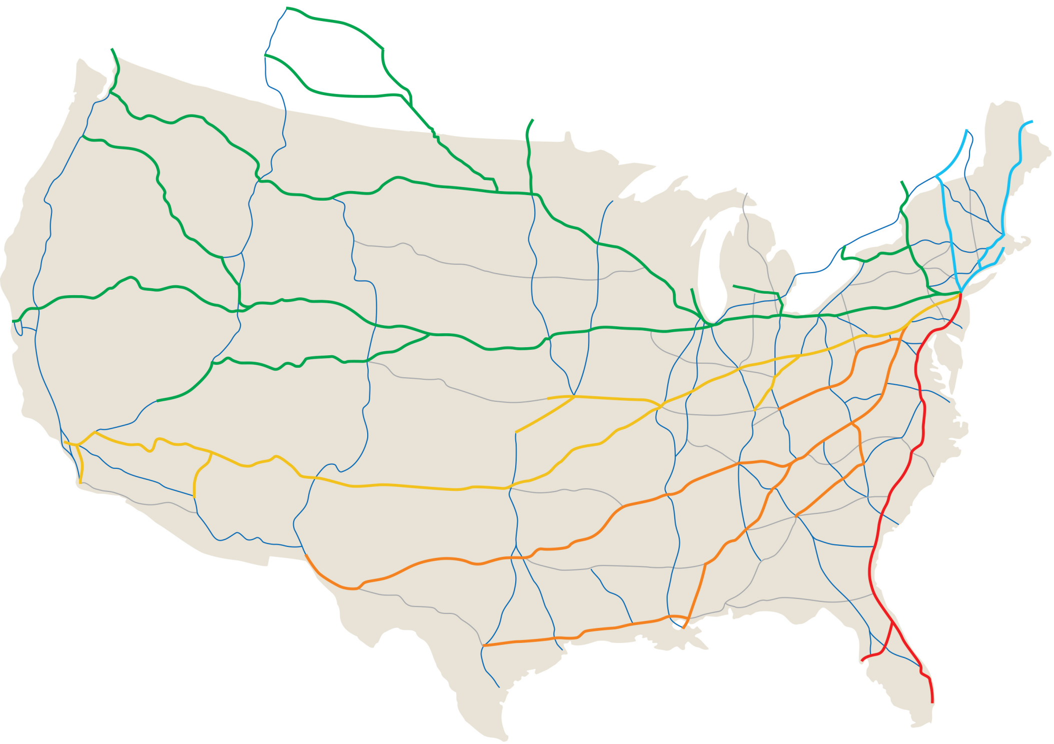


Figure 2 Map of U.S. interstate highways as polar coordinate axes centered on New York City (by the author).

In fact, regional arteries can come upon "misbehaving" highways that are actually operating at a continental level. Early in our work, we treated North America as a single UAI. Traffic engineers see the U.S. Interstate system as largely a rectangular grid.

For the purposes of visually untangling metro region maps, we found rendering the Interstate system as a polar coordinate system centered on New York City would satisfy our graphical requirements (Figure 2). We chose a color-coding scheme that went from a hot red to the south to cold blue to the north. ("Misbehaving" continental elements of the radial system are in gray, acting as their own axis system of shortcuts, as it were.)

The general form for visually untangling a metro region's skeleton takes the appearance of overlapping coordinate axis-like patterns, with each level of granularity distinguished by thickness of line and color (Figure 3). The thicker the line is, the longer the range of effect of the indicated routing. Continental-range radial highways are rendered in bright colors and the "circumferential" highways in dark blue. In the UAI of metro regions, radials are in light blue and circumferentials in yellow. Within the city center, a local grid is indicated with purple and pink axes.

Of course, things can get quite complicated inside a city center. In the real world, grids of roadways bump into barriers, collide with one another, follow old cow paths, are forced to cope with mountains and valleys, and can generally get into all kinds of mischief to keep from looking well-ordered. With the help of graphical conventions that we developed, we have found it possible to resolve any city's wayfinding skeleton into layers of UAIs of comparable geographical extent such that the coordinate axis-like appearance within each UAI can be readily perceived.

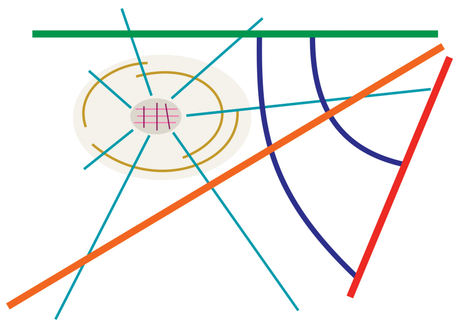


Figure 3 General form of untangling a metro region.

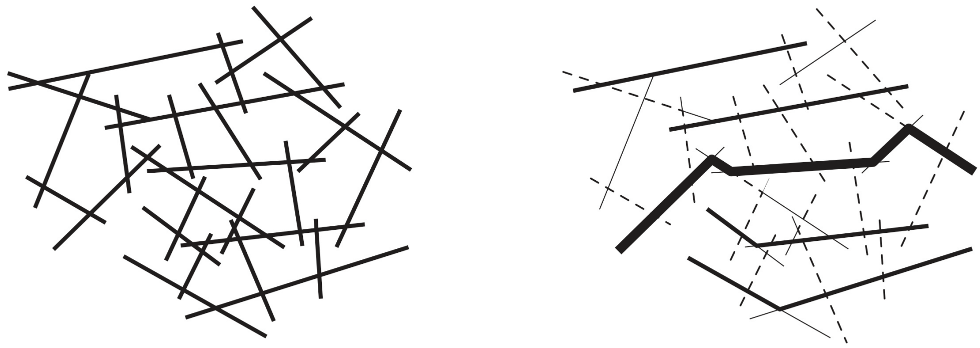


Figure 4 Pick-up Sticks Map – before untangling (left) and after (right).

For example, to point to a particularly perverse possibility, we created a pick-up sticks-like map (Figure 4), and then "untangled" it by resolving it into four implied layers of routings (thick black, thin black, dashed lines and light gray lines).

The Pick-Up Sticks Map meets our Untangling Conditions for an untangled map, as follows:

1. Skeletal arteries shall be rendered as a coordinate axis (rectangular or polar) of a recognizable type (radiating, circumferential, x-axis, y-axis) when they share the same general sense of flow for that type over the range of their encompassing UAI.
2. No axis of a given type for a given UAI shall cross another of the same type.
3. A given UAI may exhibit more than one layer of coordinate axes, as long as each layer is assigned a distinct degree of line thickness or color-type (bright colors vs gray tones, for example).
4. A coordinate axis arterial may branch at either end or both ends, as long as the branchings do not cross one another and they exhibit the same general sense of flow.

Not allowing like-axes to cross, though permitting them to branch, have proven sufficient in practice to preserve the appearance of coordinate axis-like ordering with the help of our graphical conventions: we allow for "connectors" and "separators" within our wayfinding skeletons. If you look closely at our Pennsylvania example above, we did not connect the orange arterials emanating from Pittsburgh to the ones from Harrisburg. In practice, a section of the meeting point would be rendered in gray. In general, the highways of one

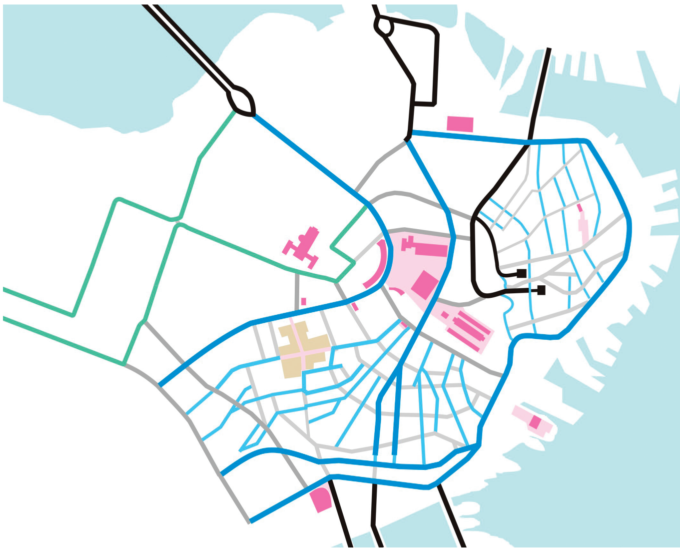


Figure 5 Boston untangled (by the author).

coordinate system can interconnect with the highways of another in a not well-behaved way, whether in the same UAI or neighboring UAIs or overlapping UAIs of differing granularity. This corresponds to how highways have been built over time, or by different entities at the same time in neighboring places. An 1890s road grid can be served by a 1930s era bypass highway which in turn can be served by a 1960s era expressway. How these highways of different eras interconnect is often not pretty (not orderly in appearance, in other words).

The solution is to allow for what we term routing objects – series of road segments that serve to link skeletal elements of different coordinate axis systems. We typically use gray tones to indicate such connectors. Note that connectors can even be quite long, serving as shortcuts cutting across an otherwise orderly grid. By having them join the family of gray connectors, they become their own level of meta-axes as it were. What we have found is that with the use of these connectors and separators, map readers appear able to readily perceive the coordinate-like systems of different types across multiple layers.

4 Dealing with Terrain and "Naturalistic" Street Plans

Given an unlimited number of layers, and the availability of connectors and separators, in theory there is no network that cannot be resolved into a technically untangled map. In practice, there is a bit of human artistry at work to produce maps in which humans can readily perceive the layers of skeletal elements as coordinate axis-like.

Terrain can obviously distort the shape that skeletons take. In Figure 5, the UAI of Boston's North End and Downtown is spanned by three thick blue arterials that roughly correspond to the shape of the encompassing peninsula. When the middle arterial branches threaten to touch the outer one, we introduce separators. As these arterials curve around the UAI, they serve to define three smaller UAIs, each with its own rectilinear-like grid of light blue and gray elements. Black lines are the major incoming highway connectors. The pair of green lines span the Back Bay/Beacon Hill UAI as an unusual one-way pairing.

Terrain can actually be a friend of order, even when it would at first appear to be the enemy. Pittsburgh's South Side Slopes (Figure 6) are a notorious tangle of switchbacks.

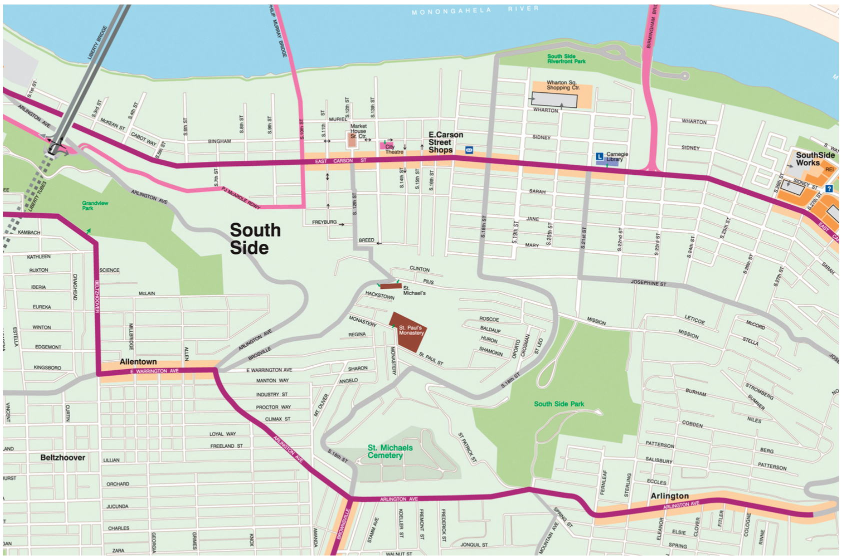


Figure 6 Pittsburgh's South Side Slopes, Untangled (by the author).

Visitors often get lost as routings twist and turn, out of sight of any orienting landmarks. But it turns out that a very simple grid lurked under this apparent disorder.

One street traces the river valley, and it is simple to follow: East Carson St. At the top of the hill, there are a series of streets that trace the ridgeline. One of the tricks the area likes to play on motorists is that it can be difficult to determine how to stay on the ridge. For example, at one five-point intersection, the road ahead that dips is actually the one that eventually rises back up and stays on the ridgeline. The roads ahead that are rising in the immediate view are the ones that just out of sight precipitously drop down to the valley floor.

We solved this by placing signs specifically to keep people on the ridge, and then placing signs for the turns to take you back to the valley floor. At its essence, the terrain is actually enforcing a simple grid of river valley artery, ridgeline artery, and the switchbacks that run between them. No one could perceive that from immediate environmental cues.

In Figure 6, the valley floor route and the ridgeline route are in purple, and the switchbacks running between them are in gray. The pink routes are "y-axis" routings for the Pittsburgh-wide UAI.

Of course, humans can be the design agent behind apparent complexity, independent of terrain. Pinehurst Village in North Carolina is infamous for its naturalistic, curving layout of streets, with few intersections at right angles and with many of them having multiple streets converging simultaneously, and curvaceously. The Village was designed by Olmsted, Olmsted & Eliot [26], including the same Olmsted responsible for the design of the intense tangle of paths that comprises the Central Park "rambles" in New York City. With heavy tree cover and primary landmarks well hidden from nearby state highways, Pinehurst poses a serious navigational challenge to visitors. This is truly a UAI with no apparent coordinate system axes in sight (Figure 7).

It turns out there is a solution for untangling this UAI. In the smaller scale map of the Pinehurst area (Figure 8), the surrounding regional state highways are marked in purple and pink. We have also noted the major landmarks: the Carolina Inn, the #2 Golf Course's clubhouse, the Holly Inn, and the village's business district (shaded in light tan).

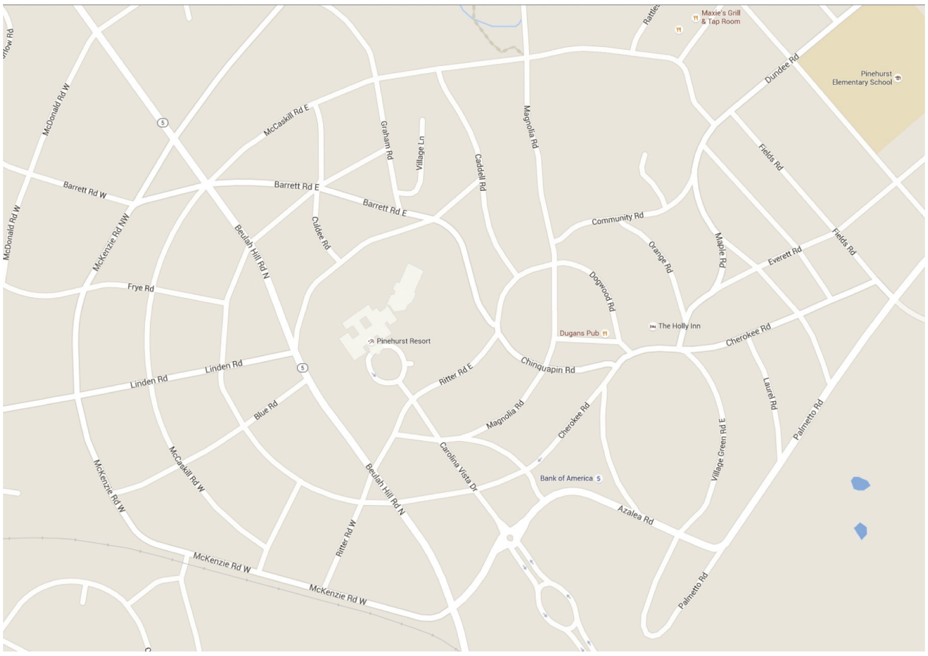


Figure 7 BEFORE – Screen shot of Google Map of Pinehurst (Map data ©2017 Google).

Highlighting arteries that optimally connect the commercial district with the surrounding regional highways yields a pattern of roads radiating out from the Holly Inn. This turns out to be appropriate, as the Holly Inn was the very first large structure in Pinehurst. Note that right in front of the Holly Inn, we use light gray as a separator between the two sets of radiating dark-gray lines to make explicit this radiating structure. The other thick light gray lines represent highly traveled shortcuts between highways as well as a circuit around the residential area west of Route 5. These act as circumferential routings with respect to the Holly Inn. Lastly, the straight, thin light-gray lines highlight the route connecting the two most important places in Pinehurst: the Carolina Inn and the Clubhouse for one of the world's most famous golf courses.

Note that there is nothing in the actual environment of the physical streets marked in the grays above, both light and dark, to distinguish them from the other original village streets; the spanning artery designations above are an artifact of the untangling process, made up to help navigation in the Village. Clearly, the Olmsteds had no intention of making navigation easy inside Pinehurst. (The straight ceremonial street connecting the Carolina Inn to the Clubhouse was a later addition, not part of the original design of curvaceous streets.)

Now imagine someone with the "after" map in Figure 8 on the screen of a smartphone, with the pulsating blue dot showing his or her position, per GPS. We would propose that such a person would have little trouble keeping track of where they are with respect to any destination, and immediately appreciate how to head to where they want to head, as compared to how they would feel with the unmodified Google map. They would have a sense of how the village "works." Zooming in and out of this map would be able to reveal far more details, within an instantly available Village-wide context.

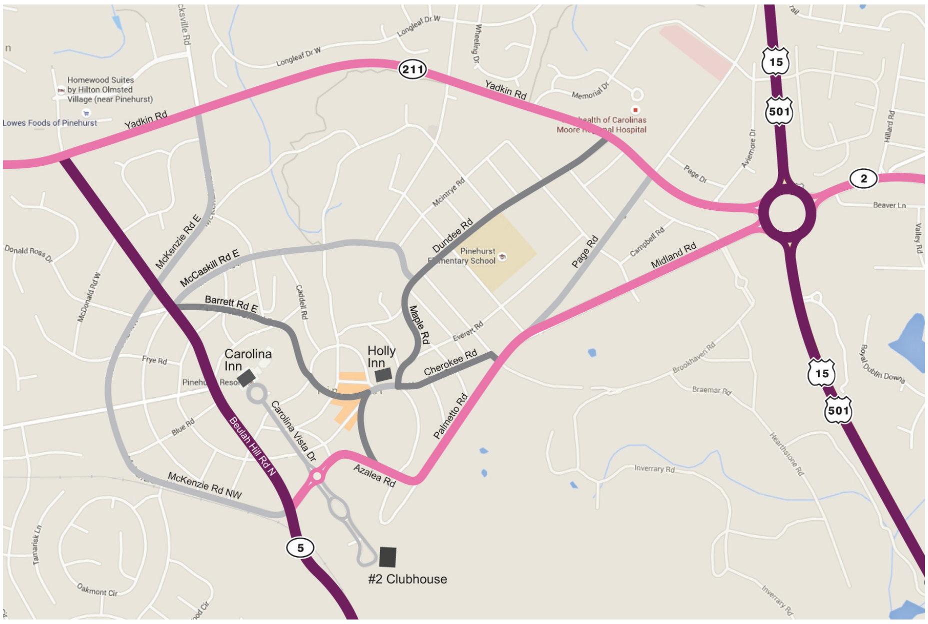


Figure 8 AFTER – untangled version of Pinehurst at smaller scale (by the author).

5 Untangled Maps: Determining Routes

In both Hirtle et al. [11] and Tomko et al. [30], the difficulties in communicating navigational instructions are explored. In particular, the latter states: "While maps – a medium to capture survey spatial knowledge – have been extensively studied in terms of spatial data quality, to our knowledge such frameworks do not exist for route spatial knowledge."

We propose that iterating UAI-untangling across a region may well provide such a mapping framework. Since in iterative UAI untangling, arterials are rendered explicitly in terms of their here-to-there function within each UAI, how does this aid the user in actual route determination? For one, there is a hierarchy of "chunking" of routes which Klippel et al [15] and others have assumed to be "crucial for . . . conceptualization of routes."

Consider a routing between a given point A and point B (Figure 9a). In general, the routing between them can be constructed in terms of UAIs as follows. There exists a UAI that contains both points and spanning elements of the arterial skeleton (for example, the Pennsylvania Turnpike running between a destination in the Pittsburgh area and one in the Philadelphia area).

There will then be a UAI that contains each of the end points that intersects with each end of the spanning artery (say, regional expressways connecting to exits at either end of the Turnpike) as illustrated in Figure 9b.

This process can be iterated to the finest-grain mesh of UAIs, with spanning arterials eventually passing close to the end points, leaving only local non-arterials to traverse to the destination (typically a matter of no more than few blocks from the nearest UAI arterial), as illustrated in Figure 9c.

At each point in the selection of a spanning arterial, the user sees an explicit set of options from which to choose that are able to be chunks of nine options or less [20, 15]. If, in a given UAI, there are more than 9 spanning arterials of a given axis type, color differences can be introduced to chunk the arterials into groups of nine or less.

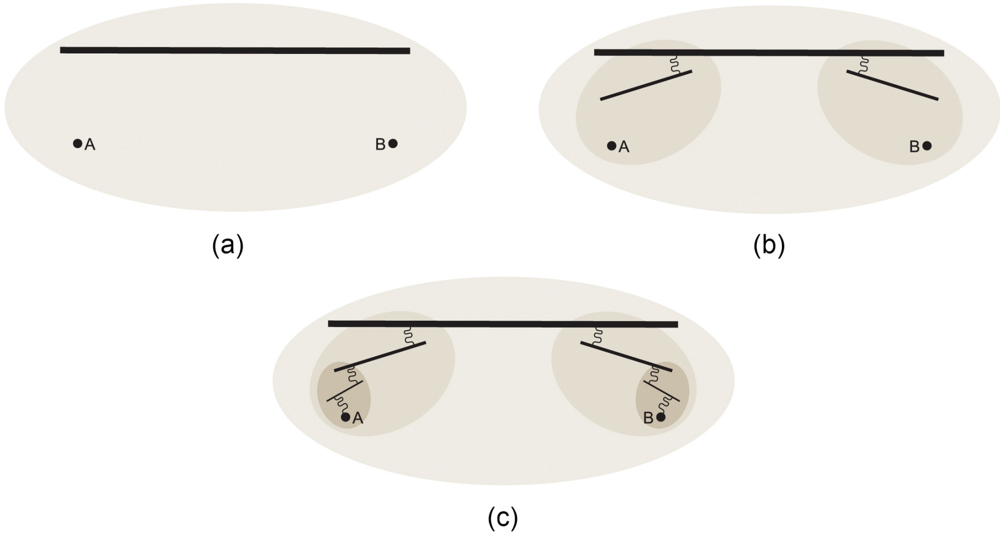


Figure 9 UAI route construction.

For example, in our polar coordinate-like interpretation of the U.S. Interstate system (Figure 2), we introduced four color groupings for the radials (ranging from hot red aiming south to ice blue aiming north). Likewise, if there are more than nine UAIs to choose from at a given level of the arterial selection process, the UAIs can be grouped in chunks of nine or fewer by introducing a hierarchy in thematic titles and area coloring.

In this way, at any point in the arterial selection process, through the organization of groupings of UAIs into a commonly understood geographical progression (e.g., state to region to neighborhood), the user faces a relatively small number of comparative choices at each level of granularity.

In our firm's work, we have applied iterative UAI untangling to more than a dozen major metro areas in the US and Canada. Over the past two years, we've also applied the process to transit maps to particularly interesting effect. For example, the official system map for the 100+ bus routes of Pittsburgh is so complex as to be largely indecipherable for most routings in the central core. In tests with our untangled version, users were able to simply "see" their way by bus with confidence, handily beating the time it takes just to enter a destination address in Google maps. (See the transit tab at `citytunr.com`, our Pittsburgh web app that is in beta, which also includes untangled "slippy maps" at four zoom levels for driving and biking.)

6 Concluding Thoughts: What Use Are Untangled Maps?

In the Theme Section Editorial of the opening issue of the *Journal of Spatial Information Science*, Tenbrink and Winter [27] discuss the difficulty the current state of automatically-generated spatial information has in being "cognitively suitable" for the user. The problem appears to boil down to integrating what is relevant to the user over granularity. On the other hand, "in spite of the complex relationship between granularity and relevance, humans typically manage to present information in an integrated and coherent way, switching flexibly and smoothly between levels of granularity according to the expected relevance for the information seeker." [27] It goes on to suggest that "research in this area can take two approaches: either an empirical approach, studying the human ability to learn about it; or an engineering approach, implementing and testing models of this capacity in spatial information systems." We propose that iterated UAI untangling as a model for capturing routing spatial knowledge is suitable for testing.

Note that we are not suggesting that simply showing untangled maps on small screens would fully replace automated systems as *the* solution. However, as Klippel et al. [16] point out, "it becomes critical to find mechanisms that preserve structurally and cognitively salient patterns to enable environmental learning and create spatial awareness." For example, we can imagine a hybrid system, in which GPS-based, turn-by-turn instructions are provided in the context of an untangled map system, perhaps most suitably on a tablet-like screen, as comes standard in a Tesla automobile, for instance. Users would be able to apprehend the untangled structure of routings as they went along, feeling oriented at each turn to both that structure and to their position relative to the desired destination. They could also decide if the "optimal" routing provided by their navigational device makes sense given the availability of nearby routings perceived amongst the untangled arteries. Quick changes to where they want to head or how they want to get there could potentially be enacted by the user without having to take time to transmit a changed destination and/or routing-preference to the GPS system. They would simply see how that new routing would work in the untangled mapping. Moreover, as the user gets near to their destination, an untangled map that makes travel to parking options transparent would provide a capability that most current systems lack.

Is there a route-determining application of untangled maps that would be better without any routing-automation beyond fine-grain parking finding? Untangled bus-transit maps may be an example. Automatic transit directions have a checkered history of directing users on convoluted itineraries when a simple one is possible [10] – if a user could determine a transit routing with just a glance at an untangled map, and at the same time attain a mastery of how transit works in a city, an automatic app might seem a fussy (and unreliable) bother by comparison.

There is another type of application we have in mind as well. In the context of the supposed coming age of the self-driving car, we heard the chief engineer for Google's effort on the "60 Minutes" TV program bragging that all someone will need to do is "plug in their destination and go." [32] That's fine if all you want to do is visit your Aunt Martha, but what if you are new to an area and want to explore the possibilities? Most automated map systems feel cumbersome at best in their indices and/or search strategies for displaying destination options – they certainly do not convey a sense of how a city "works." It would seem, then, that a map interface that could capture such spatial knowledge of a city and empower users' choices based on multi-granular spatial awareness of destination options could be a powerful application for such vehicles, an interface that potentially enables the user to "know the city" better and faster than existing systems.

Does the untangled map system truly work fast enough to be cognitively suitable for the above semi-automated and non-automated applications? When is such navigation faster and more satisfying an experience for the user than with automated systems [1]? When is the flexibility and resilience of navigating with the aid of untangled survey knowledge a clear advantage? As of yet, there has not been rigorous testing of this model; we invite researchers to test the cognitive suitability of untangled maps.

Acknowledgements. The author wishes to thank Stephen Hirtle for his patience in reading an early draft of the paper and for providing valuable comments. Also appreciated were the comments of the anonymous referees, one of whom helpfully pointed the author to Jiang and the concept of "natural roads."

References

1 Barry Brown and Eric Laurier. The normal natural troubles of driving with GPS. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1621–1630. ACM, 2012.

2 William G. Chase. Spatial representations of taxi drivers. In *The acquisition of symbolic skills*, pages 391–405. New York: Plenum, 1982.

3 Robert Firth. Comprehensive Interim Report on Pittsburgh's Visitor-Path Wayfinding Sign System. Delivered to the Carnegie Institute, Pittsburgh, PA, 1994.

4 Robert Firth. *Finding Yourself in Pittsburgh*. Pittsburgh: Informing Design, Inc., 1997.

5 Robert Firth. Maryland's Statewide Tourist Area and Corridor Signing Program. Maryland State Highway Administration, 2014. `http://www.roads.maryland.gov-/OOTS/TAC_Brochure.pdf`.

6 Christian Freksa. Strong spatial cognition. In S. Fabrikant, M. Raubal, M. Bertolotto, C. Davies, S. Freundschuh, and S. Bell, editors, *Spatial Information Theory (COSIT 2015)*, volume 9368 of *Lecture Notes in Computer Science*, 2015.

7 David Gibson. *The Wayfinding Handbook*. Pinceton Architectual Press, New York, 2009.

8 Reginald G. Golledge. *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU press, 1999.

9 Stephen Hirtle. Dinosaurs, Slide Rules and Maps. Presentation to Cognitive and Linguistics Aspects of Space, 2010(a).

10 Stephen Hirtle. *Geographical design: Spatial cognition and geographical information science*. Morgan & Claypool Publishers, 2011.

11 Stephen Hirtle, Kai-Florian Richter, Samvith Srinivas, and Robert Firth. This is the tricky part: When directions become difficult. *Journal of Spatial Information Science*, 2010(1):53–73, 2010.

12 Toru Ishikawa, Hiromichi Fujiwara, Osamu Imai, and Atsuyuki Okabe. Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. *Journal of Environmental Psychology*, 28(1):74–82, 2008. `doi:10.1016/j.jenvp.2007.09.00239`.

13 Bin Jiang. Street hierarchies: a minority of streets account for a majority of traffic flow. *International Journal of Geographical Information Science*, 23(8):1033–1048, 2009.

14 Bin Jiang, Sijian Zhao, and Junjun Yin. Self-organized natural roads for predicting traffic flow: a sensitivity study. *Journal of statistical mechanics: Theory and experiment*, 2008(07):P07008, 2008.

15 Alexander Klippel. Wayfinding choremes: Conceptualizing wayfinding and route direction elements. *Spatial Information Theory*, 1, 2003.

16 Alexander Klippel, Stephen Hirtle, and Clare Davies. You-are-here maps: Creating spatial awareness through map-like representations. *Spatial Cognition & Computation*, 10(2-3):83–93, 2010.

17 Benjamin Kuipers. Modeling spatial knowledge. *Cognitive science*, 2(2):129–153, 1978.

18 Benjamin Kuipers, Dan G Tecuci, and Brian J Stankiewicz. The skeleton in the cognitive map: A computational and empirical exploration. *Environment and behavior*, 35(1):81–106, 2003.

19 Kevin Lynch. *The image of the city*. Cambridge: MIT press, 1960.

20 George A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

21 Daniel R Montello. Spatial orientation and the angularity of urban routes: A field study. *Environment and Behavior*, 23(1):47–69, 1991.

22 Daniel R. Montello. Cognitive research in GIScience: Recent achievements and future prospects. *A research agenda for geographic information science*, pages 61–91, 2009.

23 Avi Parush, Shir Ahuvia, and Ido Erev. Degradation in spatial knowledge acquisition when using automatic navigation systems. In *International Conference on Spatial Information Theory*, pages 238–254. Springer, 2007.

24 Romedi Passini. *Wayfinding in architecture*. New York: Van Nostrand Reinhold Co. Limited., 1984.

25 Falko Schmid, Kai-Florian Richter, and Denise Peters. Route aware maps: Multigranular wayfinding assistance. *Spatial Cognition & Computation*, 10(2-3):184–206, 2010. doi:10.1080/13875861003592748.

26 Robert A. M. Stern, David Fishman, and Jacob Tilove. *Paradise planned: the garden suburb and the modern city*. The Monacelli Press, 2013.

27 Thora Tenbrink and Stephan Winter. Presenting spatial information: Granularity, relevance, and integration. *Journal of Spatial Information Science*, 2010(1):49–51, 2010.

28 Robert Thomson and Rupert Brooks. Exploiting Perceptual Grouping for Map Analysis, Understanding and Generalization: The Case of Road and River Networks. *Graphics Recognition Algorithms and Applications*, pages 148–157, 2002.

29 Sabine Timpf, Gary Volta, David Pollock, and Max Egenhofer. A conceptual model of wayfinding using multiple levels of abstraction. *Theories and methods of spatio-temporal reasoning in geographic space*, pages 348–367, 1992.

30 Martin Tomko and Kai-Florian Richter. Defensive wayfinding: Incongruent information in route following. In *International Workshop on Spatial Information Theory*, pages 426–446. Springer, 2015.

31 Martin Tomko, Stephan Winter, and Christophe Claramunt. Experiential hierarchies of streets. *Computers, Environment and Urban Systems*, 32(1):41–52, 2008.

32 Bill Whitaker. Hands off the wheel. Retrieved January 16, 2017, from. http://www.cbsnews.com/news/self-driving-cars-google-mercedes-benz-60-minutes/.

33 Alexander Zipf and Kai-Florian Richter. Using focus maps to ease map reading – developing smart applications for mobile devices. *Künstliche Intelligenz. Sonderheft Spatial Cognition*, 4(02):35–37, 2002.

Uncertainty in Wayfinding: A Conceptual Framework and Agent-Based Model*

David Jonietz[1] and Peter Kiefer[2]

1 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
jonietzd@ethz.ch

2 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
pekiefer@ethz.ch

Abstract

Though the wayfinding process is inherently uncertain, most models of wayfinding do not offer sufficient possibilities for modeling uncertainty. Such modeling approaches, however, are required to engineer assistance systems that recognize, predict, and react to a wayfinder's uncertainty. This paper introduces a conceptual framework for modeling uncertainty in wayfinding. It is supposed that uncertainty when following route instructions in wayfinding is caused by non-deterministic spatial reference system transformations. The uncertainty experienced by a wayfinder varies over time and depends on how well wayfinding instructions fit with the environment. The conceptual framework includes individual differences regarding wayfinding skills and regarding uncertainty tolerance. It is implemented as an agent-based model, based on the belief-desire-intention (BDI) framework. The feasibility of the approach is demonstrated with agent-based simulations.

1998 ACM Subject Classification I.6.m [Simulation and Modeling] Miscellaneous

Keywords and phrases Wayfinding, Uncertainty, Agent-Based Model

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.15

1 Motivation

Wayfinding – the 'goal-directed and planned movement of one's body around an environment in an efficient way' [20] – can be modeled as a sequence of wayfinding decision situations [6] in which the wayfinder chooses from a set of possible paths. These decisions are made under uncertainty, where the degree of uncertainty depends on the situation: for instance, uncertainty will be higher while performing uninformed search [34] in a foreign city than while finding the way to one's regular workplace.

Navigation aids may help alleviate uncertainty in wayfinding. Different types of navigation aids, and different approaches for the generation and communication of wayfinding instructions through digital wayfinding assistants, have been considered over the years, including turn-by-turn instructions [13], you-are-here maps [12], digital 2D and 3D maps [14], adaptive signage [15], or haptic interfaces [28]. A particularly well-studied topic is the automated creation of route instructions based on landmarks [25]. For instance, choosing landmarks based on their saliency [24] aims at reducing uncertainty by helping the wayfinder match the instruction to the environment.

* Author names in alphabetical order.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 15; pp. 15:1–15:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

While these approaches may alleviate uncertainty to some degree, a significant amount of uncertainty often remains: the user might be lacking the (spatial) abilities for interpreting the information [1], there may have been problems in the communication process [32], incongruent information that does not match the environment [29], or context factors the navigation assistant did not adapt to [23]. It is unlikely that it will be ever possible to completely erase uncertainty in wayfinding, which underlines the importance of taking the user's uncertainty into account for the design of wayfinding assistants.

A holistic understanding of uncertainty in wayfinding implies also an understanding of wayfinders' reactions to uncertainty. Brunyé et al., for instance, have found empirical evidence that the type of information source (human vs. GPS device) influences the decision made in situations of uncertainty [2]. Tomko and Richter have suggested that, under conditions of uncertainty, a wayfinder will eventually enter a particular wayfinding mode (*defensive wayfinding*) in which she is aware of a mismatch between instruction, expectations and environment, thus proceeding cautiously and investing excessive mental effort for correcting the mismatch [29]. This perspective on uncertainty in wayfinding is particularly interesting because it transcends the notion of situative uncertainty (*which factors influence uncertainty at decision point p?*) to a more process-oriented view on uncertainty (*how does uncertainty influence the wayfinder's cognitive processes over time?*). The 'defensive wayfinding' model [29], however, remains conceptual and largely informal.

Overall, we note that, while wayfinding literature has touched upon and discussed uncertainty from different perspectives, a computational model of uncertainty in wayfinding which would allow wayfinding assistants to have an explicit notion of and take their user's uncertainty into account is still missing.

Here, we take important steps towards such model: we introduce a conceptual framework which allows to model uncertainty in wayfinding as a result of non-deterministic reference system transformations (building on ideas from [11]). The conceptual framework enables to include all three aspects of a wayfinding situation into a model: the wayfinder, the instruction, and the environment. The presented framework covers both, a situational and a process-oriented view on uncertainty in wayfinding, and allows to capture individual differences (since people have different dispositions regarding their ability to deal with uncertainty [5]).

Based on this general framework, we further develop an agent-based model (ABM) of landmark-based wayfinding under uncertainty in an unfamiliar environment. In the past, numerous studies have aimed at developing artificial agents with navigation capabilities, including more general cognitively inspired computational frameworks, such as TOUR [17], but also practical applications from the ABM community (e.g. [22, 9]). More closely related to our work are studies on software agents which comprehend and follow route instructions, e.g., [18, 30], who present agents capable of following natural language route instructions, [33], whose probabilistic agent interprets ambiguous direction-based route instructions in real-world path networks, or [7], who use an agent for evaluating the reliability of a complexity-reducing route computation algorithm.

To the best of our knowledge, however, there has so far been no ABM which is based on a comprehensive concept of how uncertainty is created from the interplay of agent, environment and instruction, or its subsequent effects on the behavior of the wayfinder agent. These insights, however, are needed for developing and testing uncertainty-aware assistance systems. In this paper, thus, we demonstrate on an exemplary implementation of our model its capabilities for simulating wayfinding situations of differing complexity.

The paper is structured as follows: Section 2 explains prior work on reference system transformations in wayfinding [11]. Section 3 introduces our conceptual framework of uncertainty in wayfinding which provides the basis for the development of an ABM of wayfinding under uncertainty (see Section 4). An exemplary implementation of the ABM and simulation results are presented in Section 5, before Section 6 concludes this paper.

2 Modeling Wayfinding With Reference System Transformations

This section shortly reflects on the role of spatial reference system transformations in wayfinding (refer to [11, section 2.1] for details). Spatial reference systems are a core concept of spatial information theory as they are used to encode (externally or internally), reason about, and communicate about locations by both, humans and machines [4, 16, 27]. Three types of reference systems are relevant for wayfinding: *egocentric* (self-to-object), *allocentric* (object-to-object), and *survey reference systems* (relative to the earth's surface or other ground phenomena).

Egocentric reference systems are aligned with the body and used by the wayfinder to refer to locations in vista space [19], such as *'left of that restaurant over there'*. Such egocentric locations are the output of a wayfinding decision and serve as input for locomotion. Allocentric reference systems enable the reference to locations across individuals and independent from the current point of view. They can therefore be used for representing instructions, such as '*(take the road) in front of the restaurant'*. A reference system transformation (from allocentric to egocentric) is necessary in order to match these instructions with the current field of view. Survey reference systems, such as cartographic maps or mental spatial representations ('cognitive maps' [31]), represent locations with their relation to other locations and are therefore used for route planning. The authors of [11] modeled the output of the route planning process as a sequence of (allocentric) instructions. Route planning therefore requires a transformation from survey to allocentric locations (e.g., *'North of the restaurant symbol'* → *'in front of the restaurant'*).

The model in [11] was intentionally left underspecified in several aspects. For instance, it did not describe how the wayfinder behaves if a reference system transformation is not deterministic, which happens often in realistic use cases. This paper focusses on the uncertainty caused by non-deterministic reference system transformations, thus building on the core ideas discussed in [11].

3 Conceptual Framework

Here, uncertainty is considered as a wayfinder's lack of knowledge about relevant aspects of the wayfinding situation, which is defined by the interplay of the environment (Sections 3.1 and 3.2), the wayfinder (Section 3.3), and the instruction (Section 3.4) [6]. In this section, we focus on route following and develop a conceptual framework for modeling uncertainty along these three dimensions.

3.1 Uncertainty and Spatial Reference Systems in Wayfinding

For a wayfinder standing at a decision point, the most fundamental form of uncertainty relates to the question which path to take next (*path choice uncertainty*). Here, we argue that path choice uncertainty is caused by other types of uncertainty which originate from the cognitive sub-processes involved in wayfinding. For instance, the wayfinder may be uncertain about whether she is still on the correct route, whether the object she perceives in front of her matches the landmark refered to in a route instruction, or where she would locate her current position on a map. Following the idea of [11] (see also Section 2) of reference system transformations as a core concept for a model of wayfinding, we here suggest that

▶ **Supposition 1.** *Uncertainty in wayfinding is always related to a spatial reference system.*

Thus, uncertainty can be present for each of the three types of spatial reference systems relevant for wayfinding:

- *Egocentric uncertainty*: where in my egocentric view is location L_e? An important egocentric uncertainty is path choice uncertainty: where in my egocentric view is the path I should take?
- *Allocentric uncertainty*: where is L_{a1} located, relatively to L_{a2}? In path following [34], an important allocentric uncertainty is the *on-route uncertainty*: is my current location on the route I was planning to follow?
- *Survey uncertainty*: where is L_s in a given survey reference system, e.g., where is the wayfinder's location on a map?

Note that in our concept uncertainty in different spatial reference systems may have different degrees at the same time. For instance, a wayfinder could be very certain about her position on a map, but at the same time very uncertain about which path to take (and vice versa).

3.2 Non-Deterministic Reference System Transformations in Wayfinding

A wayfinder needs to transform information between reference systems in order to successfully solve the wayfinding problem (see Section 2 and [11]). The transformation processes are non-deterministic, which causes uncertainty:

▶ **Supposition 2.** *Uncertainty in wayfinding can be modelled as being caused by non-deterministic spatial reference system transformations.*

For instance, suppose that in an instruction-based wayfinding situation, the allocentric instruction *'(take the road) in front of the restaurant'* can be transformed to three egocentric locations – the entries to three alternative roads – as follows: ($L_{E,1}$, 0.1), ($L_{E,2}$, 0.2), ($L_{E,3}$, 0.15). Numbers describe how well these locations fit to the instruction (*fit distribution*). In this example, all three location options have rather low fit values[1], meaning that the allocentric location cannot be mapped well (e.g., none of the egocentric objects can be clearly recognized as a restaurant).

We assume that this kind of situation will increase the wayfinder's allocentric uncertainty: *'is the instruction really meant for my current location? Did I go wrong in one of my previous decisions?'*[2]. On the other hand, a high maximum fit will re-assure the wayfinder about being on route, even though she might have felt uncertain before (= decrease allocentric uncertainty). Similarly, while transforming from allocentric to a survey system, a low maximum fit may indicate that the survey system does not contain a correspondent for the allocentric location (e.g., the map shows a different area), which would increase uncertainty. These examples lead us to:

[1] Without loss of generality, we here assume a scale [0..1] for *fit*. Note that a fit distribution is not a probability distribution, i.e., fits in a particular decision situation do not sum up to 1.

[2] [29] would characterize situations with low maximum fit as having a high *detectability* of mismatch between instruction and environment.

▶ **Supposition 3.** *A non-deterministic spatial reference system transformation will increase uncertainty if the maximum fit value is low, and decrease uncertainty if the maximum fit value is high.*

Ambiguity has been identified as an important factor determining the success of wayfinding [29]. With the fit distribution introduced above, we can easily define:

▶ **Supposition 4.** *A non-deterministic spatial reference system transformation will increase uncertainty if the ambiguity of the fit distribution is high.*

Ambiguity occurs if the maximum fit value is close to the fit values of other options[3]. For instance, the instruction *'in front of the restaurant'* could have an equal fit to two options if there is more than one restaurant. Note that, in our model, ambiguity does not require the maximum fit to be particularly high. The fit distribution given in the example above (0.1, 0.2, 0.15) would be ambiguous, because this kind of distribution makes it hard for the wayfinder to decide between the three options.

3.3 Coping Strategies, Individual Differences

With the suppositions proposed so far, we have modeled *uncertainty changes over time*, which includes both, increase and decrease of uncertainty. As a next step, we look at wayfinders' (individual) reactions to uncertainty. Similar to the concept of defensive wayfinding introduced in [29], we include a wayfinder's reactions to uncertainty as follows:

▶ **Supposition 5.** *If uncertainty in a spatial reference system reaches an uncertainty threshold, the wayfinder applies one or several* coping strategies *in order to reduce uncertainty in that particular reference system. The threshold is determined by the wayfinder's characteristic* uncertainty tolerance*.*

Uncertainty tolerance describes the wayfinder's tendency to continue non-defensive wayfinding (i.e., without coping strategy) in situations of uncertainty. It is motivated by psychological research which has found that humans have a disposition with regards to their behavior in uncertain situations [5]. Here, we assume that uncertainty tolerance is influenced by at least three factors: 1) an individual disposition, 2) the wayfinder's self-estimation of her wayfinding skills (see below), 3) the impacts of getting lost for the given task context (e.g., arriving late for a dinner appointment vs. missing a plane).

Coping strategies may reduce uncertainty but require time. A wayfinder who performs coping strategies frequently will have fewer uncertain situations, leading to fewer errors, higher likelihood of reaching the destination (high effectiveness), but need more time (low efficiency). Examples for coping strategies include increasing visual monitoring of the environment, asking an instructor to point to the direction one needs to take, asking a local person to help with disambiguation of landmarks, or performing self-localization on a you-are-here map.

A second individual difference, besides uncertainty tolerance, is certainly determined by the agent's *wayfinding skills* (e.g., [8])

▶ **Supposition 6.** *The lower the wayfinding skills are, the more will the fit distribution estimated by a wayfinder during a spatial reference system transformation deviate from a ground-truth fit distribution.*

[3] We introduce one possible formula for ambiguity in Section 4.

While uncertainty tolerance is defined w.r.t. one particular reference system, the wayfinding skill is related to a type of reference system transformation. For instance, some wayfinder may have high wayfinding skills for matching landmark-based instructions to the environment (allocentric to egocentric), but low map reading skills (survey to allocentric).

3.4 Generating Uncertainty-Aware Route Instructions

The algorithms listed in Table 1 specify three possible strategies for selecting an instruction which describes the destination edge e_{dest} when arriving from edge e_{orig} at a decision point *dp*. The algorithms are executed for each decision point along a previously created route, such as the shortest or the simplest alternative [3].

Algorithm 1 generates a finite set of possible instructions and returns the one with the highest fit. Since Algorithm 1 does not take ambiguity into account, there could exist a different edge for which the same instruction has an even better fit. For instance, the approach proposed by [24] generates landmark-based instructions based on the saliency of landmark candidates (from visual, semantic, and structural attraction), but does not consider whether there are other landmark candidates of the same type at the decision point.

Algorithm 2 resolves this by comparing the fit of instruction and target edge to the fit of that instruction to all other edges. If there is a non-target edge for which the instruction fits better, the instruction is discarded and one with lower fit is tested the same way. It may happen that no instruction can be generated if all potential instructions for the target edge have a better fit to some other edge. In that case, the calling route generation algorithm would need to find a route through a different decision point.

Similarly, **Algorithm 3** first generates all instructions for the target edge and then discards all those which have a better fit somewhere else. For each of the remaining instructions the ambiguity of the instruction is computed and combined with target edge fit. The one which maximizes the combined score is returned. The rationale behind Algorithm 3 is that a wayfinder with low wayfinding skills might perceive a fit distribution which deviates from the distribution assumed by the algorithm (see Supposition 6).

4 An Agent-Based Model of Wayfinding Under Uncertainty

In this section, we describe an agent-based model of wayfinding under uncertainty. The suppositions introduced in the conceptual framework (Section 3) are here operationalized for the particular case of wayfinding with landmark-based instructions in an unknown environment (i.e., we exclude the level of survey knowledge here). It is further assumed that the level of uncertainty changes only at decision points (see discussion in Section 6).

4.1 The Environment

The environment is modeled as a directed weighted graph with decision points (DP) as nodes, and paths as edges (E). Each decision point features a (finite) set of objects (O_{dp}), which serve as potential landmarks in wayfinding instructions. A spatial configuration attribute ($config_{dp}$) describes for each *dp* the allocentric spatial relations between its adjacent path edges and inherent objects, based on a *fit distribution* over the available tuples of edge, object, and relation ($config_{dp} \subseteq E \times O_{dp} \times Rel \times [0..1]$). For instance, $(e_3, o_7, rel_x, 0.9) \in config_{dp10}$ would specify that, at decision point *dp10*, the edge e_3 can be described as being located in relation rel_x to object o_7 with a *relfit* of 0.9. Note that we remain on a high level of

Table 1 Three instruction generation algorithms for wayfinding under uncertainty (pseudo code). The algorithms assume: (a) a function which generates a finite list of possible instructions for describing a destination edge at a decision point, (b) a *fit*() function which computes how well an instruction fits to an edge, (c) a function *amb*() which computes the ambiguity of an instruction at a decision point, (d) a *combinedscore*() function which averages fit and ambiguity.

```
Algorithm 1 (local fit optimization)
in: DecisionPoint dp, Edge e_orig, Edge e_dest, out: Instruction
I ← Generate possible instructions for e_dest.
fit(i, e_dest) ← Calculate for each instruction i ∈ I the fit to e_dest.
Return i ∈ I, which maximizes fit(i, e_dest).
```

```
Algorithm 2 (max fit optimization)
in: DecisionPoint dp, Edge e_orig, Edge e_dest, out: Instruction
I ← Generate possible instructions for e_dest.
fit(i, e_dest) ← Calculate for each instruction i ∈ I the fit to e_dest.
I' ← Sort I by fit(i, e_dest) in descending order.
for each i' ∈ I'
    fit(i', e') ← Calculate the fit for i' to each e' ∈ dp.getEdges() \ {e_orig, e_dest}
    if not exists e' with fit(i', e') > fit(i', e_dest)
        return i'
return Could_not_create_instruction.
```

```
Algorithm 3 (combined fit and ambiguity optimization)
in: DecisionPoint dp, Edge e_orig, Edge e_dest, out: Instruction
I ← Generate possible instructions for e_dest.
fit(i, e_dest) ← Calculate for each instruction i ∈ I the fit to e_dest.
for each i ∈ I
    fit(i, e) ← Calculate the fit for i to each e ∈ dp.getEdges() \ {e_orig, e_dest}
    if exists e with fit(i, e) > fit(i, e_dest)
        I.remove(i)
amb(i, dp) ← Calculate for each instruction i ∈ I the ambiguity at dp.
Return i ∈ I, which maximizes combinedscore(fit(i, e_dest), amb(i, dp)).
```

abstraction here: we neither model the location of objects in coordinate space, nor explicit relations, such as ‘in front of’.

Here, in a wayfinding instruction, the corresponding object instance is referred to by an object type (from a finite set T), such as ‘a restaurant’ or ‘a blue house’. Object instances can have more than one type. Function $typefit_{dp} : (O_{dp} \times T) \rightarrow [0..1]$ describes how well an object fits to a certain type. The *fit distribution* over location candidates of the conceptual framework (refer to Section 3.2), therefore, results from a combination of the according *typefit* and *relfit* values.

4.2 The Agent

The cognitive architecture of our wayfinder agent is based on the widely-used belief-desire-intention (BDI) framework [21]. The agent’s behavior is primarily motivated by a desire (reaching the destination node), its knowledge base is represented as separate beliefs, and its intentions refer to and can trigger particular behavior. As explained in Section 3.3, agents

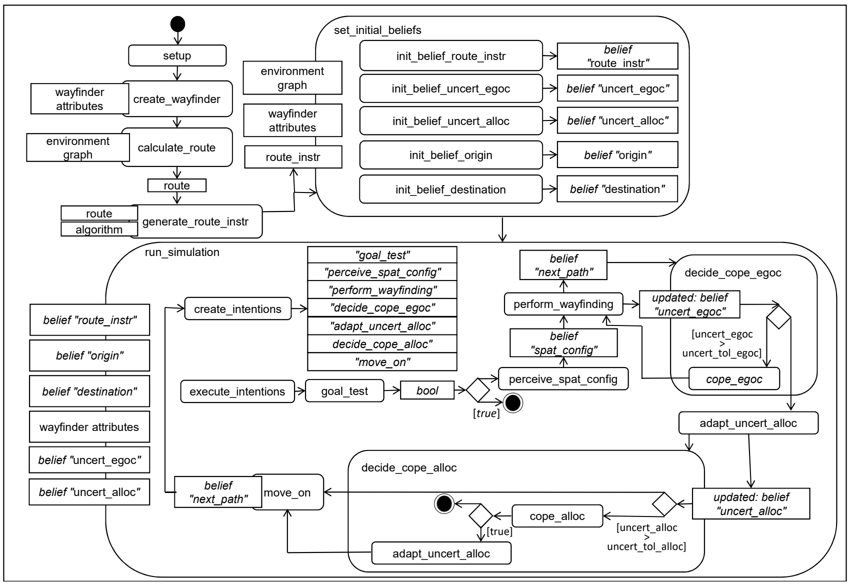


Figure 1 Overview: an agent-based model of wayfinding under uncertainty.

have individual differences w.r.t. wayfinding skills (Supposition 6) and uncertainty tolerance (Supposition 5). Wayfinding skills are here further distinguished between the skill to accurately assess the spatial relations (*relfit* values) between pairs of object and edge (*wayf_rel_skill*), and the skill to recognize the type (*typefit* values) of an object (*wayf_type_skill*).

Regarding uncertainty tolerance, we distinguish between uncertainty tolerance on egocentric and on allocentric level. The first, *uncert_tol_egoc*, reflects the tolerance regarding path choice uncertainty. A wayfinder with low *uncert_tol_egoc*, for instance, will have a high desire to make sure she makes the correct decision, which means she will often apply an egocentric coping strategy (e.g., ask somebody whether that building over there is of type t). The latter, *uncert_tol_alloc*, reflects the tolerance regarding on-route uncertainty. For instance, a wayfinder with low *uncert_tol_alloc* has a high desire to always know whether she is still on the route.

4.3 The Wayfinding Process

4.3.1 Initialization, goal test and perception

Figure 1 illustrates our model of the wayfinding process. After an environment graph has been built up in a setup procedure, a wayfinder agent is created and positioned at the start node of the route. Then, the shortest route from the start to the destination node is calculated, and, using one of the algorithms described in Section 3.4, translated into a route description, i.e., an ordered sequence $route\_instr = [i_1, i_2, ..., i_j]$ where each instruction i consists of one relation and one object type, $i_k \in Rel \times T$. While the actual route is unknown to the wayfinder agent, the *route_instr*, among other initial settings, is stored in its belief structure (*set_initial_beliefs*).

After setup and initialization, the wayfinding process starts with the creation of a stack of intentions (*create_intentions*). The wayfinder agent processes the stack incrementally, triggering the corresponding behavioral procedures (*execute_intentions*). It begins by checking whether the destination node has been reached (*goaltest*). If so, the wayfinding has been successful and the simulation stops; if not, the agent perceives the $config_{dp}$ of the dp it is currently located at and stores it in its belief structure. In order to account for different wayfinding skills (see Supposition 6), the true $config_{dp}$ is not directly perceived, but distorted by randomly altering *relfit* and *typefit* values. The probability distribution for the magnitude of this error is determined by the agent's *wayf_rel_skill* and *wayf_type_skill* attributes (by setting the standard deviation of a normal error distribution with a mean of 0).

4.3.2 Instruction matching, ambiguity and uncertainty

In the next procedure (*perform_wayfinding*), the agent interprets the route instruction w.r.t. the perceived $config_{dp}$: given instruction $i = (rel, t)$ at decision point dp, it identifies the adjacent edge e_{max} for which i fits best and stores it as a *next_path* belief.

As described in Suppositions 2 and 4, however, the agent can be more or less uncertain about the correctness of this decision, depending on the closeness of the fit values of the alternative edges (ambiguity). We here use the following formula for quantifying ambiguity:

$$ambiguity_{dp} = maxfit_{dp} - \frac{\sum_{i=1}^{n} typefit_i * relfit_i}{n}$$

where $maxfit_{dp} = max(typefit * relfit)_{dp}$ and ambiguity is hence calculated as the difference between the maximum possible product of *typefit* and *relfit* at dp and the mean of the product of the same (rel, t) for all alternative edges.

The resulting ambiguity value is taken as perceived egocentric uncertainty of choosing a path at this dp, and is stored as an *uncert_egoc* belief. In a following *decide_cope_egoc* procedure, the agent tests whether *uncert_egoc* exceeds the *uncert_tol_egoc* attribute, as described in Suppositions 5 and 6, and would therefore require a coping strategy. In this case, a *cope_egoc* procedure is triggered, which reduces uncertainty by providing the agent with the true $config_{dp}$ with undistorted fit values. Please note that other forms of *cope_egoc* strategies would also be thinkable here.

Based on the updated fit values, the agent then repeats *perform_wayfinding*, and adapts the uncertainty on the allocentric level (*adapt_uncert_alloc*). This step consists of two sub-processes: first, the influence of the $maxfit_{dp}$ value is assessed. The following formula is based on the assumption that $uncert\_alloc'$ is set to 0.0 for $maxfit_{dp} = 1.0$, set to 1.0 for $maxfit_{dp} = 0.0$, and does not change for $maxfit_{dp} = 0.5$ (with linear interpolation between):

$$uncert\_alloc' = \begin{cases} maxfit_{dp} \leq 0.5: & (2 * uncert\_alloc - 2) * maxfit_{dp} + 1\,, \\ maxfit_{dp} > 0.5: & -2 * uncert\_alloc * maxfit_{dp} + 2 * uncert\_alloc\,. \end{cases}$$

As a second sub-process, the potential increase in uncertainty is taken into account which results from a product of the agent's estimation of how likely she is located at the correct decision point $(1 - uncert\_alloc')$ and the likelihood of making a correct decision with regards to which path to follow from the dp.

$$uncert\_alloc'' = 1 - ((1 - uncert\_alloc') * (\frac{maxfit_{dp}}{\sum_{i=1}^{n} typefit_i * relfit_i}))\,.$$

The belief *uncert_alloc* is updated with the new value *uncert_alloc*" and then compared to the agent's individual *uncert_tol_alloc* attribute to check whether a *cope_alloc* strategy must be triggered (see Supposition 5). This procedure simply provides feedback to the agent whether its location is still on the intended route or not. If *cope_alloc* returns false, the simulation stops and the agent is lost. If *cope_alloc* returns true, *uncert_alloc* is set to 0. Finally, the agent moves along the *next_path* identified previously to the next dp.

5 Simulation Experiment

In this section, we describe results of a simulation with an implementation of our ABM in the agent-based simulation environment NetLogo (`http://ccl.northwestern.edu/netlogo/`). The simulation mainly serves as a validation of the feasibility of the modeling approach. We are particularly interested in answering two questions: (1) is our model capable of simulating wayfinding situations of differing complexity?, and (2) can we model individual differences among wayfinder agents?

5.1 Implementation and Parameter Settings

In the implementation we modelled wayfinders, decision points, and environmental objects as separate classes of agents, and connected the decision points with undirected links to receive a path network. For our BDI-structure, we borrowed from an implementation done by [26]. Random graph networks with 70-80 decision points were created. In order to maintain realistic structural network properties, we enforced small-world structures with average node degrees between 2-4.75, which roughly corresponds to the characteristics which [10] identified for real-world urban networks.

A total of 24.576 simulation runs were executed, resulting from the following systematic variation of parameters (256 agents × 3 instruction generation algorithms × 32 environments):

- **Agent**: different combinations of skill and uncertainty tolerance levels: *wayf_rel_skill*, *wayf_type_skill*, *uncert_tol_egoc*, *uncert_tol_alloc*, each $\in \{0.00, 0.33, 0.66, 1.00\}$
- **Environment**: 2 random environments for each of 16 different complexities by varying: *mean_main_rel_distr* and *mean_main_type_distr* (each $\in \{0.5, 0.75\}$), *mean_minor _rel_distr* and *mean_minor_type_distr* (each $\in \{0.25, 0.5\}$). These parameters determined the mean of normal distributions for the generation of *typefit* and *relfit* distributions. We assumed the existence of 10 abstract types of objects, and assigned up to two of them as main types to object instances, which would likely receive higher *typefit* values than the other, minor types. The same was done for the *relfit* values of combinations of edge and object. Thus, lower mean values for main, and higher mean values for minor *typefit* and *relfit* have a higher probability to increase ambiguity and decrease $maxfit_{dp}$ at decision points.
- **Instruction generation algorithm**: Algorithm 1, 2, 3 (see Section 3.4, Table 1).

5.2 Simulation Test

We analyse the results in terms of effectiveness (Has the agent successfully reached the destination?) and efficiency (How often have those agents which reached the destination performed a coping strategy?).

For the analysis of *effectiveness*, we aggregate results regarding the *challenge of the wayfinding situation*, which is determined by the agent's wayfinding skills and the ambiguity of the environment as follows:

- *High challenge*: within the upper half of the total average ambiguity distribution along the route, and the lower half of the total average $maxfit_{dp}$ distribution along the route, and low wayfinding skills (*wayf_rel_skill* and *wayf_type_skill* <= 0.33)
- *Low challenge*: within the lower half of the total average ambiguity distribution along the route, and the upper half of the total average $maxfit_{dp}$ distribution along the route, and high wayfinding skills (*wayf_rel_skill* and *wayf_type_skill* >= 0.66)

Table 2 Wayfinding success rates for combinations of: uncertainty tolerance (lines), challenge of the decision situation (columns) and instruction generation algorithm (A1, A2, A3, see Table 1).

	high challenge			low challenge		
	A1	A2	A3	A1	A2	A3
high uncertainty tolerance	0.055	**0.063**	0.046	0.142	0.186	**0.228**
low uncertainty tolerance	0.159	**0.497**	0.426	0.191	**0.519**	0.475

These were combined with high (*uncert_tol_egoc* and *uncert_tol_alloc* >= 0.66) or low (*uncert_tol_egoc* and *uncert_tol_alloc* <= 0.33) uncertainty tolerance values, yielding in the results listed in Table 2.

The table shows the normalized success rates (successful runs / total runs) and the best performing route instruction algorithm. It can be seen that the success rates differ to a great degree, high challenge generally leading to lower success rates, especially in the case of a high uncertainty tolerance (agent refrains from using coping strategies). With regards to the performance of the different route instruction algorithms, it can be seen that in most cases, the more elaborate algorithms clearly outperform simple Algorithm 1. The differences between Algorithm 2 and 3, however, are less clear. Especially in case of low uncertainty tolerance, the potentially higher $maxfit_{dp}$ provided by Algorithm 2 might represent the better choice, whereas the ambiguity-reducing strategy of Algorithm 3 can provide an advantage for agents with a high uncertainty tolerance in relatively low challenging environments.

With regards to *efficiency*, we observe a clear effect of wayfinding skills on the normalized number of *cope_egoc* strategies (low skills: 0.465 vs. high skills: 0.306). Moreover, with growing *uncert_tol_egoc* and *uncert_tol_alloc*, not surprisingly, the number of coping strategies drop sharply (e.g. for *cope_egoc* strategies: low uncertainty tolerance: 0.681 vs. high uncertainty tolerance: 0.080).

Hence, as can be seen particularly from the differing results on effectiveness listed in Table 2, our ABM was indeed capable of simulating wayfinding situations of varying complexity. Individual differences between wayfinder agents are also observable and especially apparent in our results on efficiency.

6 Discussion and Conclusion

We have presented a conceptual framework of uncertainty in wayfinding and used it as a basis for an ABM. In an exemplary implementation, we demonstrated its capability to model uncertainty as a result of non-deterministic reference system transformations in different wayfinding situations consisting of agent, environment and instruction.

While uncertainty has certainly been addressed in wayfinding research before (e.g., [29]), our framework addresses the topic from a novel perspective and has features which make it attractive for future applications to uncertainty-aware wayfinding assistants. In particular, it covers the dynamic aspect of uncertainty (over several decision points) and allows to differentiate between uncertainty on different levels of spatial reference systems.

Although our model currently involves highly abstracted representations of the agent and the environment, it is illustrated how it can inform the choice of an algorithm for the generation of route instructions based on the individual characteristics of wayfinder and environment. Such understanding on the conceptual level can be of value for the design of future wayfinding assistant systems which take their users' uncertainty into account.

However, there are also some shortcomings. Some concepts and processes in the ABM are drastically simplified, including the particular graph representation of the environment with spatial configurations of highly abstracted object types and spatial relations. Still, however, we expect that our fundamental concepts can be applied to more realistic environmental models as well. A further point of simplification is the representation of the wayfinding skills as determinants of a random error distribution in fit perception. Moreover, due to our focus on wayfinding in unknown environments, we did not model transformations from and to survey reference systems, such as maps. The ABM assumes that uncertainty changes only at decision points. This is a simplification because (ambiguous) reference system transformations may also occur while moving between decision points when the observed environment does not match the wayfinder's expectations (route monitoring).

Still, however, our work is valuable as a conceptual basis for the development of uncertainty-aware wayfinding assistance systems. For future work, we aim to apply the simulation to real-world urban networks. Further, user experiments would be required to gain a deeper insight on how humans perceive and react to uncertainty in different wayfinding decision situations. Moreover, it would be interesting to investigate the influence of specific coping strategies on egocentric, allocentric, and survey uncertainty.

References

1 Gary L. Allen. Spatial abilities, cognitive maps, and wayfinding. *Wayfinding behavior: Cognitive mapping and other spatial processes*, pages 46–80, 1999.

2 Tad T. Brunyé, Stephanie A. Gagnon, Aaron L. Gardony, Nikhil Gopal, Amanda Holmes, Holly A. Taylor, and Thora Tenbrink. Where did it come from, where do you go? Direction sources influence navigation decisions during spatial uncertainty. *The Quarterly Journal of Experimental Psychology*, 68(3):585–607, 2015.

3 Matt Duckham and Lars Kulik. "Simplest" paths: automated route selection for navigation. In *International Conference on Spatial Information Theory*, pages 169–185. Springer, 2003.

4 Andrew U. Frank. Formal models for cognition – taxonomy of spatial location description and frames of reference. In *Spatial Cognition*, pages 293–312. Springer, 1998.

5 Adrian Furnham, Joseph Marks, et al. Tolerance of ambiguity: A review of the recent literature. *Psychology*, 4(09):717, 2013.

6 Ioannis Giannopoulos, Peter Kiefer, Martin Raubal, Kai-Florian Richter, and Tyler Thrash. Wayfinding decision situations: A conceptual model and evaluation. In *Proceedings of the Eighth International Conference on Geographic Information Science (GIScience 2014)*, pages 221–234. Springer International Publishing, 2014.

7 Shazia Haque, Lars Kulik, and Alexander Klippel. Algorithms for reliable navigation and wayfinding. In *International Conference on Spatial Cognition*, pages 308–326. Springer, 2006.

8 Mary Hegarty, Daniel R. Montello, Anthony E. Richardson, Toru Ishikawa, and Kristin Lovelace. Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34(2):151–176, 2006.

9 Serge P. Hoogendoorn and Piet H. L. Bovy. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190, 2004.

10 Bin Jiang. A topological pattern of urban street networks: universality and peculiarity. *Physica A: Statistical Mechanics and its Applications*, 384(2):647–655, 2007.

11 Peter Kiefer, Simon Scheider, Ioannis Giannopoulos, and Paul Weiser. A Wayfinding Grammar Based on Reference System Transformations. In *Spatial Information Theory (COSIT 2015)*, volume 9368 of *Lecture Notes in Computer Science*, pages 447–467. Springer International Publishing, 2015.

12 Alexander Klippel, Stephen Hirtle, and Clare Davies. You-are-here maps: Creating spatial awareness through map-like representations. *Spatial Cognition & Computation*, 10(2-3):83–93, 2010.

13 Alexander Klippel, Heike Tappe, Lars Kulik, and Paul U. Lee. Wayfinding choremes – a language for modeling conceptual route knowledge. *Journal of Visual Languages & Computing*, 16(4):311–329, 2005.

14 Christian Kray, Christian Elting, Katri Laakso, and Volker Coors. Presenting route instructions on mobile devices. In *Proceedings of the 8th International Conference on Intelligent User Interfaces*, pages 117–124. ACM, 2003.

15 Christian Kray, Gerd Kortuem, and Antonio Krüger. Adaptive navigation support with public displays. In *Proceedings of the 10th International Conference on Intelligent User Interfaces*, IUI'05, pages 326–328, New York, NY, USA, 2005. ACM.

16 Werner Kuhn. Semantic reference systems. *International Journal of Geographical Information Science*, 17(5):405–409, 2003.

17 Benjamin Kuipers. Modeling spatial knowledge. *Cognitive science*, 2(2):129–153, 1978.

18 Matt MacMahon, Brian Stankiewicz, and Benjamin Kuipers. Walk the talk: connecting language, knowledge, and action in route instructions. In *Proceedings of the 21st National Conference on Artificial Intelligence*. July, 2006.

19 Daniel R. Montello. Scale and multiple psychologies of space. In *Spatial Information Theory*, pages 312–321. Springer, 1993.

20 Daniel R. Montello. Navigation. In *Cambridge handbook of visuospatial thinking*, pages 257–294. 2005.

21 Anand S. Rao, Michael P. Georgeff, et al. BDI agents: From theory to practice. In *ICMAS*, volume 95, pages 312–319, 1995.

22 Martin Raubal. Ontology and epistemology for agent-based wayfinding simulation. *International Journal of Geographical Information Science*, 15(7):653–665, 2001.

23 Martin Raubal and Ilija Panov. A formal model for mobile map adaptation. In *Location Based Services and TeleCartography II*, pages 11–34. Springer, 2009.

24 Martin Raubal and Stephan Winter. Enriching wayfinding instructions with local landmarks. In *International Conference on Geographic Information Science*, pages 243–259. Springer, 2002.

25 Kai-Florian Richter and Stephan Winter. Landmarks. *Springer Cham Heidelberg New York Dordrecht London*, 10:978–3, 2014.

26 Ilias Sakellariou, Petros Kefalas, and Ioanna Stamatopoulou. Enhancing netlogo to simulate BDI communicating agents. In *SETN 2008, LNAI 5138*, pages 263–275. Springer, 2008.

27 Simon Scheider. *Grounding geographic information in perceptual operations*, volume 244. IOS Press, 2012.

28 Maximilian Schirmer, Johannes Hartmann, Sven Bertel, and Florian Echtler. Shoe me the way: A shoe-based tactile interface for eyes-free urban navigation. In *Proc. of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 327–336. ACM, 2015.

29 Martin Tomko and Kai-Florian Richter. Defensive Wayfinding: Incongruent Information in Route Following. In *Spatial Information Theory (COSIT 2015)*, volume 9368 of *Lecture Notes in Computer Science*, pages 426–446. Springer International Publishing, 2015.

30 Ladina B. Tschander, Hedda R. Schmidtke, Carola Eschenbach, Christopher Habel, and Lars Kulik. A geometric agent following route instructions. In *International Conference on Spatial Cognition*, pages 89–111. Springer, 2002.

31 Barbara Tversky. Cognitive maps, cognitive collages, and spatial mental models. In *Spatial Information Theory*, pages 14–24. Springer, 1993.

32 Paul Weiser and Andrew U. Frank. Cognitive Transactions – A Communication Model. In *Conference on Spatial Information Theory (COSIT), number 8116*, pages 129–148. Springer International Publishing, 2013.
33 Matthias Westphal and Jochen Renz. Evaluating and minimizing ambiguities in qualitative route instructions. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 171–180. ACM, 2011.
34 Jan M. Wiener, Simon J. Büchner, and Christoph Hölscher. Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, 9(2):152–165, 2009.

Timing of Pedestrian Navigation Instructions*

Ioannis Giannopoulos[1], David Jonietz[2], Martin Raubal[3], Georgios Sarlas[4], and Lisa Stähli[5]

1 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
igiannopoulos@ethz.ch
2 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
jonietzd@ethz.ch
3 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
mraubal@ethz.ch
4 **Institute for Transport Planning and Systems, ETH Zürich, Zürich, Switzerland,**
sarlasg@ethz.ch
5 **Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland**
staehlli@student.ethz.ch

Abstract

During pedestrian navigation in outdoor urban environments we often utilize assistance systems to support decision-making. These systems help wayfinders by providing relevant information withing the context of their surroundings, e.g., landmark-based instructions of the type "turn left at the church". Next to the instruction type and content, also the timing of the instruction must be considered in order to facilitate the wayfinding process. In this work we present our findings concerning the user and environmental factors that have an impact on the timing of instructions. We applied a survival analysis on data collected through an experiment in a realistic virtual environment in order to analyze the expected distance to the decision point until instructions are needed. The presented results can be used by navigation systems for instruction timing based on the characteristics of the current wayfinder and environment.

1998 ACM Subject Classification I.6.3 Applications

Keywords and phrases navigation, wayfinding, instructions, timing, survival analysis

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.16

1 Introduction

Various studies have investigated design implications for navigation instructions: instructions should be simple [20] as well as connected to landmarks [14, 15], and the mode of the instruction (e.g., visual or auditory) is of secondary importance for user performance [9]. Besides requirements on how to compose navigation instructions, there is almost no research about *when* to provide the user with such information. The following general conclusion was stated by Winter [29] in the context of landmark-based navigation systems: *"People feel comfortable if they recognize reference features early, before arrival at a decision point. They*

* All authors contributed equally (names in alphabetical order).

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore;
Article No. 16; pp. 16:1–16:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

feel confirmed that they are on track, and they do not need to break movement at the decision point, but can interpret the next wayfinding instruction in advance." (p. 350).

Even though there are certain obvious implications for the timing of instructions based on this statement, there are no specific guidelines with regards to spatial and temporal dependencies of navigation instructions for pedestrian navigation aids. This type of guidelines has already been identified and discussed for in-vehicle navigation systems [21], and has been addressed heavily by research since the 1980's mainly due to safety implications.

Pedestrian navigation does not show as many safety concerns connected to instruction timing as car navigation. Nevertheless, instruction timing can be an essential component for wayfinding experience. Sub-optimal instructions timing could lead to unnecessary interaction with the assistance aid [6] as well as increase the cognitive load of the user. Therefore, there is a need to establish design guidelines for pedestrian navigation that include instruction timing.

The main contribution of this paper is to address a research gap in pedestrian navigation considering instruction timing. We report on the results that were gained through a time-to-event analysis (Survival Analysis, [12]) on data collected in a user experiment performed in a realistic virtual urban environment. Through this analysis we identified factors which have an impact on timing, provide a survival rate function (i.e., the probability distribution that a wayfinder would have already asked for instructions at a given distance) as well as the estimated parameters for applying the model in navigation assistance systems.

The rest of this paper is structured as follows: we continue with related work and provide an application scenario that is followed by the relevant research questions. Next, we introduce the utilized methods, describe the experiment and provide the results. We close with a general discussion and provide conclusions and future research directions.

2 Related Research

2.1 Aided Wayfinding

When we want to reach a destination, we usually follow an intended, predefined route. In fact, a prerequisite for such coordinated and goal-directed movement is our ability to successfully navigate, which, according to Montello [16], activates the two separate sub-processes wayfinding and locomotion. The former, which lies within the focus of this study, has been defined as "the process of determining and following a path or route between an origin and destination" [8], and thus emphasizes the planning aspect of navigation.

Planning ahead is especially relevant due to the fact that the destination of a wayfinding process is typically not directly accessible, i.e., it is located outside of our immediate perceptual field [18]. In such cases, we either rely on our own memory or external wayfinding aids, such as verbal or graphical route descriptions and depictions, for making the correct choices at decision points [25]. Apart from a mere cartographic representation of the environment, route descriptions can include both destination descriptions [26], with the focus on *where* to go, and turn-by-turn directions [30], focusing on *how* to get there. The latter usually consist of a set of communicative statements which include environmental features (e.g. landmarks or decision points), delimiters (e.g. distance designations), verbs of movement (e.g. walk or turn), and state-of-being verbs (e.g. you stand in front of the train station) [1].

In general, however, both aided and unaided wayfinding are still challenging tasks. Giannopoulos et al. [7], for instance, identified the environment, the route instruction, and the user as determining factors for the complexity of wayfinding situations. For instance, a more complex environmental layout, e.g. an intersection with a higher number of connecting streets, is harder to mentally represent and to become familiar with [26]. Even in a less complex environment, though, wayfinding may fail due to inappropriate route instructions, which, for instance, can differ in terms of their sequential ordering from the decision points actually encountered in the environment [1]. Finally, users do not belong to a homogeneous group, but rather differ in terms of their spatial abilities, preferences, environmental familiarity, and specific needs [17]. As a result, they may require different times for route instruction comprehension and following, or prefer different representations of route instructions. In order to provide optimal route instructions and decrease the complexity of the wayfinding process, thus, these factors have to be taken into account.

2.2 Timing of Route Instructions

A further aspect which has so far been largely overlooked in the discussion on optimizing route directions is their timing. An instruction provided too early might result in a user simply forgetting it until finally reaching the related decision point. An instruction which is given too late, on the other hand, can lead to a user completely missing the decision point and the required behavior, or making an uncomfortably sharp turn [22]. Apart from safety issues, such shortcomings are likely to reduce system acceptance.

Despite the importance of appropriate instruction timing, to the best of our knowledge, there is currently no empirical evidence on how to determine this point in time for mobile pedestrian navigation systems. There have been, however, several studies which aimed to provide practical guidelines or identify the determining factors on instruction timing for car navigation, and which can be of relevance for other modes as well. Thus, in [22] for instance, experts provided test drivers with simple route instructions that were intentionally given either too late, too early or at an optimal point in time. Based on posterior ratings of the temporal appropriateness of the instructions by the test persons, the authors could identify the significant factors, which included the distance and time to the next junction, the driving speed, the type of required maneuver, and the complexity of the route instruction. These findings are in accordance with the U.S. Federal Highway Administration's general guidelines for navigation systems [4], which add the factors weather and driver characteristics such as age to the list. A further empirical study also found age and gender of the car driver, the speed and type of turn (left or right), as well as the number of vehicles ahead to be significant for defining optimal distances prior to intersections for auditory route instructions [5]. Further environmental aspects were examined by Schraagen [23], who found the visibility of road signs to be of relevance for deciding when to present a new route instruction. Verwey et al. [27], finally, proposed to stack instructions which refer to decision points encountered in close succession (within less than 10 sec) to decrease cognitive load.

In general, due to the significant systemic differences between walking and driving, it is highly questionable whether these findings can be transferred to pedestrian navigation. Thus, in unobstructed spaces, the walking speed is relatively constant, and therefore possibly not of relevance for instruction timing. Regarding other factors discussed, such as the distance or time to the next decision point, visibility relationships, the complexity of the required maneuver, or the user characteristics, in contrast, a relation to route instruction timing seems more plausible, and should be investigated further. An additional aspect could be the type of route instruction, since, for example, reading a textual instruction would require more time than looking at a graphical depiction or listening to an auditory instruction [13]. Another worthwhile direction could be to develop adaptable systems which allow for personalization of the timing conditions for different types of instructions [31].

2.3 Modeling

Modeling of the relationship between a dependent variable and a set of independent variables (covariates) has been the focus of interest in many fields. Essentially, modeling constitutes a way to identify and quantify the impact that the covariates exert on the dependent variables. Statistical modeling enables making statements about how much "faith" we can put in those estimates and also provide ways to improve the overall fit, and in turn the predictive accuracy of the developed models on the basis of a set of mathematical assumptions. Without doubt, the most popular and widely employed model is the family of linear regression models. However, such models have some underlying assumptions which if not met, can result in biased and inconsistent coefficient estimates, and thus incorrectly estimated and perhaps misspecified models. One of the main assumptions is the normality of the error terms and different ways have been developed to by-pass this, resulting in new modeling frameworks. When it comes to the problem of analyzing time-to-event data (also called *Survival Analysis*), linear regression fails to provide "correct" estimates mainly due to the underlying distribution of the modeled process. Time-to-event models have been widely applied to a plethora of problems in different domains, varying from biometrics to industrial engineering to transportation research. An overview of applications is presented in [3, 11, 12]. Departing from the time setting, Waldorf [28] explored and verified the conceptual equivalence between survival models applied to both temporal and spatial processes, focusing on the "at-risk" concept. In addition, he highlighted the limited number of applications of hazard models in spatial settings. Following his study, a number of applications have built upon that and utilized hazard models for such distance related problems (e.g. [2]).

3 Instruction Timing

3.1 Application Scenario

The following application scenario aims at highlighting the importance of instruction timing on the wayfinding process by outlining three scenarios with different types of outcomes depending on when instructions are available to the user.

Consider the following typical situation: Alice has just arrived at the main station of Zurich as a fist time visitor. She wants to walk to her hotel from the main station and types the address into her navigation system that will guide her through the city by providing audio instructions.

- **Scenario A**: *Alice is approaching the first intersection and just when she thinks she would like to know where to go next, the navigation system instructs her to "Turn left at the restaurant with the red façade". Alice was already observing the next intersection and immediately spots the restaurant and turns left at the intersection.*
- **Scenario B**: *Alice just passed the first intersection and the navigation system already provides the next instruction. By the time Alice arrives at the next intersection, she is not sure anymore if she is supposed to turn left or right.*
- **Scenario C**: *Alice passed the second intersection. She continues walking and gets closer to the intersection without receiving any instructions. Alice is getting nervous and starts feeling uncomfortable.*

In Scenario A the instruction is provided to the user with the right timing whereas in Scenario B the instruction is given too early and in Scenario C too late respectively. Scenarios B and C demonstrate two situations where the wayfinder might get confused and cognitively overloaded, leading to poor user experience and possibly poor wayfinding performance.

3.2 Research Questions

The focus of this work lies on the investigation of the environmental and user properties that have an effect on instruction timing. Furthermore, we are interested in predicting the timing of an instruction based on the characteristics of the wayfinder and the environment. The main research questions of interest are the following:
RQ1: Do wayfinders prefer to receive a navigation instruction multiple times?
RQ2: Which properties of the environment have an impact on navigation instruction timing?
RQ3: Which properties of a wayfinder have an impact on navigation instruction timing?
RQ4: Is it possible to predict when an instruction should be optimally provided?

With these research questions we aim to investigate the topic of instruction timing for pedestrian navigation allowing us to include or exclude certain characteristics (e.g., spatial abilities) for further research. The first research question investigates the possibility that a navigation instruction should be given more than once. The second and third research questions address the properties (e.g., visibility of a decision point, user's spatial abilities) that influence instruction timing. Finally, the fourth research question addresses the possibility of estimating a model that will allow us to predict the timing of instructions based on the characteristics of the wayfinder and the environment.

4 Method

4.1 Implementation

A prototype of a navigation system based on local landmarks [19] connected to audio instructions was implemented in a realistic virtual environment. Navigation instructions were triggered by the participants by pushing a button of a joystick.

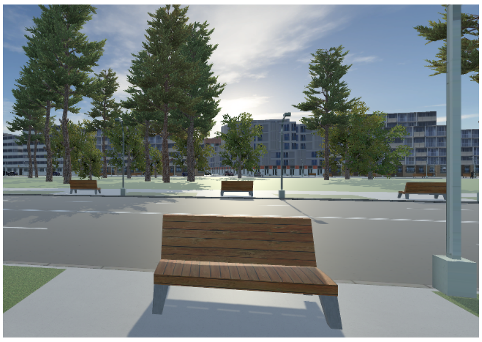
The used hardware consisted of a Logitech 3D Precision Pro joystick to enable movement and interaction with the navigation system, an HP XB31 digital projector for displaying the virtual environment and a gaming computer for rendering, executing the experiment and logging all user data.

The realistic virtual environment including street layout, building blocks and façade textures was designed using the ESRI CityEngine[1] with the aid of the Complete Street Rule which features realistic street furniture, such as traffic lights and benches (see Figure 1a). The generated city was imported into the Unity3D[2] game engine where a realistic skybox was added as well as data collection scripts and the interaction with the navigation system. The correct interaction with the navigation system (i.e., the instruction only for the next decision point was played) was achieved by using colliders that were placed and filled up the space between decision points.

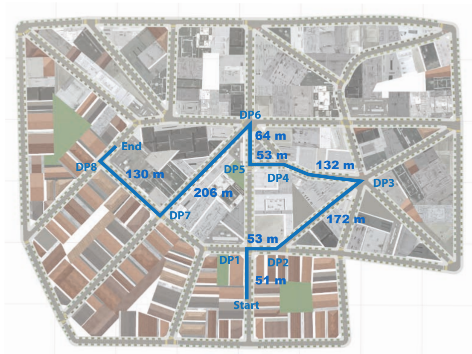
4.2 Experiment

A user study was conducted in the virtual environment to evaluate preferences for instruction timings. The participants were divided into two groups: the first group was able to request instructions per decision point as often as they wanted (multiple-clicks) whereas the second group was limited to one instruction per decision point (one-click). The navigation instructions were given as audio instructions (e.g., 'turn right at the building with the green façade')

1 `http://www.esri.com/software/cityengine`
2 `http://www.unity3d.com`



(a)



(b)

Figure 1 (a) shows an example scene of the virtual environment. (b) shows the complete route of the navigation path with 8 decision points and the length of the segments in meters.

connected to colored façades in the virtual environment simulating local landmarks. The audio instructions were given using a female American-English computer-generated voice.

Participants were standing in front of a height-adjustable table where the joystick was placed in the middle and used to navigate through the virtual environment. The virtual environment was projected to the opposite wall so that participants faced it at a distance of about 3 meters.

4.2.1 Design

The user experiment was designed as a Between-Subjects study. Participants were randomly assigned to the two conditions, trying to balance gender and spatial abilities. The navigation path they were required to move along consisted of 8 decision points (see figure 1b). We designed a route with several decision points, having different numbers of connections, with the intention to investigate the effect of decision points with varying structure and complexity [7]. The considered types were Cross-intersection, T-intersection, Y-intersection and Star-intersection. Each type occurred twice on the route, once with the next intersection being visible from the beginning and once not. This allowed us to investigate the impact of decision point visibility.

4.2.2 Procedure

First, the participants were informed about the experiment and the procedure. Second, they had to provide their demographic information, as well as to fill in the *Santa Barbara Sense of Direction Scale* questionnaire for the self-estimation of their spatial abilities (*SBSODS*, [10]).

After filling in the questionnaires, the actual task started. Participants were instructed to stay on the sidewalks and use the pedestrian crossings. They started walking (constant speed 5m/s) along the navigation path in a first-person view. Based on the condition, they were able to ask once or multiple times for instructions by pushing the joystick button.

The user location (x,y) and the corresponding decision point were logged whenever a navigation instruction was requested by pushing a button of the joystick.

4.2.3 Participants

A total of 45 people participated in the user experiment. Due to tracking problems, one participant had to be dropped. The one-click condition had 23 participants (11 female) with a mean age of 25.7 years (*SD = 5.4*). The multiple-clicks condition had 21 participants (11 female) with a mean age of 27.7 years (*SD = 8.2*). This results in a total number of 352 cases (44 participants * 8 decision points). The participants in both conditions came from various professional (e.g., Geography, Marketing) and cultural (e.g., Swiss, Greek) backgrounds.

4.3 Modeling Approach

Given the nature of the time-to-event process that we aim to model, the family of hazard based models is exploited. Naturally, the choice of a model with a fully parametric hazard function advances as the most appropriate, offering the medium for simultaneously describing the basic underlying survival distribution and quantifying how that distribution changes as a function of the covariates [12]. In addition, in cases where the distribution of the survival times can be adequately approximated by an existing mathematical function, the use of a fully parametric model is preferred since it can facilitate the transferability and generalization of the model estimates. Two such models exist, differing essentially in their underlying assumptions of how the hazard, and subsequently the survival function, is modeled. In particular, the proportional hazards model assumes that the covariates have a constant multiplicative effect on the hazard function while the accelerated failure time (AFT) models [12] assume that the effect of the covariates on the hazard function is multiplicative on the time scale, thus not constant.

Based on the problem at hand, assuming a constant impact of the covariates on the hazard rate would constitute a rather restrictive choice that is not aligned with our expectations. The choice of the AFT models on the other hand, assuming that time (distance in our case) has an effect on the impact of the covariates, appears as a more plausible alternative of the underlying survival function we want to model, being capable of facilitating both accelerating and decelerating effects on the survival time.

Let us denote the traveled distance from the previous decision point where instructions were asked (in analogy to the time-to-event concept) as t, having a cumulative distribution function T such as $F(t) = Pr(T \leq t)$. The survival function gives the probability of observing a survival distance higher than t, denoted as $S(t) = \Pr(T > t) \ = 1 - F(t)$. The probability of a process ending at point t, given that it has lasted up to that point, is called hazard rate and is defined as:

$$h(t) = \frac{\mathrm{f}\,(\mathrm{t})}{S(t)}\,. \tag{1}$$

In the case of the AFT models with a Weibull survival function, a convenient way to characterize the distribution of time is $T = e^{\beta_0+\beta_i x} \ * \varepsilon$, with β's representing the effect of the covariates and ε an error component. This function can be easily linearized as $ln(T) = \beta_0 + \beta_i x_i + \sigma * \varepsilon^*$, with $\varepsilon^* = ln(\varepsilon)$ following the extreme minimum value distribution, denoted as $G(0, \sigma)$ with σ called the scale parameter. Then the hazard and the survival function can be written as:

$$h(t, \chi_i, \beta_i, \lambda) = \frac{\lambda t^{\lambda-1}}{\left(e^{\beta_0+\beta_i x_i}\right)^{\lambda}} = \lambda t^{\lambda-1} e^{-\lambda(\beta_0+\beta_i x_i)} = \lambda\gamma\left(te^{-\beta_i x_i}\right)^{\lambda-1} e^{-\beta_i x_i} \tag{2}$$

$$S(t, \chi_i, \beta_i, \sigma) = exp\{-t^{\lambda} exp[(-1/\sigma)(\beta_0 + \beta_i x_i)]\ \} \tag{3}$$

with $\lambda = 1/\sigma$ *and* $\gamma = exp(-\beta_0/\sigma\ )$. With this formulation, the equation for the median survival time is:

$$t_{50}\left(\chi_i, \beta_i, \sigma\right) = \left[-ln\left(0.5\right)\right]^{\sigma} e^{\beta_0+\beta_i x_i}\ . \tag{4}$$

5 Results

Before proceeding to the estimation of the model, the distance values are normalized to the range of [0,1] in order to make them comparable and compatible for the estimation process. As distance we consider the traveled distance up to the point in space where the participants requested instructions. Subsequently, we center our focus on estimating an AFT model with a Weibull survival distribution in terms of maximum likelihood, using a robust "sandwich" standard error estimator capable of identifying clusters of residuals, hence relaxing the independence-of-observations assumption. The estimation was conducted using the open-source statistical software R, making use of the *Survival* package [24]. It should be mentioned that other parametric forms of survival functions (e.g., exponential, Gaussian, logistic etc.) have been checked as well, however the Weibull provided the best fit.

The variable selection process was conducted iteratively on a goodness-of-fit basis by minimizing the Akaike Information Criterion (AIC) which penalizes for the number of included parameters, hence accounting for over-fitting. The reported estimates are presented in Table 1 and they are all found to be statistically significant at the 5% level. An estimate with a positive sign implies a longer survival (i.e., instructions will be necessary at a later point), while a negative one implies the opposite. The rate of how longer or shorter the survival becomes varies along the survival function, but on its median it can be easily quantified by the following formula:

$$TR\left(x_{11}, x_{12}\right) = \frac{t_{50}\left(x_{12}, \beta_1, \sigma\right)}{t_{50}\left(x_{11}, \beta_1, \sigma\right)} = \frac{\left[-ln\left(0.5\right)\right]^{\sigma} e^{\beta_0+\beta_{i-1}x_i-1+\beta_1 * x_{12}}}{\left[-ln\left(0.5\right)\right]^{\sigma} e^{\beta_0+\beta_{i-1}x_i-1+\beta_1 * x_{11}}} = e^{\beta_1 \Delta x_1}\ . \tag{5}$$

Essentially, obtaining the betas allows us to estimate the survival and hazard functions for different sets of covariates (and hence individuals and spatial variants) and consequently we can proceed to obtain point estimates of quantiles of the distribution (e.g., the median) that can be of potential interest for prediction purposes.

In the plots of Figure 2, the estimated mean survival function per intersection type is presented (mean in the sense that the mean value of the covariates is plugged into the formula) along with the observed non-parametric survival function to provide a visual evaluation of the model's fit. In the plot of Figure 3a, different variations of the characteristics of Alice are presented to highlight the impact of the covariates on the survival function. All included attributes were found to be statistically significant, while the magnitude of their impact (parameters) is highlighted in the figures with different variables' values. Last, in Figure 3b, a scatter plot of the predicted median survival values versus the observed ones for the sample of observations is presented.

6 Discussion

This section interprets and discusses the results from the user experiment. The mean distance at which the wayfinders asked for instructions was at 45.4% (*SD = 25.2%, Median = 44.7%*) of the total segment length. More than half of the times, i.e., in 68.7% of the cases, the wayfinders asked for instructions after the decision point was visible from the distance (the visibility point occurred in average at 48% of the segments' length. The cases where the next decision point was visible from the beginning were excluded).

Table 1 The β estimates (N = 352, Log Likelihood = 28.3, χ^2 = 107.38 with p<.001, AIC = −34.65).

	Value	SE	Z	P
Intercept	*−1.627*	*0.151*	*−10.76*	*<0.001*
DP Visibility	*2.031*	*0.480*	*4.23*	*<0.001*
Long Segments	*−0.504*	*0.226*	*−2.23*	*0.026*
Condition	*−0.203*	*0.051*	*−3.97*	*<0.001*
Y-Intersection	*0.488*	*0.114*	*4.27*	*<0.001*
Star-Intersection	*0.488*	*0.127*	*3.84*	*<0.001*
T-Intersection	*0.243*	*0.117*	*2.07*	*0.039*
Age	*0.019*	*0.005*	*3.91*	*<0.001*
Age older than 27	*−0.348*	*0.084*	*−4.16*	*<0.001*
Low Spatial Abilities	*−0.173*	*0.075*	*-2.29*	*0.022*
Log(scale)	*−0.639*	*0.085*	*−7.53*	*<0.001*

RQ1: Do wayfinders prefer to receive a navigation instruction multiple times? Although participants in the multiple-clicks condition could ask for instructions as often as they wanted, they did not. Only 14.4% of the cases (25 out of 168) participants asked more than once for instructions. The maximum number of times that an instruction was asked for regarding a certain decision point was 3 times and occurred only for 1.7% of the cases. This result suggests that receiving a wayfinding instruction once is considered sufficient.

RQ2: Which properties of the environment have an impact on navigation instruction timing? According to the results (see Table 1), there are environmental properties that have a significant impact on instruction timing. The visibility of a decision point (DP Visibility), which is the location on the segment from which the wayfinders could perceive that a decision point is coming ahead of them, has a significant effect on the timing of instructions. Furthermore, the length of the segment (Long Segments) that has to be traveled as well as the type of intersection (Cross-, Y-, Star-, and T-Intersection) have also a significant effect. The estimates for the visibility and type of decision point all have a positive sign, revealing that the further away the visibility point and depending on the type of the decision point (based on the estimates, the order is Cross-, T-, Y-, and Star-Intersection), the later the wayfinders will need instructions. Subsequently, as shown in Table 1, the longer the segment the sooner (in terms of normalized distance) instructions will be necessary. The effect of the decision point type on the survival rates is illustrated in Figure 2.

RQ3: Which properties of a wayfinder have an impact on navigation instruction timing? The relevant wayfinder properties (see Table 1) that have a significant impact on instruction timing are the age of the wayfinder, the age threshold (Age older than 27), which categorizes them as younger and older wayfinders (based on the mean age of our sample, 27 years) as well as their spatial abilities, which were clustered into low (below a SBSOD score of 3), medium, and high (Low Spatial Abilities). Furthermore, the condition (single- or multiple-clicks) also had a significant effect. The inclusion of gender specific effects were found to be statistically insignificant. The estimates for age have a positive sign, showing that the older the wayfinder the later instructions will be necessary, however, taking into account the negative sign of the age threshold, there is a diminishing effect for ages over 27. The estimates for the spatial abilities have a negative sign, which reveals that if the wayfinder has low spatial

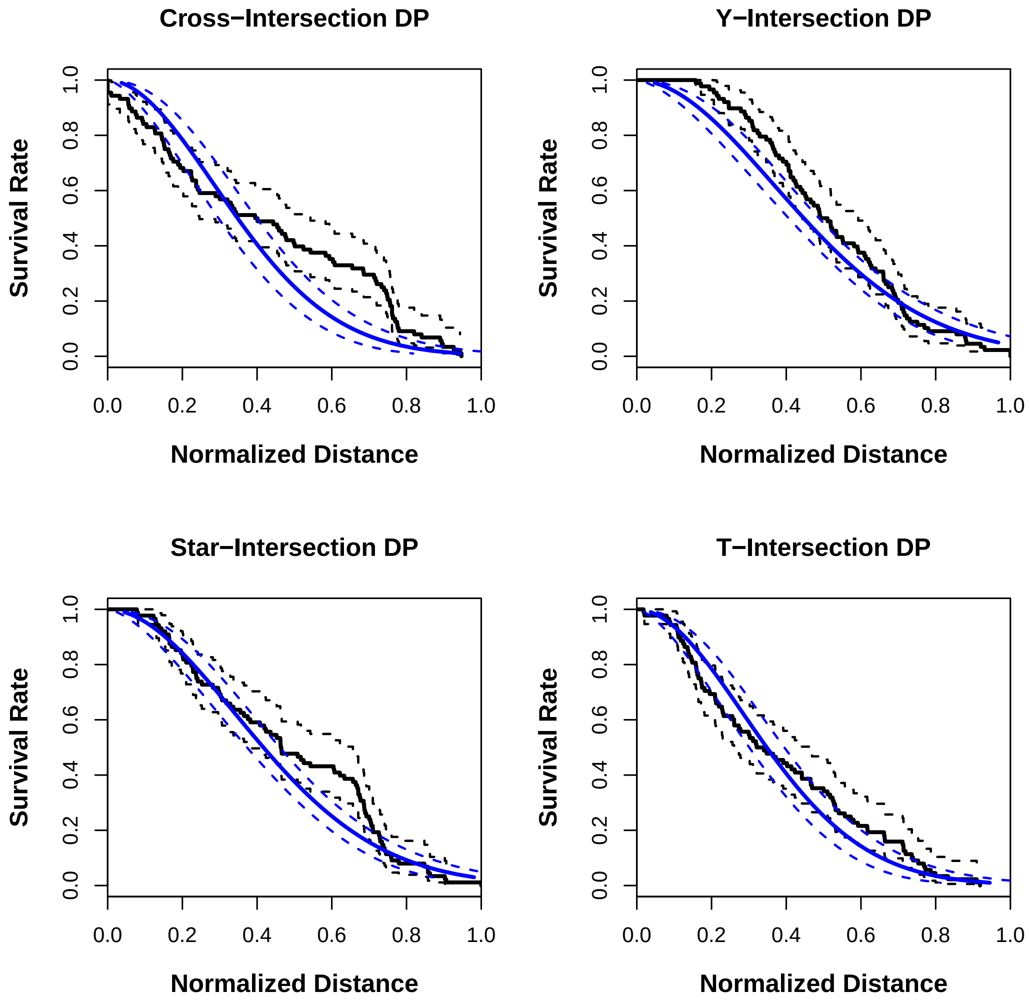


Figure 2 The mean survival functions (blue) for the four types of intersections vs. the mean observed ones (black). The 95 % confidence intervals are presented with dotted lines.

abilities (below a SBSOD score of 3), the sooner instructions will be necessary. These results are exemplified in Figure 3a. The red line depicts the survival rate for agent Alice, who is 22 years old, having medium to high spatial abilities (SpatAb: 0). The red dotted line illustrates the survival rate for agent Alice if she would have low spatial abilities. A possible interpretation of this result could be that the higher the spatial abilities, the higher the confidence of the wayfinder concerning the interpretation and mapping of instructions just before the decision point. Another interpretation could be that wayfinders with high spatial abilities wait longer in order to minimize the possible space where the given instructions can be mapped. Figure 3a illustrates also the effect of age. The blue lines represent the survival rates for Alice, who is 45 years old. The 45-years-old agent Alice will ask for instructions later than the younger agent Alice. One possible interpretation of this effect could be that the older we get, the more experience we gain, and similarly to the previous interpretation of high spatial abilities, we wait until the possible space where an instruction can be mapped is less.

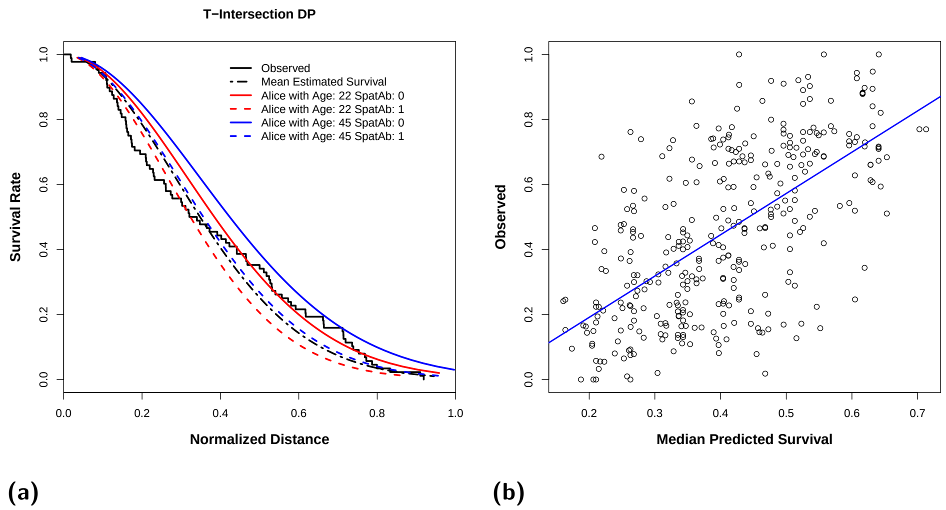


Figure 3 (a) illustrates an example of the change in the survival distribution for different agent user characteristics. (b) shows the predicted median survival values vs. the observed ones for our sample.

RQ4: Is it possible to predict when an instruction should be optimally provided? Figure 3b indicates that such a model allows us to make statements on when an instruction should be given (Adjusted R-squared: 0.389, $p < 0.001$). More specifically, observing the spread, no clear patterns can be identified along with no heteroscedasticity issues. This indicates that no main factor was left out of the model. Taking into account the statistical significance of the reported estimates, we can draw concrete conclusions that these factors have an impact on the process of instruction timing.

Plugging the estimates into equation (3), the survival rate function can be easily calculated for individual wayfinders and different environmental settings (as we did for for the examples in Figure 3a). Given that we provide estimates for a survival rate function, exact prediction of the time when an instruction should be provided is not feasible. Instead, the survival function can be applied in order to determine, based on a given situation, when to provide an instruction on the basis of different criteria (e.g., median survival estimate). A navigation assistance system could benefit from such a survival function by reacting according to the characteristics of the current wayfinder and environment.

Making use of equation (5), the impact of a change in a factor can be quantified on the median survival rate. For instance, a difference on the median survival rate between low and high spatial abilities is approximately 16%.

7 Conclusions and Future Work

In this work we identified relevant environmental and user properties that have an impact on instruction timing. Furthermore, we report the estimated covariates that can be used to calculate a survival function for a given wayfinder and environmental setting. This function can be calculated on the fly by an assistance system and time the instructions accordingly. Since this function provides a probability distribution, the assistance system has also to decide the criteria of the timing, e.g., provide the instruction based on the median survival estimate. Since the observed instruction timings were retrieved based on user preferences, in future work, we will investigate if the proposed way of timing instructions is also increasing the wayfinding experience and performance. Furthermore, we will focus on the generalization of instruction types and environments as well as consider in more detail the environmental complexity [7]. Additionally, we will perform experiments in real urban environments to investigate and compare the external validity of the current work.

References

1 Gary L. Allen. From knowledge to words to wayfinding: Issues in the production and comprehension of route directions. In *Proceedings of the International Conference on Spatial Information Theory*, COSIT'97, pages 363–372, London, UK, 1997. Springer-Verlag.

2 Panagiotis Ch. Anastasopoulos, Mouyid B. Islam, Dionysia Perperidou, and Matthew G. Karlaftis. Hazard-based analysis of travel distance in urban environments: longitudinal data approach. *Journal of Urban Planning and Development*, 138(1):53–61, 2012.

3 Chandra R. Bhat and Abdul R. Pinjari. Duration modeling. In *Handbook of Transport Modelling: 2nd Edition*, pages 105–131. Emerald Group Publishing Limited, 2007.

4 Thomas A. Dingus, Melissa C. Hulse, Steven K. Jahns, J. Alves-Foss, S. Confer, A. Rice, I. Roberts, Richard J. Hanowski, D. Sorenson, et al. Development of human factors guidelines for advanced traveler information systems and commercial vehicle operations: Literature review. *United States. Federal Highway Administration*, 1996.

5 Kellie George, Paul Green, and Jill Fleming. Timing of auditory route-guidance instructions. Technical report, The University of Michigan, 1996.

6 Ioannis Giannopoulos, Peter Kiefer, and Martin Raubal. GeoGazemarks: Providing gaze history for the orientation on small display maps. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, ICMI'12, pages 165–172. ACM, 2012. doi:10.1145/2388676.2388711.

7 Ioannis Giannopoulos, Peter Kiefer, Martin Raubal, Kai-Florian Richter, and Tyler Thrash. Wayfinding decision situations: A conceptual model and evaluation. In *International Conference on Geographic Information Science*, GIScience'14, pages 221–234. Springer, 2014.

8 Reginald G. Golledge. Human wayfinding and cognitive maps. In Reginald G. Golledge, editor, *Wayfinding behavior: Cognitive mapping and other spatial processes*, chapter 1, pages 5–45. Johns Hopkins University Press, Baltimore, MA, 1999.

9 Alice F. Healy, Vivian I. Schneider, Blu McCormick, Deanna M. Fierman, Carolyn J. Buck-Gengler, and Immanuel Barshi. Which modality is best for presenting navigation instructions? *Journal of Applied Research in Memory and Cognition*, 2(3):192–199, 2013.

10 Mary Hegarty, Anthony E. Richardson, Daniel R. Montello, Kristin Lovelace, and Ilavanil Subbiah. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425–447, 2002.

11 David A. Hensher and Fred L. Mannering. Hazard-based duration models and their application to transport analysis. *Transport Reviews*, 14(1):63–82, 1994.

12 David W. Hosmer, Stanley Lemeshow, and Susanne May. *Applied survival analysis.* Wiley Blackwell, 2011.

13 Christian Kray, Christian Elting, Katri Laakso, and Volker Coors. Presenting route instructions on mobile devices. In *Proceedings of the 8th International Conference on Intelligent User Interfaces*, IUI'03, pages 117–124, New York, NY, USA, 2003. ACM.

14 Andrew J. May, Tracy Ross, Steven H. Bayer, and Mikko J. Tarkiainen. Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing*, 7(6):331–338, 2003.

15 Pierre-Emmanuel Michon and Michel Denis. When and why are visual landmarks used in giving directions? In *International Conference on Spatial Information Theory*, COSIT'01, pages 292–305. Springer, 2001.
16 Daniel R. Montello. *Navigation*, chapter 7, pages 257–294. Cambridge University Press, 2005.
17 Daniel R. Montello and Martin Raubal. Functions and applications of spatial cognition. In David Waller and Lynn Nadel, editors, *Handbook of Spatial Cognition*, pages 249–264. American Psychological Association, 2012.
18 Daniel R. Montello and Corina Sas. *Human Factors of Wayfinding in Navigation*, pages 2003–2008. CRC Press/Taylor and Francis, Ltd., 2006.
19 Martin Raubal and Stephan Winter. Enriching wayfinding instructions with local landmarks. In *International Conference on Geographic Information Science*, GIScience'02, pages 243–259. Springer, 2002.
20 Kai-Florian Richter and Matt Duckham. Simplest instructions: Finding easy-to-describe routes for navigation. In *International Conference on Geographic Information Science*, GIScience'08, pages 274–289. Springer, 2008.
21 Tracy Ross and S. Brade. An empirical study to determine guidelines for optimum timing of route guidance instructions. In *Design of the Driver Interface*, pages 1–5. IEE, 1995.
22 Tracy Ross, Gill Vaughan, and Colette Nicolle. Design guidelines for route guidance systems: Development process and an empirical example for timing of guidance instructions. In Y. Ian Noy, editor, *Ergonomics and Safety of Intelligent Driver Interfaces*, pages 139–152. L. Erlbaum Associates Inc., 1997.
23 Jan M. Schraagen. Information presentation in in-car navigation systems. *Driving future vehicles*, pages 171–185, 1993.
24 Terry Therneau. A package for survival analysis in S. R package version 2.37-4, 2014.
25 Martin Tomko and Kai-Florian Richter. Defensive wayfinding: Incongruent information in route following. In *International Conference on Spatial Information Theory*, COSIT'15, pages 426–446. Springer International Publishing, 2015.
26 Martin Tomko and Stephan Winter. Pragmatic construction of destination descriptions for urban environments. *Spatial Cognition & Computation*, 9(1):1–29, 2009.
27 Willem B. Verwey, Håkan Alm, John A. Groeger, Wiel H. Janssen, Marja J. Kuiken, Jan M. Schraagen, Josef Schumann, Wim van Winsum, and Heinz Wontorra. GIDS functions. In *Generic intelligent drive support*, pages 113–146. Taylor & Francis, Inc., 1993.
28 Brigitte Waldorf. Spatial patterns and processes in a longitudinal framework. *International Regional Science Review*, 26(3):269–288, 2003.
29 Stephan Winter. Route adaptive selection of salient features. In *International Conference on Spatial Information Theory*, COSIT'03, pages 349–361. Springer, 2003.
30 Markus Wuersch and David Caduff. Refined route instructions using topological stages of closeness. In Ki-Joune Li and Christelle Vangenot, editors, *Web and Wireless Geographical Information Systems: 5th International Workshop, W2GIS 2005, Lausanne, Switzerland, December 15-16, 2005. Proceedings*, pages 31–41. Springer Berlin Heidelberg, 2005.
31 Hubert Zimmer, Stefan Münzer, and Jörg Baus. From resource-adaptive navigation assistance to augmented cognition. In Matthew W. Crocker and Jörg Siekmann, editors, *Resource-Adaptive Cognitive Processes*, pages 35–53. Springer, 2010.

On Avoiding Traffic Jams with Dynamic Self-Organizing Trip Planning*

Thomas Liebig[1] and Maurice Sotzny[2]

1 **TU Dortmund University, Dortmund, Germany**
`thomas.liebig@tu-dortmund.de`
2 **TU Dortmund University, Dortmund, Germany**
`maurice.sotzny@tu-dortmund.de`

Abstract

Urban areas are increasingly subject to congestions. Most navigation systems and algorithms that avoid these congestions consider drivers independently and can, thus, cause novel congestions at unexpected places. Pre-computation of optimal trips (Nash equilibrium) could be a solution to the problem but is due to its static nature of no practical relevance.

In contrast, the paper at-hand provides an approach to avoid traffic jams with dynamic self-organizing trip planning. We apply reinforcement learning to learn dynamic weights for routing from the decisions and feedback logs of the vehicles. In order to compare our routing regime against others, we validate our approach in an open simulation environment (LuST) that allows reproduction of the traffic in Luxembourg for a particular day. Additionally, in two realistic scenarios: (1) usage of stationary sensors and (2) deployment in a mobile navigation system, we perform experiments with varying penetration rates. All our experiments reveal that performance of the traffic network is increased and occurrence of traffic jams are reduced by application of our routing regime.

1998 ACM Subject Classification I.2.11 Distributed Artificial Intelligence

Keywords and phrases situation-aware trip planning, self-organizing traffic, reinforcement learning

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.17

1 Introduction

During the transition towards smart cities, intelligent traffic systems are used to detect current traffic hazards [10], to predict future traffic states [20] or to provide situation dependent navigation suggestions to drivers [13]. Due to the complex nature of everyday traffic, precise travel time prediction has proven to be an algorithmically challenging problem. Its accuracy is inherently dependent on various inputs, such as spatio-temporal variables, road supply, road demand, vehicle usage, and overall network quality [4].

The routes, however, should avoid current and upcoming traffic jams. This can easily be done individually, by a navigation device or a routing app (e.g. Google) but this could become problematic, as it does not consider greedy route choice amongst the drivers. Every driver uses comparable edge weights and optimal roads are overrepresented. This may lead to novel unexpected congestion on optimal roads during peak periods. And, in turn, optimal roads are no longer optimal.

* This research has been financed by the European Union through the Horizon 2020 688380 VaVeL project.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 17; pp. 17:1–17:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Figure 1 Motivating example to dynamic self-organizing routing. Whereas a car may take the apparently optimal route from S to D (middle) it may also avoid causing unnecessary congestion and use the route depicted to the right. Best viewed in color.

This problem could be approached in two ways. If one knows all trips in advance, one could find optimal static weights amongst the drivers to gain optimal flow through the city. This approach results in a Nash equilibrium, as discussed in [19]. The second approach is to apply dynamic routing and perform self-organization, this approach is yet unexplored and subject to the paper-at-hand.

In contrast to static routing methods, which do not care about other persons decision, in dynamic self-organizing routing (also from a given start to a target location) the drivers continuously answer following questions:

- Which turn should I make?
- Which effects will my decision have?

A visual representation of the expected behaviour is presented in Figure 1. Given a vehicle travelling from location S to D it might choose amongst two possible paths. Though the upper route choice (depicted in the middle) uses the main road and may pertain a better static (also predicted) cost, the vehicle avoids this road as it realizes that it causes unnecessary congestion on this road. The alternative (right) has a reasonable dynamic cost and avoids unnecessary congestion on the main road.

Our approach to this problem is to apply reinforcement learning, as we just observe which decisions the vehicles have done plus their result, but none of the cases that would have happened if some vehicles would have turned differently. We therefore apply a method for learning from bandit feedback to the routing problem. Our experiments reveal that this approach successfully increases performance of the traffic network and reduces traffic jams.

The paper is structured as follows. Section 2 discusses other, non–self-organizing, approaches to congestion preventing traffic control. In Section 3 we present the mapping of the problem to bandit feedback learning and introduce a recent approach to this problem. Section 4 presents the experiments we performed. Here we test the overall achieved performance and the performance we achieve for different penetration rates and deployment scenarios, i.e. stationary or moving sensors. Finally, in Section 5 we discuss future research directions.

2 Related Work

In this section we fist present literature on trip planning problems. Afterwards, we will briefly discuss related approaches for traffic control.

Before digging into our approach on dynamic trip planning, we present some fundamentals on (static) trip planning. The task to plan a route from one start location to a target location is called trip planning, when multiple means of transportation (also called ‘travel modes’) are involved this becomes multi-modal trip planning. The integration of transportation systems with personal constraints, residential and city services systems can offer real promise for implementing an intelligent transportation infrastructure that can efficiently address issues beyond congestion, resiliency and safety. Trip planning operates on a graph representation of the road and transit network the so-called traffic network G consisting of vertices V (e.g. junctions) and connecting edges E (e.g. streets). A cost function maps each edge to a positive number that denotes how much it would 'cost' to travel the corresponding segment. The cost function needs to be consistent throughout the traffic network, but can be defined in several ways, such that it holds the most relevant aspects: for example length of the segment, travel time, or comfortableness. With a given start and end location in the traffic network, trip planning searches the path that connects start and goal and minimizes the cost.

Several algorithms exist to compute this minimizing path. Dijkstra [5] proposes a best-first traversal of the graph where the candidates for traversal are hold in a priority-queue. In the slightly modified version of the algorithm A* [8] the order in the priority-queue for the traversal not only depends on the cumulated costs to reach a vertex in the graph but also on the expected costs to reach the goal from this vertex. Bound by Minkowski's inequality, whereas $||x + y||_p \leq ||x||_p + ||y||_p$ (known as triangle inequality for $p = 2$), A* prunes the search space in comparison to Dijkstra's Algorithm. A sound heuristic for the remaining cost estimation is the geographical distance that is always lower than the road-based distance. In case of static cost functions contraction hierarchies [6] are a data structure that speeds-up the A* algorithm and enables trip calculation in large traffic networks at European scale. Instead of searching the shortest path directly within the traffic network, contraction hierarchies reduce the search space to the most important ones. In a preprocessing step these important segments are identified (based on the topology) and the network is extended by edges between these important links.

According to Hoogendorn [9], individual movement is performed in three layers:

- *Strategic Level*: In the strategic level the driver chooses its target and the strategy how to get there. This is the self estimated best route, among a collection of different alternatives. This can be done based on experience. Examples could be the global shortest path or the familiar path to a given destination.
- *Tactical Level*: Short-terms decisions are made at the tactical level, avoiding jams or switching to a faster route for instance. Thus, the person chooses the path to avoid obstacles. Basic rules for motions are defined at the tactical level, which include accelerating, decelerating and stopping.
- *Operational Level*: In the operational level, the motion to the next intermediate point is performed, for example, decision for a movement direction and speed or planning of the next step.

Based on this characteristic, it is clear that on each layer of this hierarchy smart methods could improve performance of traffic. The game-theoretic Nash equilibrium [18], applied in [19] operates on the *strategic level* and the route is chosen such that any driver may not get a better travel time by changing its own travel plan. Recently proposed system for self-organising traffic [22], uses slots instead of traffic lights and operates on the *operational level* of motion.

In contrast, our approach works on the *tactical level*, as it predicts and avoids jams. As opposed to [13], our approach allows altering the route at driving time and prevents creation of jams by the given navigation advices. Latter aspect is also focus of the research performed in [16, 15]. Their approach is to obfuscate the given signals such that the traffic distributes better, whereas we use regular sensor data as available from traffic loop networks or navigation devices and provide space time dependent suggestions, which are (based on the different locations of the traffic participants) individual.

3 Combination of Routing Decisions and Congestion Feedback for Reinforcement Learning

Traffic closely resembles a bandit feedback learning environment (compare [1] for an introduction to bandit learning). Bandit learning is a reinforcement learning task, where the behaviour of some blackbox (e.g. a bandit) should be learned just by the feedback we observe, several actions can be taken (in the bandit problem this equals drawing an arm). However, only the result of the actions can be observed and it is unknown what would have happened otherwise. Vehicles serve as agents which move in a road network. The actions are represented by the roads a vehicle can choose at an intersection. Once a road was chosen, a reward will be assigned for that particular road depending on its actual state. The reward for all other roads which could have been chosen remains unknown. This lack of fully labeled data makes a supervised learning approach particularly complex.

The *Policy Optimizer for Exponential Models* (POEM) [21] is able to learn solely based on the reward values provided by the environment. Additionally, POEM does not perform on-line learning, but rather uses logged data. This abstraction is known from bandit problems, where a reward should be optimized from the sole information gained after turning the arm of the bandit. This presents a more robust approach, since a learned model can be thoroughly tested before deployment. The system will also not evolve over time, which could lead to unpredictable behavior. This is particularly undesirable in the context of vehicle routing.

The following sections outline how POEM[1] can be used to predict congestion in road networks. The results of POEM are then utilized to dynamically route vehicles around congested areas using A*.

3.1 Learning Setting

The choices a vehicle takes at each intersection are made according to a specific policy. The Nash equilibrium [18] finds a local optimum amongst all policies (using complete knowledge on future traffic demand) such that no vehicle may gain any advantage over this policy by altering its own policy, whereas the central idea of the reinforcement learning algorithm POEM [21] is to use logged data to improve an existent policy h_0.

In [21] POEM assigns a structured output to an arbitrary input based on its probability of being correct. Therefore, before applying POEM to congestion avoidance, a suitable mapping of the routing problem to a policy h_0, along with an input space X and output space Y must be modeled. Additionally, a cardinal loss feedback mapping δ is required, which serves as the reward function about all selected input/output combinations.

The input space X was chosen as $X := [0,1]^m$. Here, each $\vec{x} = (x_1, \dots, x_m)^T \in X$ represents a feature vector of (normalized) sensor measurements for a road segment. For instance, a road's density, occupancy, mean speed, vehicle count or waiting time can be used. Any value not in $[0,1]$ was scaled using min-max scaling.

The output space must be a set of suitable, structured outputs. As POEM should be applied to the problem of congestion control, a single label indicating whether a road is congested or not already provided adequate results. Thus, let $Y := \{(0), (1)\}$, where (0) indicates a road is not congested, and (1) corresponds to congestion, respectively.

The policy $h_0(Y \mid \vec{x})$ is a probability distribution over the output space. In other words,

[1] For implementation and more theoretical information on POEM, we point the interested reader to [21] and the website at `http://www.cs.cornell.edu/~adith/POEM/index.html`.

it assigns a probability to each output $\vec{y}$ given any input $\vec{x}$ based on how likely $\vec{y}$ is to be correct under conditions $\vec{x}$. Hence, predictions are made by sampling $\vec{y} \sim h_0(Y \mid \vec{x})$. The goal of POEM is then to improve this policy. Initially no such policy exists for the constructed input and output spaces. This is a common problem when applying POEM. Therefore, a default policy is used (compare [21]). Let $h_0(\vec{y} \mid \vec{x}) := 0.5$, meaning both labels are assigned a probability of 0.5 for all $\vec{x}$.

Lastly, in order to improve an existing policy, POEM requires a cardinal loss feedback mapping $\delta : X \times Y \to \mathbb{R}$. This was achieved by applying one of the following two primitive congestion detection methods to the sensor readings: the primitive density congestion metric, $\delta_{density}$, would assume a road as congested when its density was greater than one seventh of its jam density [2]. The primitive mean speed congestion metric, δ_{speed} would assume a road as congested when its mean speed was less than ten kilometers per hour of its allowed maximum speed.

3.2 Application

In order to not only detect congestion but also reduce it, vehicles must receive frequent information updates about the current state of the road network. Then, POEM will be used to predict the next state of the road network. This information will consequently be used by vehicles to bypass roads which are deemed congested. Thus, those results must also be applied in a routing algorithm, such as Dijkstra or A*.

Let $G = (V, E, c, q)$ be a graph representing a road network. Here, c and q are the default cost and heuristic functions. Additionally, assume all vehicles have knowledge about a congestion labeling policy $h \in \mathcal{H}_{lin} \cup \{h_0\}$ [21]. When using dynamic routing, vehicles will receive updates about roads in regular intervals $T \in \mathbb{N}$. The update can then be written as $u_T : E \mapsto X$.[2] Then, when a vehicle receives update u_T it is able to predict how likely a road is to be congested during interval $T + 1$ using h.

The described model receives sensor information only about whole road segments, rather than individual lanes, which might be problematic, as congestion does not always arise on every lane equally. That challenging situation is most likely to occur at junctions where each lane will allow a vehicle to go in a different direction. We address this problem by aggregating sensor data for each connected edge pair (using a line graph of G, compare [7]). Additionally, the resulting data allows more precise congestion detection as individual turning lanes are separated in the model.

In order to bypass arising congestion, a vehicle must recalculate its route with respect to the newly received update u_T. This is achieved by increasing the weight of an edge which will likely be congested:

$$p^0_{(e_1,e_2)} := h((0) \mid 0.5u_T(e_1) + 0.5u_T(e_2)), \tag{1}$$

$$c' : E^2 \to \mathbb{R}_+,\ (e_1, e_2) \mapsto \frac{c(e_2)}{p^0_{(e_1,e_2)}}. \tag{2}$$

The denominator shows the previously mentioned aggregation of sensor data. For notational simplicity c' is defined for all elements of E^2. However, in practice only a subset of E^2 is used where e_1 is incident or equal to e_2.

The function c' calculates the new weight of an edge e_2 depending on its preceding edge it was reached by. For instance, a vehicle on an edge $e_1 = (u, v)$ would calculate the weight

[2] Here, it is assumed updates are received equally for all edges.

for edge $e_2 = (v, w)$ using $c'(e_1, e_2)$. A vehicle which starts its route on edge e_2 would use $c'(e_2, e_2)$.

Essentially, c' divides the default weight of an edge by its probability of not being congested in interval $T + 1$. This means the weight of an edge will remain almost unchanged when no congestion is expected. The increase will conversely depend on how likely congestion is to arise.

Finally, it was assumed sensor data updates are available for every road. In real-world road networks permanently installed sensors are much more scarcely distributed throughout the network. This problem can be partly alleviated by directly implementing sensors in the vehicle (e.g. using navigation applications provided by smartphones, or self-driving cars). However, some roads will still remain uncovered. Here, u_T can map to $\{0\}^m$. For the previously defined features in X (a road's density, occupancy, mean speed, vehicle count and waiting time) its dimension m would equal to 5. This will cause h to assign a probability of 0.5 to both labels (as defined by $\mathcal{H}_{lin}$ in [21]). Another solution could be to map u_T to the average of all sensor readings in an interval. Thus, uncovered roads would reflect the average state of a road network.

3.3 Logging

For POEM, no interactive control over actions is required, as it was specifically designed to learn using logged data. Hence, with respect to the previously defined setting, POEM requires a dataset:

$$D := \{(\vec{x}_i, \vec{y}_i, \delta_i, p_i) \mid i \in \mathbb{N}_{\leqslant n}\},\ p_i = h(\vec{y}_i \mid \vec{x}_i)\,. \tag{3}$$

This dataset will be created during the logging phase. All edges are assigned weights using c' and routes are calculated using an implementation of A*, which produces shortest routes for any admissible heuristic. Additionally, POEM is initially applied using the default policy h_0, which will scale all weights equally by a factor of two. The scaling will not affect A*, meaning no route changes will occur, which in turn simplifies learning on previously collected data.

The data itself can either be collected by each vehicle or a centralized authority monitoring each vehicle. For both approaches a data entry cannot be created before any feedback is available. Thus, intermediate results must be cached.

First the aggregated feature vector $\vec{x}_i$ is logged. The respective label $\vec{y}_i$ with its corresponding probability p_i are then determined using:

$$\vec{y}_i = \begin{cases} (0), & h((0) \mid \vec{x}_i) > 0.5\,, \\ (1), & h((1) \mid \vec{x}_i) > 0.5\,, \\ random((0),(1)), & otherwise\,. \end{cases} \tag{4}$$

Here, $random((0),(1))$ means a label is chosen randomly, uniformly distributed. Lastly, the feedback is logged using either $\delta_{density}$ or δ_{speed}. The respective results will inherently depend on the previously chosen label.

In the next section we test the performance in three cases: complete knowledge, stationary sensors, and moving sensors.

4 Experiments

The deployment of our self-organizing routing algorithm in an urban area could be done in two ways. Either the data of an existing stationary traffic information system is used (e.g. a

SCATS [10] system) and fed into a navigation platform that can be used by the citizens. The other option is to turn vehicles directly into sensors and retrieve segment-wise statistics on travel-time, density and traffic flow directly from the navigation app. In the latter case, one might be worried about individual privacy because mobility statistics are recorded centrally, however recent work [12] provides an approach to protect individual privacy in this case using homeomorphic encryption. This approach encrypts the data such that it still allows for analysis on the cryptotext but just the result can be decrypted. In the following we will test these two deployment settings using stationary and moving sensors and compare it to Nash equilibrium and uninformed routing.

For comparability of experiments with different routing algorithms it is essential to guarantee the same traffic demand (i.e. origin/destination pairs) over time. For repeatability of the same origin/destination setting, we perform analysis with a microscopic traffic simulator, SUMO [11]. The simulator models individual vehicles (on a microscopic level, so it controls also acceleration and deceleration) and is largely applied in traffic simulation and applications. It allows us to control traffic demand and provides us complete knowledge on the performance of the street network and on the routing performance. In contrast to arbitrary toy experiments, we aim at modeling sound traffic scenarios, thus, we use an open simulation scenario in the city of Luxembourg [3] which enables reproduction of 24 hours of mobility in this city.

The common procedure of SUMO is to generate the route of each vehicle before the simulation starts, which is why its live routing capabilities are rather limited. However, SUMO provides the *Traffic Control Interface* (TraCI), a network interface which allows full control over the current simulation. We used this to implement a Java application (SUMO-CA) which simulates a central authority. In order to calculate vehicle routes, SUMO-CA loads a road network and converts it to a directed, weighted multi-graph. When running a simulation, SUMO-CA will receive and parse sensor measurements in regular intervals. This information is utilized to predict the next state of the road network using POEM (compare Section 3). Finally, those results are used to update vehicle routes as shown in Section 3.2. SUMO-CA is additionally capable of logging the dataset discussed in Section 3.3.

Unless stated otherwise, each experiment will start at 7:45 o'clock (simulation time) and runs over a period of roughly 35 minutes, or exactly 2048 seconds. The reason why this particular window was chosen is that roads generally are more susceptible to congestion during rush hour. Additionally, a size of 2048 seconds allows rerouting intervals to be easily scaled using a factor of two. Finally, in order to create more realistic jams on arterial roads, SUMO was set to scale the original demand by a factor of 1.3.

4.1 Measuring Relative-Weighted Difference

Evaluating vehicle detours is problematic. Neither absolute nor relative differences will adequately represent measured detours. The reasoning behind this is that long routes will allow longer, absolute detours, whereas, short routes will allow longer, relative detours. Hence, a different metric is required. We propose usage of the weighted relative detour as follows.

Let $y_A, y_B \in \mathbb{R}^*_+$ be arbitrary measurements of one vehicle when algorithms A and B were applied respectively. Then weighted relative detour diff_{rw} will then calculate the relative difference, while at the same time weighting it using the absolute difference.

$$\mathit{diff}_{rw}(y_A, y_B) := |y_A - y_B| * \frac{y_A - y_B}{y_A + y_B} \quad (5)$$

4.2 Charts

Various charts present the evaluation results.[3] The x-axis shows multiple methods which were evaluated. The baseline is a uninformed uniform cost search (UCS), where each road was assigned its static, default weight and every vehicle chooses its path individually by A*. In this case congestions are likely to appear. Next, our approach with all evaluated rerouting interval sizes is presented. Lastly, a Nash equilibrium (NASH) is shown as a baseline.

On the y-axis we use different metrics. Vehicle throughput is measured in number of cars that reach their goal within the simulation time. Edge travel time, trip detour duration, and trip wait time duration are, if not stated otherwise, denoted in seconds. Edge travel time is a traffic network indicator and denotes the travel time per street segment. Trip detour duration highlights the deviation of each traveled trip in comparison to the UCS routing. The wait time is the time the vehicle is actually waiting in a jam.

In order to adequately present distributions box plots [23, 17] were used. Here, boxes represent the lower, middle and upper quartiles, whereas whiskers will represent the second and 98th percentiles. The outliers were omitted in the graphs.

4.3 Experiment One – Complete Knowledge

This experiment will route 100%, 75%, 50% and 25% of vehicles, chosen randomly, uniformly distributed, using 100% of available live sensor data. In other words, every road is equipped with a permanent sensor, which measures its vehicle count, average speed, occupancy, density and waiting time. This represents the best case scenario regarding information availability. The feedback was created using $\delta_{density}$.

The results show that vehicle throughput increases considerably, even at a usage rate of 25%, which can be seen in Figure 2. It additionally confirms that long update intervals may cause more congestion at high penetration rates. This is particularly visible in Figures 2, 3 and 4.

4.4 Experiment Two – Stationary Sensors

The results in previous experiment one were achieved by placing a sensor on every road. In real-world, sensors are much more scarcely distributed throughout the network [10]. This scenario with stationary traffic sensors is evaluated in experiment two and described in this section. Here, we evaluate sensor coverage of 25% or 10% of the roads. The locations were chosen randomly, uniformly distributed. Just like in experiment one, each sensor measures every vehicle. However, just 40% of vehicles will receive navigation updates, which is a more attainable penetration rate of navigation systems.

The results of experiment two, depicted in Figures 5, 6 and 7, reveal that even at lower penetration rates, the POEM algorithm successfully labels congested roads. Although the travel time per edge does not decrease as noticeable as it did in experiment one, road users still have a time-wise advantage. The results could possibly be improved by placing sensors not evenly throughout the road network, but rather around congestion prone areas. This sensor placement is subject to future research. Here we assume that the intelligent traffic system (pertaining the traffic loops) is already installed in the city (as is in most major cities) and situation of the loops cannot easily be altered.

[3] The charts are placed at the end of the paper in two-column layout.

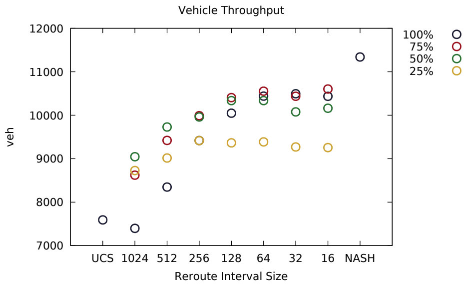


Figure 2 This chart shows results of experiment one. For an interval size of 64 seconds, throughput increases by over 50% when 100%–50% of vehicles were rerouted. Interestingly, an interval size of 1024 seconds noticeably decreased throughput when 100% of vehicles were rerouted. Here, vehicles are rerouted only once and most likely chose similar diversions on small byways, which creates more congestion.

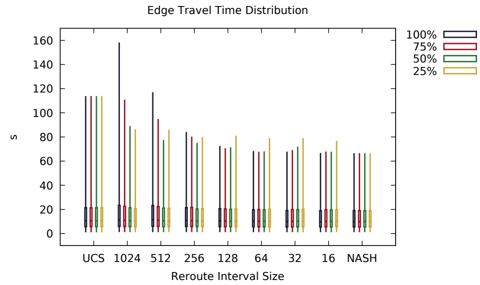


Figure 3 This chart shows results of experiment one. The noticeable increase in the 98th percentile for a routing interval of 1024 seconds where 100% of vehicles were routed coincides with results in chart 2. The same applies to the gradual decrease in travel times per edge.

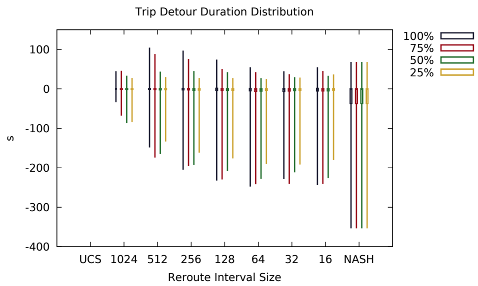


Figure 4 This chart shows results of experiment one. Here, the relative, weighted detours with respect to UCS are presented. It shows most vehicles are unaffected. However, it also shows the benefits considerably outweigh the drawbacks.

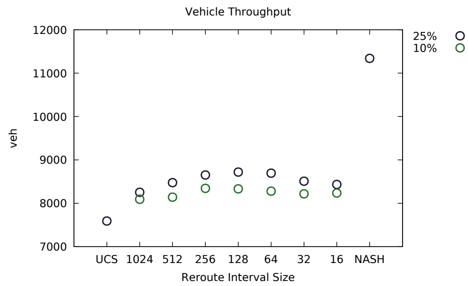


Figure 5 This chart shows results of experiment two. The chart shows that even with sensor data collected at only 10% of all roads, throughput was increased by as much as 10%.

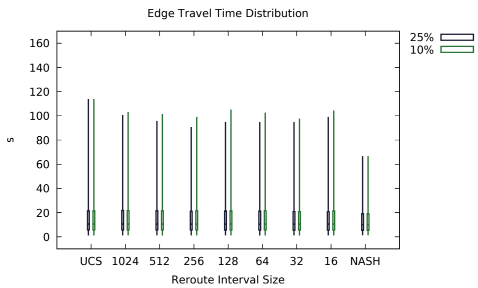


Figure 6 This chart shows results of experiment two. Although a general decrease in travel time per edge can be seen, it is not as noticeable as it was with a sensor on each road, shown in figure 3. However, this can likely be improved by placing sensors on particularly congested roads.

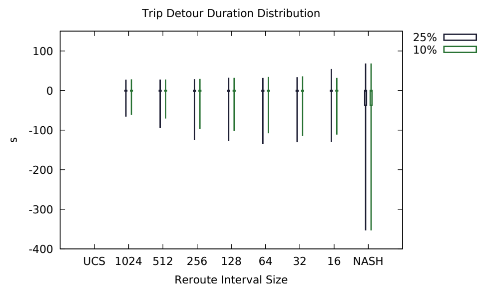


Figure 7 This chart shows results of experiment two. Here, the relative, weighted detours with respect to UCS are presented. Again, the benefits outweigh the drawbacks.

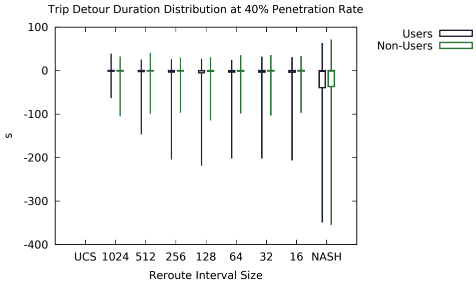


Figure 8 This chart shows results of experiment three. The users of such a navigation system have a clear advantage over non-users compared to UCS. However, non-users also benefit noticeably.

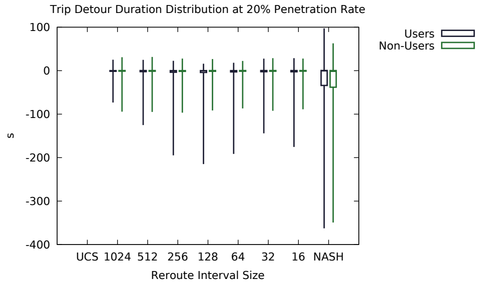


Figure 9 This chart shows results of experiment three. It can be seen that a penetration rate of 20% still provides users with similar advantages. However, performance is considerably lower for smaller interval sizes.

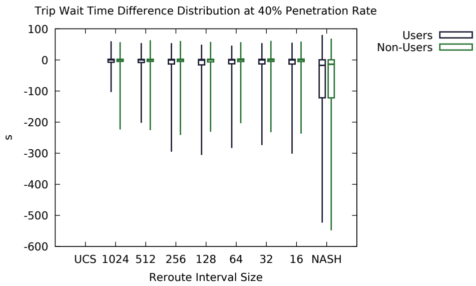


Figure 10 This chart shows results of experiment three. For smaller interval sizes, a user's waiting time will generally decrease more than that of a non-user. However, it is considerably outperformed by a Nash equilibrium.

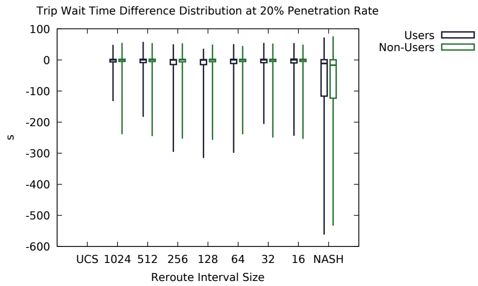


Figure 11 This chart shows results of experiment three. The chart shows that even at a lower penetration rate of 20%, it is still possible for users to outperform non-users with respect to waiting time.

4.5 Experiment Three – Moving Sensors

The sensors used in previous experiments measure every vehicle on their respective position. This information could be gathered using one of many permanently installed sensors, such as an induction loop. Alternatively, smartphone navigation applications can be used as sensors. The measured data will be incomplete, as its only collected for a subset of roads and vehicles. The incompleteness results in erroneous measurements of density values. However, measured vehicle mean speeds will remain largely unaffected. Hence, δ_{speed} was used as feedback. The experiment will evaluate penetration rates of such applications of 40% and 20%.

For vendors of such applications, routing performance regarding all users is not particularly interesting. Their interest mostly focuses on how great of an advantage users will have compared to non-users. This is why the focus lies primarily on those results.

Figures 8 and 9 show that, compared to UCS, users and non-users of such an application would benefit from its use. However, users will have a considerably greater advantage. Figures 10 and 11 show that, when looking only at medium sized intervals, time spent waiting due to congestion or traffic lights also decreases more for users. Generally, larger intervals perform poorly compared to UCS. Here, many vehicles most likely chose similar diversions and did not distribute evenly throughout the network. In turn more congestion is created on low-capacity roads.

5 Discussion and Future Work

Previous experiments revealed that the application of reinforcement learning to the routing problem is beneficial. All our experiments highlight that, by usage of our self-organizing routing regime, performance of the traffic network is increased and occurrence of traffic jams are reduced. We used regular update intervals at which the route could become updated, but this does not imply that in each step the route is altered. However, in order to achieve acceptance of the users, the travel times per user have to be reduced. Future research has to show whether or not the models learned in one geographic region could be transferred to another; we plan evaluation in the city of Cologne. In this case the demand data and road network is provided by TAPAS Cologne [24]. However, the road network quality in this scenario is considerably worse than that of Luxembourg, as it was not manually revised, but rather is a raw *OpenStreetMap* import.

High performing navigation systems have many advantages. An individual person benefits directly, as more precise predictions would shorten travel times and simplify travel planning. This also applies to logistics providers, which could optimize routing schedules and thus increase overall performance. These individual advantages would then extend to the community. Self-organized routes decrease overall traffic volume and congestion, which would serve not only the people, but also the environment [14].

References

1. Jean-Yves Audibert and Rémi Munos. Introduction to bandits: Algorithms and theory. *ICML Tutorial on bandits*, 2011.
2. Harideo Chaudhary. Application of the theory of a single first order equation to traffic flow. *Journal of the Institute of Engineering*, 9(1):175–181, 2014.
3. Lara Codeca, Raphaël Frank, and Thomas Engel. Luxembourg SUMO traffic (LuST) scenario: 24 hours of mobility for vehicular networking research. In *Vehicular Networking Conference (VNC), 2015 IEEE*, pages 1–8. IEEE, 2015.
4. Serdar Çolak, Antonio Lima, and Marta C González. Understanding congested travel in urban areas. *Nature communications*, 7, 2016.
5. Edsger W. Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.
6. Robert Geisberger, Peter Sanders, Dominik Schultes, and Daniel Delling. Contraction hierarchies: Faster and simpler hierarchical routing in road networks. In *International Workshop on Experimental and Efficient Algorithms*, pages 319–333. Springer, 2008.
7. F. Harary and R. Norman. Some properties of line digraphs. *Rendiconti del Circolo Matematico di Palermo*, 9(2):161–168, May 1960. doi:10.1007/BF02854581.
8. Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.
9. S. P. Hoogendoorn, P. H. L. Bovy, and W. Daamen. Microscopic Pedestrian Wayfinding and Dynamics Modelling. In M. Schreckenberg and S. D. Sharma, editors, *Pedestrian and Evacuation Dynamics*, pages 123–155, 2002.
10. Dermot Kinane, François Schnitzler, Shie Mannor, Thomas Liebig, Katharina Morik, Jakub Marecek, Bernard Gorman, Nikolaos Zygouras, Yannis Katakis, Vana Kalogeraki, et al. Intelligent synthesis and real-time response using massive streaming of heterogeneous data (insight) and its anticipated effect on intelligent transport systems (its) in dublin city, ireland. *ITS, Dresden, Germany*, 2014.

11 Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of SUMO – Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012.

12 Thomas Liebig. Privacy preserving centralized counting of moving objects. In Fernando Bacao, Maribel Yasmina Santos, and Marco Painho, editors, *AGILE 2015*, Lecture Notes in Geoinformation and Cartography, pages 91–103. Springer International Publishing, 2015. doi:10.1007/978-3-319-16787-9_6.

13 Thomas Liebig, Nico Piatkowski, Christian Bockermann, and Katharina Morik. Dynamic route planning with real-time traffic predictions. *Information Systems*, 2016.

14 Ruilin Liu, Hongzhang Liu, Daehan Kwak, Yong Xiang, Cristian Borcea, Badri Nath, and Liviu Iftode. Themis: A participatory navigation system for balanced traffic routing. In *2014 IEEE Vehicular Networking Conference (VNC)*, pages 159–166. IEEE, 2014.

15 Jakub Mareček, Robert Shorten, and Jia Yuan Yu. Signalling and obfuscation for congestion control. *International Journal of Control*, 88(10):2086–2096, 2015.

16 Jakub Mareček, Robert Shorten, and Jia Yuan Yu. r-extreme signalling for congestion control. *International Journal of Control*, pages 1–13, 2016.

17 R. McGill, J. W. Tukey, and W. A. Larsen. Variations of Box Plots. *The American Statistician*, 32(1):12–16, 1978.

18 John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

19 Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

20 Marco Stolpe, Thomas Liebig, and Katharina Morik. Communication-efficient learning of traffic flow in a network of wireless presence sensors. In *Proceedings of the Workshop on Parallel and Distributed Computing for Knowledge Discovery in Data Bases (PDCKDD 2015)*, CEUR Workshop Proceedings, page (to appear). CEUR-WS, 2015.

21 Adith Swaminathan and Thorsten Joachims. Batch learning from logged bandit feedback through counterfactual risk minimization. *Journal of Machine Learning Research*, 16:1731–1755, 2015.

22 Remi Tachet, Paolo Santi, Stanislav Sobolevsky, Luis Ignacio Reyes-Castro, Emilio Frazzoli, Dirk Helbing, and Carlo Ratti. Revisiting street intersections using slot-based systems. *PloS one*, 11(3):e0149607, 2016.

23 J. W. Tukey. *Exploratory Data Analysis*. Number 1 in Exploratory Data Analysis. Addison Wesley Publishing Company, 1970.

24 Sandesh Uppoor, Oscar Trullols-Cruces, Marco Fiore, and Jose M. Barcelo-Ordinas. Generation and analysis of a large-scale urban vehicular mobility dataset. *IEEE Transactions on Mobile Computing*, 13(5):1061–1075, 2014.

Global Landmarks in a Complex Indoor Environment

Cristina Robles Bahm[1] and Stephen C. Hirtle[2]

1 **Information Systems Technology, La Roche College, Pittsburgh, PA, USA**
cristina.bahm@laroche.edu

2 **School of Computing and Information Sciences, University of Pittsburgh, Pittsburgh, PA, USA**
hirtle@pitt.edu

Abstract

Wayfinding in complex indoor environments can be a difficult and disorienting activity. Many factors contribute to this difficulty, including the variable number of floors and half-floors paired with many different and often unpredictable ways to get from one floor to another. In order to explore how the spatial information of floor to floor transitions is represented cognitively, a user study was conducted at the Carnegie Museums of Art and Natural History that drew on experienced participants from the Visitor Services Department. The participants were asked to give wayfinding descriptions to and from several landmarks in the museums with the majority of the routes spanning multiple floors. It was found that floor to floor transition points were often represented as landmarks with notable locations in the Museums being represented with both functional and referential aspects. A functional aspect of a floor to floor transition points meant that its purpose in the wayfinding description was to provide a means to get from one floor to another. A referential quality meant that a floor to floor transition points was simply an indemnity and did not serve as a way to move vertically through the environment. This finding informs the discussion on global landmarks and their representation and salience in large complex indoor environments.

1998 ACM Subject Classification I.2.1 Applications and Expert Systems, I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Navigation, wayfinding, indoor environments

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.18

1 Introduction

Human beings engage in wayfinding on a daily basis through a variety of indoor, outdoor, and transitional spaces [13]. If the final destination is familiar, then one often knows the path to take and can do so typically without complications. Conversely in a difficult environment, it is useful to determine why one gets lost and how this can be prevented in the future [4]. For this reason, large complex locations become perfect places to study since it is in these locations that wayfinding difficulties are likely to arise [11].

Environments such as large museums, large libraries [15], and large convention centers [12] present a unique and interesting set of wayfinding difficulties that require a distinct set of heuristics to understand fully [24]. Many aspects of large complex indoor environments make it difficult to "get one's bearings" when attempting to get from point A to point B [4, 11, 12]. One difficulty is that staircases, or floor to floor transition points, are often not depicted well on wayfinding aids. Battles and Fu [2] examined a variety of wayfinding strategies that are

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 18; pp. 18:1–18:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

adopted by travellers using a schematic map of a multilevel building. Frankenstein et al. [7] showed how the role of background knowledge is used to evaluate indoor landmarks. Other work in this area includes the examination of individual differences in indoor wayfinding abilities using space syntax as a tool to measure the complexity of the space [12].

This study blends the idea of complex indoor environments and transitional spaces by examining the way spatial information, and in particular the floor to floor transition points, is represented in cognitive maps in the context of global landmarks. Since this work is focused on an indoor environment, we consider how the floor to floor changes might be represented as a type of transitional space in the cognitive maps of participants familiar with the environment. In this context, we define a floor to floor transition point as a space where a traveller is neither on one floor or another, but somewhere in a transitional area between two coherent spaces. Part of the difficulty of these spaces lies in the fact that when individuals leave the transitional area, their direction of movement may be the same as when they entered or might differ by any number of degrees, depending on the number of switch-backs. As such, staircases and other floor to floor transitions are important to examine in detail, given that they are often points that people find confusing [12, 15].

In order to examine floor to floor transition points, we look at the wayfinding descriptions that might be given to patrons by the staff at a public museum. The primary goal of this study is to examine the cognitive maps formed by employees who are familiar with an environment and, in particular, the role of floor to floor transition points. Thus, this research adds to the existing literature by providing insights into the internal representations of floor to floor transition points in complex indoor environments space, specifically in the context of global landmarks.

2 Research design

2.1 Study environment

The environment chosen for the study was the Carnegie Museums of Art and Natural History, which consist of two contiguous buildings, one built in 1895 housing the Museum of Natural History and a second adjoining building built in 1974 housing the Museum of Art. Total area for the museums is approximately 45,900 square meters. The attendance per year is approximately 330,000 visitors of all ages. The floor design of the museums, as described by the Head of Visitor Services is “a maze.” This environment was chosen because it is a large complex indoor space with several floor to floor transition points that do not connect floors in predictable ways.

In addition to the difficult floor to floor transition points, this building is also difficult to navigate for several additional reasons [11], including a lack of visual access, difficulty in creating a mental map, and the unpredictable layout of the floors and hallways. The Head of Visitor Services at the Carnegie Museums of Art and Natural History gave further insight into what he perceives the problems to be with wayfinding in the museums. Below are the reasons he cited for why the museums are difficult to navigate.

Multiple “half” floors: One of the challenges with wayfinding in the museums is the number of half floors throughout the space. For example, visitors often enter through the back of the museums because it is the entrance nearest to the parking garage. However, this entrance lies on a landing between the lower (basement) level and first (main) floor, which makes it difficult to represent on wayfinding aids. Often, the back entrance is shown as being on the first floor.

Both museums housed in one building: In addition to the size and complexity of the environment, the building houses both the Museum of Natural History as well as the Museum of Art. The experience designed to be a singular one since both museums overlap physically, but in reality people usually come to visit one or the other. This makes it difficult to communicate to patron's ideas about the space such as that you have to go through the Museum of Art to get to the Museum of Natural History.

No distinct entrance: Lastly, the museums lack a distinct main entrance. The most used entrance to both museums is the rear entryway because it comes from the parking lot and is located behind the Museum of Art. Because of the proximity of this entrance to the Museum of Art, patrons looking for the Museum of Natural History often get lost trying to find a distinct entrance for the Museum of Natural History regardless of signage indicating where the entrance is. In addition, the entrance is located near an entrance for employees and school groups. Often patrons who intend to enter through the back entrance of the museums end up entering through the "employee only" entrance.

These difficulties suggest that the Carnegie Museums of Art and Natural History can provide a rich study space in which to explore the role of floor to floor transition points in cognitive maps. In particular, the complex floor plan, multiple "half floors" and difficult mental map construction make it a rich environment for the study of indoor navigation.

2.2 Participants

Rather than examining the mental maps of the visitors to the museum, this study focused on the employees in the Visitors Services Department. This group is familiar with the space and is often tasked with working at the various help desks throughout the museums where they aid patrons in finding their way around the museums. Because of this experience, they are likely to have a robust internal representation of the environment. More importantly, they are accustomed to giving wayfinding descriptions that include just the public spaces and are communicated in ways that visitors to the museum would understand. 20 individuals participated in the study, 10 men and 10 women ranging in age from 19 to 77 years (SD = 15.15 years). At the time of the study they had been employed at the museums an average of 31.7 months (SD = 41.34 months).

2.3 Data collection and analysis

Participants were seated and asked to give 22 wayfinding descriptions from 17 origin and destination locations/landmarks in both verbal and sketch map form. Half of the participants were asked to give route descriptions 1–11 in sketch map form and 12–22 in verbal form. The other half gave descriptions 1–11 in verbal form and 12–22 in sketch map form. All participants were videotaped and instructed to give the description as if they are giving directions to a patron who was not familiar with the environment. In the sketch map portion of the study, participants were given a blank piece of paper and were asked to draw the path that would take the patron from the origin to the destination on the provided paper. In the verbal description portion, participants were asked to verbally give their descriptions. Participants were free to imagine the direction they were facing and generally used left/right/up/down as primary directional terms. Participants did not have access to the museum's maps during the study and were not corrected if the wayfinding description they gave was not correct or included the closed area.

Participants were then asked to complete a Santa Barbara Sense of Direction Scale [9] to measure their individual spatial ability. Finally, the study concluded with a map placement activity. During this activity, participants were given a copy of the current maps for the museum with the labels of all the locations/exhibits removed. Participants were asked to place 20 exhibits in their correct floor and location. The exhibits chosen included those that were the beginning and ending exhibits in the wayfinding portion of the study.

2.4 Sketch map analysis methods

Each sketch map provided during the study was assessed for accuracy and complexity. In order to assess accuracy the sketch map was compared to the real environment [16, 20]. To assess whether the placement of a landmark is accurate two criteria must to be met:

1. the landmark appears correctly in the sequence of landmarks encountered along the wayfinding description
2. the path connecting two landmarks accurately reflects any turns that would need to be taken in order to adequately get from Landmark A to Landmark B.

Only sketch maps that met the requirements for accuracy were further analyzed. After removal of inaccurate sketch maps the dataset consisted of 168 sketch maps. Each sketch map was classified and sorted into one of the sketch map complexity types as specified by Appleyard [1] with the purpose being to assess the amount and quality of information in the cognitive map of the participants. According to this method, the complexity of a sketch map can be classified as containing either sequential elements or spatial elements, and by the amount of detail. Sequential maps can be further ranked in terms of complexity as (1) Fragment maps, (2) Chain maps, (3) Branch and Loop maps, (4) Network maps. Spatial maps can be ranked in complexity as (1) Scattered, (2) Mosaic, (3) Linked, (4) Patterned. As the rating goes up so does the complexity, meaning that a patterned sketch map shows more complexity than a scattered map for the spatially dominated maps. For sequentially dominated maps, network maps are more complex than a fragment map.

The frequencies of landmarks, path segments, and nodes [20] in the sketch map data were also counted. The purpose of this analysis being to allow for a measurement of which landmarks are important and which routes contain the most data. Landmarks with higher frequencies across all descriptions are likely the most important landmarks in the dataset.

2.5 Verbal analysis methods

Verbal data was transcribed and coded by the researcher. Landmark based theme analysis methods were applied to all landmarks mentioned, not just floor to floor transition points. The purpose of this analysis was to begin to determine how the coarseness of the spatial information communicated by participants linguistically compares to the coarseness the information in their representations of space [10]. The analysis of verbal and horizontal prepositions were examined as well as verbalizations connected to any mentioned landmarks. For example, if a participant says "The room is to the left of the big statue" this will be coded as the horizontal preposition "Left" with the room being related to the big statue.

Verbalizations focused on axial parts, distance of regions, and paths and trajectories were also assessed. Verbalizations that showed a relationship as axial parts would show a connectedness between the landmarks and often a symmetrical representation of importance for the landmarks on a cognitive map. Word such as "on top of" or "in front of" would connect two landmarks as axial parts. Distance of regions verbalizations showed a relationship between a pair of landmarks and their distance to each other. Words such as "near" and "far"

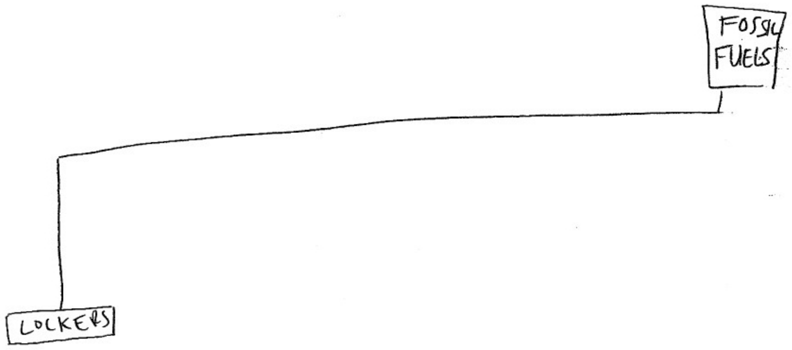


Figure 1 Sketch map from route 1 showing a less complex route.

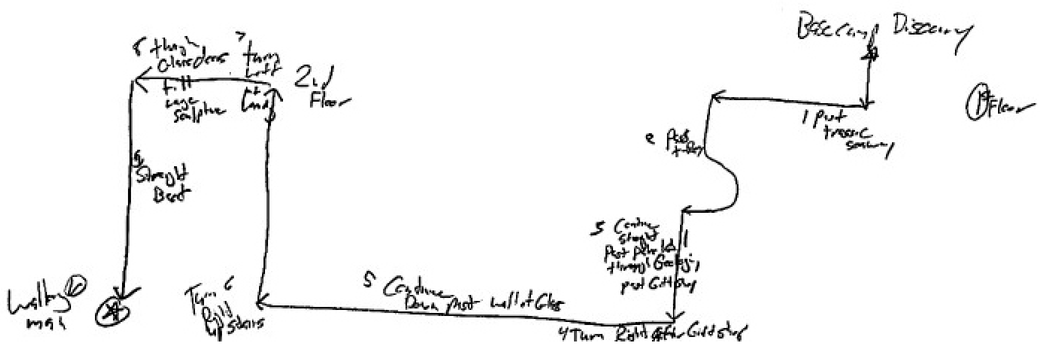


Figure 2 Sketch map from route 18 showing a more complex route.

illustrate a conceptual distance between two landmarks. Paths and trajectories verbalizations showed whether or not two landmarks were considered to be on the same path or in the same trajectory.

Verbalizations were also analyzed for the relationships between two landmarks. This method focuses on prepositions by taking into account figure and ground objects in addition to the preposition itself. Consider the following example from Landau and Jackendoff [14]:

- The bike (figure) is near the garage (ground object).
- The garage (figure) is near the bike (ground object).

Although these two sentences communicate a spatial relationship between two objects, they have different figures and ground objects making their implication about the importance of the two objects different. Ground objects usually have, properties that facilitate search and “in many contexts, they should be large, stable, and distinctive” [14].

3 Results

3.1 Sketch map analysis results

Each of the 20 participants completed sketch maps for 11 route descriptions, resulting in 220 sketch maps available for analysis. The level of detail varied greatly across destinations and across participants as shown by two example sketch maps shown in Figure 1 and Figure 2. These figures display two different wayfinding descriptions drawn by two different participants.

Table 1 Number of maps rated and their map type.

Overall Map Type	Map Type	Rating	Number of Maps
Sequential	Fragmented	1	15
Sequential	Chain	2	66
Sequential	Branch and Loop	3	17
Sequential	Netted	4	6
Spatial	Scattered	1	13
Spatial	Mosaic	2	30
Spatial	Linked	3	39
Spatial	Pattered	4	29

Sketch map complexity was analyzed using the methods described previously. The number of sketch maps that met the criteria for each Appleyard [1] classification is shown in Table 1 with examples from the study shown in Figure 3. Although each type of map was shows, the most used map type was a sequential chain map.

3.2 Verbal analysis results

Due to a technology error that resulted in data loss, the verbal data from eight participants was not able to be analyzed. The remaining data included 14 participant's verbal wayfinding descriptions for nine routes making the total number of descriptions collected 126 descriptions.

Not surprisingly, in terms of horizontal and vertical prepositions, verbalizations that included floor to floor transition points also were often accompanied by "up" and "down" but were the most often accompanied by the word "to." In total, 139 wayfinding descriptions given by participants across all routes contained these two words. The frequency of horizontal and vertical prepositions for all landmarks as well as the type of relationship the preposition indicates are shown in Table 2.

An analysis of verbal prepositions focused on determining figure and ground objects showed that floor to floor transition points were verbalized as ground objects in 59.9 % of the verbalizations. This slight preference for verbalizing a floor to floor transition point shows that participants may have thought of the transition points as reference points when giving wayfinding descriptions.

3.3 Vertical transitional space analysis

The verbal analysis was conducted on the subset of 14 participants also included an analysis that focused on floor to floor transition points as vertical transitional spaces. This analysis was based on the verbalizations found to be indicative of an indoor/outdoor transitional space introduced by Kray et al. [13]. The purpose of this analysis was to determine the importance of floor to floor transition points as landmarks in the study space with the potential representation of a floor to floor transition point as being type of transitional space. In this analysis, we extend the original theory by examining the use of transitional words by looking at their frequency in the verbal descriptions. Table 3 shows that the grand staircase was the most commonly mentioned vertical transition, but that another twelve locations were mentioned by at least one participant, which included seven staircases, four elevators, and one ramp.

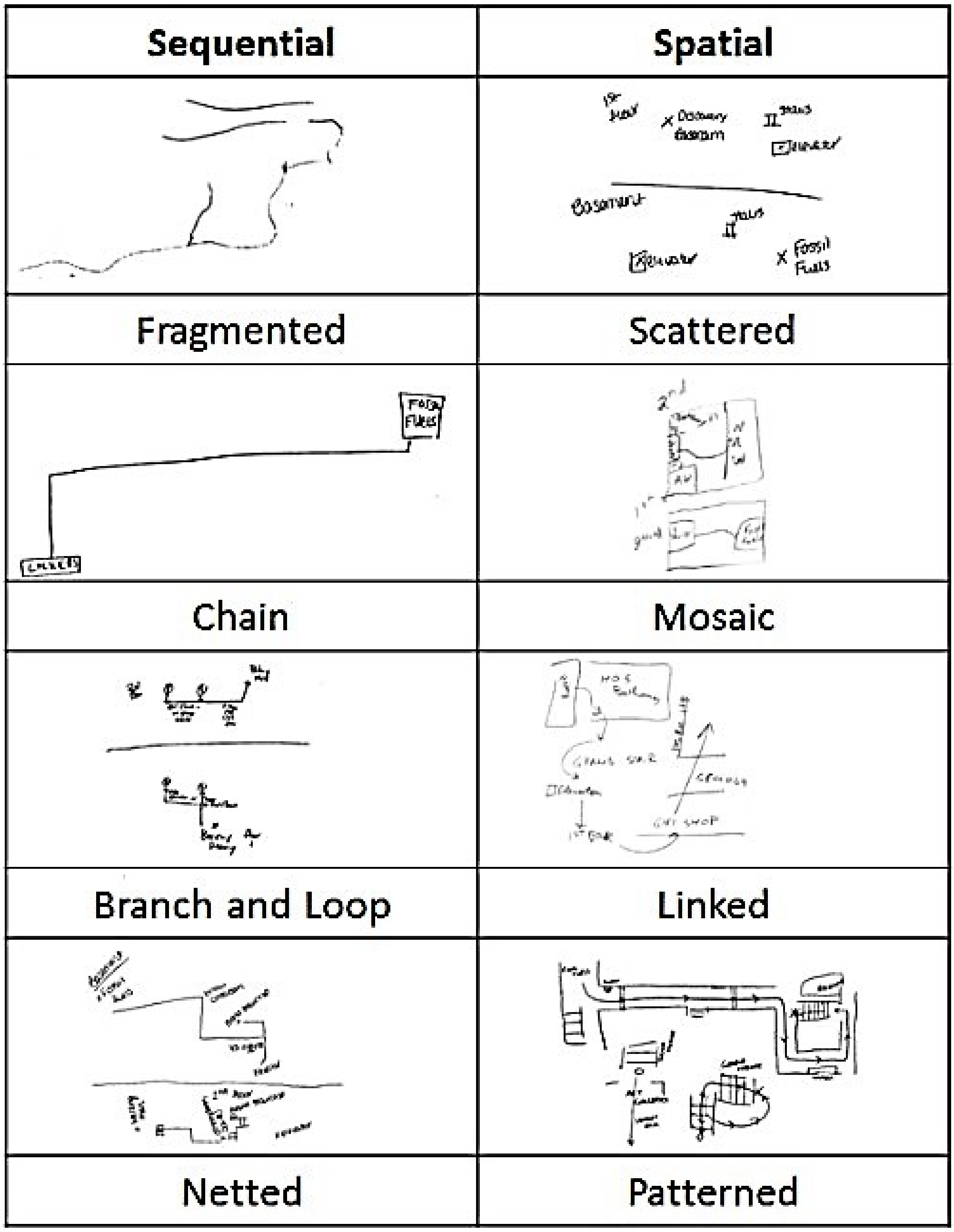


Figure 3 Examples of all types of maps based on Appleyard (1970).

Table 2 Frequency of verbalizations to describe paths.

Word	Word Type	Frequency
To	Verbalization-Paths-and-Trajectories	281
Down	Verbalization-Paths-and-Trajectories	133
Up	Vertical-Preposition	115
Left	Horizontal-Preposition	92
Right	Horizontal-Preposition	72
In	Verbalization-Relative-Distance-of-Region	71
From	Verbalization-Paths-and-Trajectories	64
On	Verbalization-Relative-Distance-of-Region	33
End	Verbalization-Axial-Parts	21
Around	Verbalization-Paths-and-Trajectories	17
Front	Verbalization-Axial-Parts	12
Top	Verbalization-Axial-Parts	11
Towards	Verbalization-Paths-and-Trajectories	10
Bottom	Verbalization-Axial-Parts	7
Behind	Verbalization-Axial-Parts	6
Side	Verbalization-Axial-Parts	5
Along	Verbalization-Relative-Distance-of-Region	3
Over	Verbalization-Paths-and-Trajectories	2
Far	Verbalization-Relative-Distance-of-Region	2
Above	Vertical-Preposition	2
Over	Vertical-Preposition	2
Backward	Verbalization-Paths-and-Trajectories	1
Away	Verbalization-Paths-and-Trajectories	1

4 Discussion

Employees who must navigate large indoor spaces on a daily basis while providing guidance to others have likely encoded noted locations, regions, and relationships into a cognitive map. Although the complexity of the cognitive maps varied as shown in Figure 3, transition points from region to region and notable landmarks were present in most representations. In line with the past literature, the floor to floor transition points at the Carnegie Museum of Art and Natural History were represented as important landmarks in the cognitive maps of participants in both sketch map and verbal tasks [6]. The Grand Staircase in particular was a floor to floor transition point that was verbalized and drawn often. The vertical transitional space analysis showed that the Grand Staircase was the floor to floor transition point verbalized the most as a possible transitional space.

From here we begin to ask what characteristics of a floor to floor transition point, particularly the Grand Staircase, makes its representation in the cognitive map of a participant distinctive? And were there any landmarks that were more distinctive than others? By defining a landmark as anything that stands out from a scene [19] this discussion explores the characteristics of floor to floor transition point representations that make them ideal candidates as global landmarks. The focus here is on the characteristics of the floor to floor transition points that allowed them to become distinct. These characteristics include:

Table 3 Frequency of floor to floor transition points on maps.

FTF Type	Name	Frequency
Staircase	Grand Staircase	40
Staircase	Back Staircase	25
Staircase	Spiral Steps	18
Staircase	Scaife Steps	11
Staircase	Library Steps	3
Staircase	Portal Steps	2
Staircase	Jane Steps	2
Staircase	HOA Steps	1
Elevator	Back Elevator	21
Elevator	Silver Elevator	21
Elevator	Rental Locker Elevator	12
Elevator	Scaife Elevator	2
Ramp	Basement Ramp	3

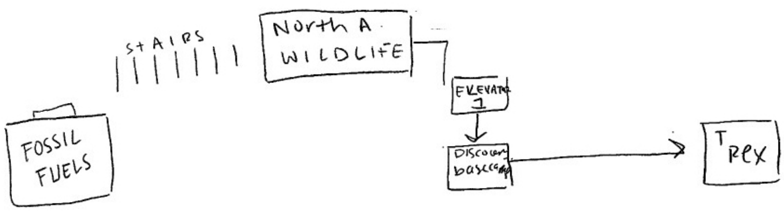


Figure 4 Example of generic unnamed stairway and unnamed elevator.

1. Descriptive names for distinct floor to floor transition points.
2. Dual representation – both functional and referential for distinct floor to floor transition points.
3. Where the floor to floor transition points lie structurally in the museums.

4.1 Descriptive names for distinct floor to floor transition points

As with landmarks in any context, floor to floor transition points were represented at varying degrees of importance in the cognitive maps of participants [18, 23]. Most wayfinding descriptions used generic floor to floor transition points such as "the stairs" or "the elevator." The verbal analysis included phrases such as *"what you're going to do is take the stairs up to discovery basecamp and make a left at the top of those stairs"* which show a generic, unnamed, communication of a floor to floor transition point. Sketch map data showed that most floor to floor transition points were thought of generically. Figure 4 shows the generic representation of a stairway as well as an elevator, which are used for movement of the traveler. Figure 5 shows a stairway used for movement, but also a Spiral Staircase which is the landmark to orient the traveler along the path.

Although most floor to floor transition points were mentioned generically, some were explicitly named. These were the Grand Staircase, the Spiral Staircase, the Silver Elevator, and the Gold Elevator. These were referred to by name in several of the verbal and sketch map descriptions. It is interesting to note that in these cases participants used these specific names. This communicates a global understanding of what these landmarks were, indicating

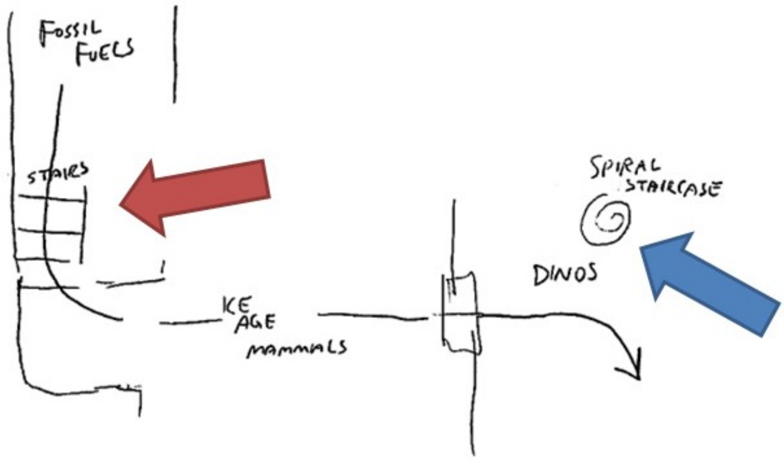


Figure 5 Figure showing functional floor to floor transition points (red arrow added) as well as referential floor to floor transition points (blue arrow added).

that they might be important. An interesting observation is that these floor to floor transition points, in particular, are visually distinguishable from other labeled entities in the museums. It is likely that this distinguishability is what makes the Grand Staircase, the Spiral Staircase, the Silver Elevator, and the Gold Elevator important landmarks in the environment [17, 21].

4.2 Dual representation – both functional and referential for distinct floor to floor transition points

A dual representation of a floor to floor transition point means that the point is represented not only as a way to get from floor to floor, but as a reference point for wayfinding in general. A functional quality of a floor to floor transition points meant that its purpose in the wayfinding description was to provide a means to get from one floor to another. A referential quality meant that a floor to floor transition points was simply an indemnity and did not serve as a way to move vertically through the environment. All floor to floor transition points were represented as being functional in at least one description. However; some floor to floor transition points that were represented referentially as well. Figure 5 shows a sketch map from the study that shows two stairways: One being included for function (red arrow) and one being a landmark (blue arrow). An interesting observation is that the referential floor to floor transition points are also given a descriptive name while the functional floor to floor transition points are generic.

In this particular environment of the Carnegie Museums of Art and Natural History, the floor to floor transition point that was most often represented as both a functional and descriptive landmark in both verbal and sketch map descriptions is the Grand Staircase. Figure 6 shows an example of the Grand Staircase being portrayed as a functional floor to floor transition point, while Figure 7 shows the Grand Staircase as a landmark floor to floor transition points.

An interesting example from the verbal analysis showing the Grand Staircase as a landmark was as follows: *"go out to the front of the building by the Grand Staircase and take the elevator down to two."* In this case the Grand Staircase is being referred to by its descriptive name, but then the participant tells the addressee to use the elevator to go down to two. This verbalization shows a deliberate instruction to use the elevator to complete

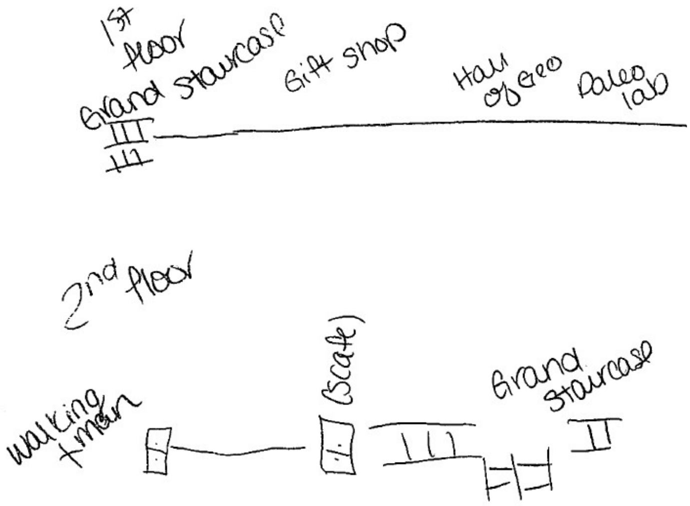


Figure 6 The Grand Staircase as a named functional floor to floor transition points, moving the traveller from the 1st to the 2nd floor.

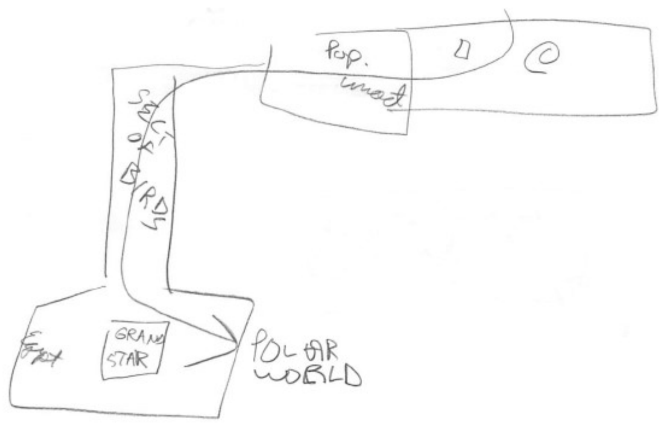


Figure 7 The Grand Staircase ("Grand Stair") as a floor to floor transition point landmark.

the function of going from floor to floor while referring to the Grand Staircase to provide orientation information.

4.3 Where the floor to floor transition points lie structurally in the museums

Where a floor to floor transition point lies in the overarching structure of the museum is important in determining the importance of the floor to floor transition points as a landmark. The literature shows that a landmark is structurally important if it is located somewhere significant in the structure of the space [21]. The three dimensional nature of the museums means that "in order to change floors in a building, for example, it is necessary to move to a location that allows vertical movement such as a staircase" [3]. This vertical movement meant that several of the important landmarks in the cognitive maps of participants were the floor to floor transition points. When applying this definition to the museums, the Grand

Staircase emerges as a particularly important landmark. The Grand Staircase spans all four floors and also sits between the natural history and art museums making it the most important landmark in this case study. Furthermore, while the Grand Staircase was used both for travel and a referent, the Spiral Staircase was used primarily as a referent. The verbal analysis shows that a landmark with the ability to be verbalized as being "down" or "up" from another landmark is of particular importance since these were the most used verbalizations after the word "to."

4.4 Towards a definition of global landmarks

Landmarks, in general, provide a structured knowledge of an environment, usually in terms of an anchor point [5]. Particular to global landmarks, they provide a point of reference for the participant, allowing for orientation and a sort of "compass" effect [22]. One of the difficulties in indoor wayfinding is the fact that it is easy to define landmarks on the local level but not on the global scale [8]. The concept of a local landmark is easily transferred to an indoor environment due to the natural chunking of spatial information in an indoor environment [12]. The unique characteristics of floor to floor transition points make them ideal candidates for global landmarks.

By examining the floor to floor transition points at the Carnegie Museums of Art and Natural History in terms of their ability to be named, their representation as either functional or landmarks, and their location in the structure of the building an idea of a global landmark begins to emerge. The Grand Staircase, the Spiral Staircase, the Gold Elevator all met the first two criteria. However, central location of the Grand Staircase made it the strongest candidate for a global landmark amongst the four.

5 Conclusions

Through the collection of suggested routes by trained staff at a large museum, this study was able to identify floor to floor transitional spaces in large complex indoor environments, which share numerous properties with traditional outdoor/indoor transitional spaces. The study also investigated the possibility of floor to floor transition points as global landmarks in indoor environments.

In addition, the study uncovered some interesting asymmetries between drawings and instructions to be explored in future work. For example, there were several ramps that took patrons to half floors, which were noted in the verbal descriptions, but rarely drawn on the maps as unique features. This difference between the sketch map and verbal descriptions supports the theories that there are differences in how we describe spaces when asked to describe them verbally or spatially [16]. However, in order to fully determine this, a full set of verbal data would need to be taken in conjunction with sketch map data. Finally, the results can give guidance in terms of automatic route generation by determining what elements would constitute the best global landmarks especially as transitioning from one area, or floor, to another in a complex indoor environment.

References

1 Donald Appleyard. Styles and methods of structuring a city. *Environment and behavior*, 2(1):100–117, 1970.

2 Andrew Battles and Wai-Tat Fu. Navigating indoor with maps: Representations and processes. In *CogSci*, 2014.

3 Simon J. Buchner, Christoph Holscher, and Gerhard Strube. Path choice heuristics for navigation related to mental representations of a building. In *Proceedings of the European Cognitive Science Conference*, pages 504–509, 2007.

4 Laura A. Carlson, Christoph Hölscher, Thomas F. Shipley, and Ruth Conroy Dalton. Getting lost in buildings. *Current Directions in Psychological Science*, 19(5):284–289, 2010.

5 Helen Couclelis, Reginald G. Golledge, Nathan Gale, and Waldo Tobler. Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7(2):99–122, 1987.

6 Michel Denis. The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de psychologie cognitive*, 16(4):409–458, 1997.

7 Julia Frankenstein, Sven Brüssow, Felix Ruzzoli, and Christoph Hölscher. The language of landmarks: the role of background knowledge in indoor wayfinding. *Cognitive processing*, 13(1):165–170, 2012.

8 Nicholas A. Giudice, Lisa A. Walton, and Michael Worboys. The informatics of indoor and outdoor space: a research agenda. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pages 47–53. ACM, 2010.

9 Mary Hegarty, Anthony E. Richardson, Daniel R. Montello, Kristin Lovelace, and Ilavanil Subbiah. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425–447, 2002.

10 P. Bryan Heidorn and Stephen C. Hirtle. Is spatial information imprecise or just coarsely coded? *Behavioral and Brain Sciences*, 16(02):246–247, 1993.

11 Stephen C. Hirtle and Cristina Robles Bahm. Cognition for the navigation of complex indoor environments. *Indoor Wayfinding and Navigation*, pages 1–12, 2015.

12 Christoph Hölscher, Tobias Meilinger, Georg Vrachliotis, Martin Brösamle, and Markus Knauff. Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology*, 26(4):284–299, 2006.

13 Christian Kray, Holger Fritze, Thore Fechner, Angela Schwering, Rui Li, and Vanessa Joy Anacta. Transitional spaces: between indoor and outdoor spaces. In *International Conference on Spatial Information Theory*, pages 14–32. Springer, 2013.

14 Barbara Landau and Ray Jackendoff. Whence and whither in spatial language and spatial cognition? *Behavioral and brain sciences*, 16(02):255–265, 1993.

15 Rui Li and Alexander Klippel. Wayfinding in libraries: Can problems be predicted? *Journal of Map & Geography Libraries*, 8(1):21–38, 2012.

16 K. Lohmann. The use of sketch maps as a basis for measures of spatial knowledge. In *Understanding and Processing Sketch Maps. In: Proceedings of the COSIT 2011 workshop. AKA Verlag, Heidelberg*, 2011.

17 Kevin Lynch. *The image of the city*, volume 11. MIT press, 1960.

18 Pierre-Emmanuel Michon and Michel Denis. When and why are visual landmarks used in giving directions? In *International Conference on Spatial Information Theory*, pages 292–305. Springer, 2001.

19 Clark C. Presson and Daniel R. Montello. Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology*, 6(4):378–381, 1988.

20 Michael J. Rovine and Gerald D. Weisman. Sketch-map variables as predictors of wayfinding performance. *Journal of Environmental Psychology*, 9(3):217–232, 1989.

21 Molly E. Sorrows and Stephen C. Hirtle. The nature of landmarks for real and electronic spaces. In *International Conference on Spatial Information Theory*, pages 37–50. Springer, 1999.

22 Sibylle D. Steck and Hanspeter A. Mallot. The role of global and local landmarks in virtual environment navigation. *Presence: Teleoperators and Virtual Environments*, 9(1):69–83, 2000.

23 Thora Tenbrink and Stephan Winter. Variable granularity in route directions. *Spatial Cognition & Computation*, 9(1):64–93, 2009.

24 Perry W. Thorndyke and Barbara Hayes-Roth. Differences in spatial knowledge acquired from maps and navigation. *Cognitive psychology*, 14(4):560–589, 1982.

A Crowdsourced Model of Landscape Preference

Olga Chesnokova[1], Mario Nowak[2], and Ross S. Purves[3]

1 **Department of Geography, University of Zürich, Zürich, Switzerland**
olga.chesnokova@geo.uzh.ch
2 **Department of Geography, University of Zürich, Zürich, Switzerland**
mario.h.nowak@gmail.com
3 **Department of Geography, University of Zürich, Zürich, Switzerland**
ross.purves@geo.uzh.ch

Abstract

The advent of new sources of spatial data and associated information (e.g. Volunteered Geographic Information (VGI)) allows us to explore non-expert conceptualisations of space, where the number of participants and spatial extent coverage encompassed can be much greater than is available through traditional empirical approaches. In this paper we explore such data through the prism of landscape preference or *scenicness*. VGI in the form of photographs is particularly suited to this task, and the volume of images has been suggested as a simple proxy for landscape preference. We propose another approach, which models landscape aesthetics based on the descriptions of some 220000 images collected in a large VGI project in the UK, and more than 1.5 million votes related to the perceived scenicness of these images collected in a crowdsourcing project. We use image descriptions to build features for a supervised machine learning algorithm. Features include the most frequent uni- and bigrams, adjectives, presence of verbs of perception and adjectives from the "Landscape Adjective Checklist". Our results include not only qualitative information relating terms to scenicness in the UK, but a model based on our features which can predict some 52% of the variation in scenicness, comparable to typical models using more traditional approaches. The most useful features are the 800 most frequent unigrams, presence of adjectives from the "Landscape Adjective Checklist" and a spatial weighting term.

1998 ACM Subject Classification I.7.0 Document and Text Processing

Keywords and phrases VGI, crowdsourcing, semantics, landscape preference

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.19

1 Introduction

The advent of new sources of spatial data, and in particular those which are generated not through a top-down, regulated process, but bottom-up, by individuals with varying backgrounds and motivations, has brought with it new opportunities for research. In particular, the advent of spatial data associated with natural language, typically in the form of tags or unstructured text provide a potential route to exploring ways in which space is described in language, albeit typically in corpora where we as researchers have very little control. The data studied in such research can be produced in a number of ways, and differing, but overlapping, definitions have been assigned to such data including those related to volunteered geographic information (VGI), crowdsourcing, user-generated content, social media, citizen science and so on [7]. These definitions are important since they have implications for the ways in which data are produced, and in turn the ways in which they can reasonably be interpreted.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 19; pp. 19:1–19:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

One obvious, and much studied, source of such data are the tags and descriptions associated with georeferenced images. Here, researchers typically assume that images and their descriptions often capture information about named locations, their properties and, occasionally, notions related to sense of place (e.g. [28, 16]). Indeed, Fisher and Unwin [8] presciently recognised this potential in 2005, stating that “GI theory articulates the idea of absolute Euclidean spaces quite well, but the socially-produced and continuously changing notion of place has to date proved elusive to digital description except, perhaps, through photography and film. (p. 6).” Nonetheless, in practice analysing text and extracting information related to place has proved challenging, and many studies have either focussed above all on exploring the properties of text related to location, with limited or no opportunities for validation, or on using counts of images as a proxy for some spatially varying phenomena and generating appropriate statistical models (e.g. [5, 36, 37]).

The act of georeferencing images typically implies that an individual wishes to relate a particular image to an event (not relevant in the context of this work) or a location. The act of producing an image however is not random, and neither is the act of choosing to share an image with others in an online source [11]. Images capturing locations presumably capture perceptually salient elements of a landscape, and thus, accompanied by their descriptions might provide us with clues as to how landscape is conceptualised and parcelled up into cognitive entities [22]. Understanding landscape, and the ways in which it is perceived is not merely an abstract research question, but one with considerable direct policy and societal relevance, since landscapes are the subject of national and international policies and regulation. Contemporaneously with the emergence of new data sources such as those described above, has been an increasing realisation in many areas of policy that there is a need to include not only top-down definitions of landscapes in policy work, but also to capture bottom-up ways in which landscapes are perceived and experienced. Even seemingly simple notions such as landscape aesthetics have proved remarkably challenging to generalise and model spatially, and although methods based in the social sciences can capture well the diversity of opinions about individual locations, they are ill-suited to characterising large regions [37].

In this paper we set out to demonstrate, through the use of two, related, datasets, how we can firstly, capture through textual descriptions, elements of a landscape which are perceived as more or less attractive across a large region. To do so, we combine descriptions of georeferenced images which are an excellent example of VGI *sensu* Goodchild [12] with a large crowdsourced data containing *scenicness* rating for more than 220000 images. We then develop and evaluate a predictive model of scenicness, which as its primary input uses text describing images, and thus aims to model scenicness as a function of language.

1.1 Related work

In the following we briefly set out related work from two key areas. Firstly, we summarise concepts related to landscape aesthetics and its assessment. Secondly, we explore examples of research which have used novel data sources to explore landscape properties in a range of ways.

Theories seeking to explain landscape perception and aesthetics typically focus on both evolutionary and cultural influences [19, 15]. Evolutionary approaches assume that preferences with respect to landscape relate to the ability of landscapes to meet human needs such as ‘prospect’ (i.e. the ability to command a landscape through sight) and ‘refuge’ (the potential to conceal oneself in a landscape) [1]. Other, related concepts include the ability to ‘make sense’ of the environment (coherence and legibility of landscapes), and ‘involvement’ or ability to function well in the environment (complexity and mystery of landscapes) [18]. Cultural influences on landscape preference are recognised in the emergence of work on landscape and language, for example, through the study of ethnophysiography [22] which notes the importance of cultural influences and the absence of universally shared landscape elements.

Irrespective of the theoretical perspective taken, typical approaches to capturing landscape perception have focussed on in-situ methods using, for example, interviews and participatory mapping [2, 27]. However, the need to be on site makes such approaches poorly suited to capturing dynamic landscape preferences over large areas, and also makes it difficult to control potential influences. Such limitations, and the simple need to generate more lab-based reproducible experiments, led to the development of approaches based around photographs of landscapes where participants can be presented with images controlling the visual field [31], seasonal changes, or introducing extra factors (e.g. presence of animals [17] or anthropogenic objects [20]).

The advent of VGI, and the realisation that such data might contain diverse, independent and decentralised information, provided opportunities to replicate previous work on geographic concepts [34], and to demonstrate that such data were a reliable source of information about landscape characteristics and the ways in which landscapes were categorised [6, 28]. In parallel, the need to generate landscape indicators related to cultural ecosystem services and landscape preferences over large areas has led some of researchers to use the position and number of images taken as a proxy indicator of landscape preference [36, 37], or to incorporate the number of individuals taking pictures [3, 11] and their origins [10]. Others have realised that the images themselves contain information central to understanding landscape preference, and have analysed image content to explore cultural ecosystem services [30]. The importance of scenicness in a policy context, and the possibilities offered by new data sources are recognised in recent work exploring the link between wellbeing and scenicness using crowdsourced data, and attempting to model scenicness using user generated content [33, 32].

In this paper we seek to build on previous work in two key ways. Firstly, in-situ and lab-based studies of landscape preference have typically worked, of necessity, with relatively small groups of participants in focussed, often coherent, landscapes. Our study, by using VGI at the scale of Great Britain, allows us to explore landscape preferences across a whole country, and to explore regional differences between such preferences. Secondly, attempts to model scenicness have typically focussed on using spatial data in some form as explanatory variables (for example number of images, elevation, number of visible pixels, landcover type, etc.). We take an approach which we argue is likely to be closer to the way in which a particular landscape is perceived, and build a model of scenicness which uses language (in the form of words and phrases extracted from written descriptions) as explanatory variables.

In the following, we first describe the datasets on which we carried out our experiment, and the steps we took in processing, analysing and modelling scenicness with these data. We then present our results, demonstrating that the words used to describe scenic areas make clear distinctions especially between scenes perceived to be more or less anthropogenically influenced. Our model of scenicness is capable of explaining about 52% of the variance in scenicness in space, which is comparable to typical state of the art approaches. We then discuss the implications of these results, before concluding with some suggestions for future research.

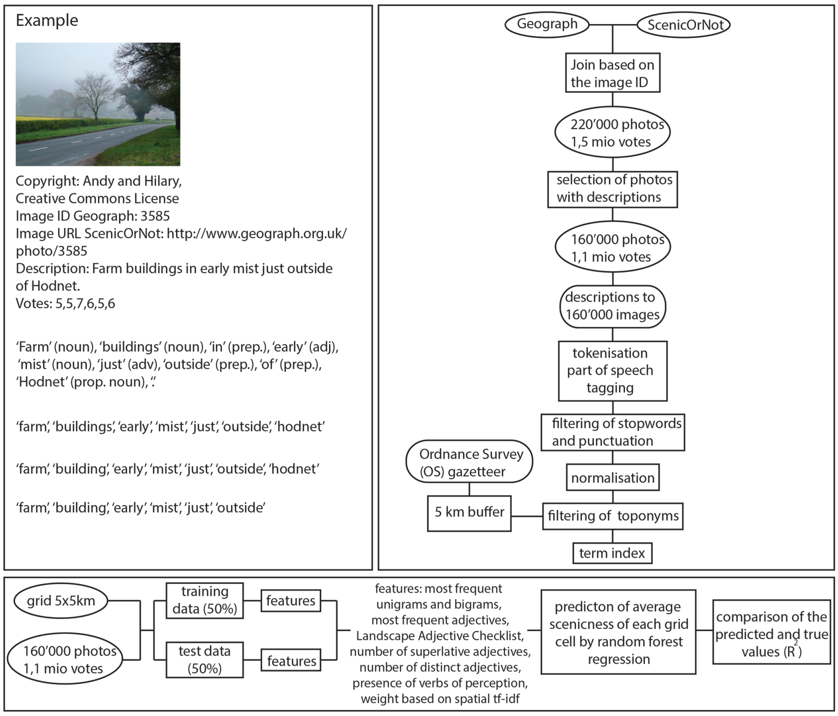


Figure 1 Steps of the data acquisition and preprocessing with an example.

2 Data and methods

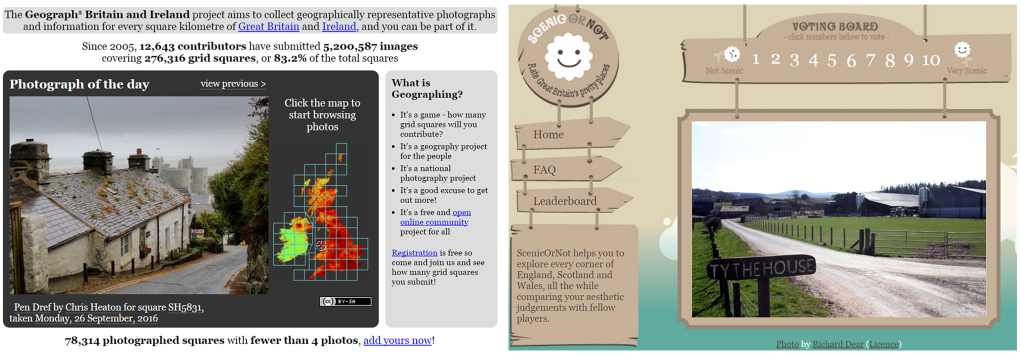
As set out above, our aims are twofold. Firstly, we wish to identify which terms are typically used with more or less scenic images, as described by votes in ScenicOrNot project and, secondly, based only on terms describing images to develop a spatially contiguous model of scenicness at the country level. In the following we describe the datasets used, and in particular aspects relevant to our work. We then set out our approach to processing the corpus, before describing the features used in producing our spatial model of scenicness. Fig. 1 gives a visual overview of the material which follows.

2.1 Data and study region

We use two unique, and related, datasets in this work. The Geograph[1] dataset (Fig. 2a) is a crowdsourced collection with more than 12000 contributors, launched in 2005, with the aim of collecting "geographically representative photographs and information for every square kilometre of Great Britain and Ireland." The project takes the form of a game, with users receiving points for uploading georeferenced images and associated descriptions, and content is moderated. The entire dataset is available under a Creative Commons Licence, and in this paper we used a version downloaded in June 2016 consisting of ca. five million images.

The ScenicOrNot[2] project (Fig. 2b) was initiated in 2009 by MySociety and is currently hosted by the Data Science Lab at Warwick Business School. The goal of the project is

[1] http://www.geograph.org.uk/
[2] http://scenicornot.datasciencelab.co.uk/



(a) Interface of the Geograph project **(b)** Interface of the ScenicOrNot project

Figure 2 Interface of the Geograph project (Copyright Chris Heaton, Creative Commons Licence) and the ScenicOrNot project, where users can rate a Geograph photograph from 1 (not scenic) to 10 (very scenic).

to crowdsource scenicness ratings using Geograph images. In contrast to Geograph, where it is reasonable to assume that users uploading images typically also took the pictures in question (and thus visited the landscape), the ScenicOrNot project is purely internet based. Participants, about whom no demographic information is collected, are presented with a series of random images, with neither associated locations or descriptions, and asked to rate them on a scale of 1 (not scenic) to 10 (scenic) for scenicness. More than 220000 Geograph images had amassed some 1.5 million votes by June 2016 in the ScenicOrNot collection.

In the following our corpus consists of the 160000 Geograph images which both have a description, and are associated with three or more votes in ScenicOrNot.

2.2 Corpus processing

Our aim in corpus processing was to explore how terms used in describing Geograph images were associated with scenicness ratings. Since our starting point are natural language captions, standard corpus processing steps were applied. In the following, we briefly describe these steps, which were, in the main, carried out using the Python-based NLTK[3] library.

Each image description was in parallel tokenised, and part of speech tagged. The tokens were then filtered for stopwords and punctuation, before being normalised by changing all tokens to lower case and reducing tokens to their lemmas. Our aim was to build a term index, with associated features, for use in exploring the semantics of scenic locations.

Since we were explicitly not interested in the names of locations, we filtered toponyms from descriptions using gazetteer look-up in a 5km window around the coordinates associated with images. We used a freely available gazetteer, based on the 1:50000 maps from the Ordnance Survey for this process. This approach aims to strike a balance between removing local toponyms, which may be the subject of considerable semantic ambiguity (e.g. does bath refer to a place to bathe or the historic city) and retaining tokens which are being used in a non-toponymic sense.

Having performed these steps we are left with a term index, where unique entries are made up of tuples containing normalised tokens (unigrams and bigrams) present once or more in a description, part of speech tagging and the images IDs with which they are associated. Since

[3] `http://www.nltk.org/`

each term can be present in one or more images, and each image is ranked three or more times, we assign an average scenicness to every term in our index. Importantly, identical tokens having different parts of speech will have different values of average scenicness. Furthermore, since we store image IDs, we also have access to all locations associated with a term, the array of votes and an overall frequency of the term, based on the number of images described using a given term. Using our term index, it is possible to generate lists of terms, ranking or filtering by, for example, average scenicness, part of speech or frequency.

2.3 Feature choice and modelling scenicness

The final step in our approach was to create a spatially contiguous model of scenicness based on our term index. We predict scenicness for 5km grid cells, using Random Forests regression, which is a state of the art non-linear, non-parametric method in supervised machine learning, and which requires no assumptions with respect to the data distribution [4]. Our choice of 5km was motivated by the underlying 1km granularity of the Geograph data and its associated spatial distribution. We report briefly on sensitivity to resolution in the discussion.

A key task in creating such a model is the choice of appropriate features. Our basic approach was to use training data associated with 5km grid cells, where average scenicness was associated with features based on our terms. Only descriptions consisting of at least five tokens, after filtering as described above, were used in the model. The simplest possible feature set would be one based purely on unigrams, that is to say individual tokens from image descriptions found in grid cells (e.g. ‘hill’, ‘mountain’, ‘shop’, etc.).

However, in natural language processing [21] it is typical to also consider n-grams, and here we also experimented with bigrams (e.g. sequences of two tokens such as ‘steep hill’, ‘rugged mountain’, ‘closed shop’) as features. By reducing the feature space it is often possible to maintain model predictive capacity, while improving performance, and we also experimented by reducing the number of unigrams considered to the n-most frequent. Other features of our data, and previous work on landscape description, suggest additional potential model features which are listed below:

- adjectives alone: since adjectives are assumed to be strong indicators of subjectivity and sentiment; [14], we used unigrams consisting only of frequent adjectives;
- “Landscape Adjective Checklist”: presence of adjectives pertaining specifically to landscape in Craik’s list [24];
- the number of superlative adjectives as identified during part of speech tagging, with the assumption that superlatives are more likely to be used in more scenic areas;
- the number of distinct adjectives found in a description, with the assumption that more adjectives are used in more scenic areas;
- the presence of a verb of perception [39], where we assume that the presence of verbs of perception may indicate descriptions more relevant with respect to scenicness (e.g. by reducing the weight of descriptions focussing on historical events at a location);
- a weight based on spatial tf-idf [29]: here terms which are used frequently in an individual grid cell, but rarely in the collection as a whole are given a higher weight.

2.4 Training and test data

In any supervised model it is necessary to generate both training and test datasets. However, the way in which the data are split can have important implications for not only the quality of the model, but also for any implications which can be drawn from the results. Since an important property of crowdsourced data are user-generated biases in data production [13], we

(a) Scenicness between 1–3,
$n_{images} = 14072$, $n_{users} = 2137$.

(b) Scenicness between 3–5,
$n_{images} = 155822$, $n_{users} = 4170$.

(c) Scenicness between 5–7,
$n_{images} = 79752$, $n_{users} = 3340$.

(d) Scenicness between 7–10,
$n_{images} = 3134$, $n_{users} = 851$.

Figure 3 Average scenicness for 150 most frequent nouns extracted from image descriptions - font size indicates relative frequency within scenicness range.

considered these, as well as the desired spatial contiguity of our model in generating training and test data. Thus, our models were trained (and tested) on the following configurations, with 50% assigned to training and test data respectively in both cases:

- fully random: image descriptions are simply selected at random from the full corpus;
- user dependent random: since we expect individual users to write characteristic descriptions, and since Geograph is subject to participation inequality, meaning that a single user may contribute a large proportion of the descriptions in a single area, we select random images while allowing individual users only to appear in either training or test datasets.

3 Results and interpretation

3.1 Semantics of scenicness

The word clouds in Fig. 3 exemplify our results, illustrating the average scenicness of nouns after part of speech tagging of image descriptions. A number of features are worthy of observation here. At a first glance, the lowest rated scenicness values are related to nouns which are clearly in developed areas (e.g. 'motorway', 'housing', 'shop', 'stadium'). The highest rated scenicness nouns include Gaelic words, terms related to natural processes, wildlife and some esoteric examples (e.g. 'coire', 'avalanche', 'otter', 'backcloth'). However, these classes contain a small proportion of the total set (ca. 6%), with only some 1% of nouns being found in the most scenic class. Thus, many of these nouns belong to the long tail of our data, and although they reflect a clear split between developed areas and more natural landscapes (associated with Gaelic placenames in the Highlands of Scotland) we should be careful not to overinterpret these terms.

Unsurprisingly, since each image was rated at least three times, and many of the nouns are associated with multiple images, the vast majority (94%) of nouns have average scenicness ratings of between 3 and 7. Exploring these classes, it becomes apparent that the clear split so visible in the two extreme classes is much less prominent. Thus, we find that nouns such as 'village', 'lane' and 'wood' are all rated on average 3–5, even though these might be terms typically expected to be associated with more rural, and thus potentially more scenic images. However, exploring the nouns rated 5–7 it again becomes clear that differences exist. Here, many more nouns appear to relate to perceived natural (as opposed to rural) scenes (e.g. 'moorland', 'summit', 'ridge').

3.2 Predicting scenicness

We tested the goodness of fit of our Random Forest regression using the features as described above, and two different configurations of test and training data. Independent of the configuration chosen, we only predicted scenicness values for grid cells where at least two descriptions were present in both training and test data.

Goodness of fit improved as we increased the number of unigrams in the model until we reached the 800 most frequent unigrams. Including presence of adjectives from the "Landscape Adjective Checklist" by Craik and weighting according to spatial tf-idf further increased goodness of fit to a maximum value of around 52% (52.4% in the case of fully random and 52.0% in the case of user dependent random division on training and test data).

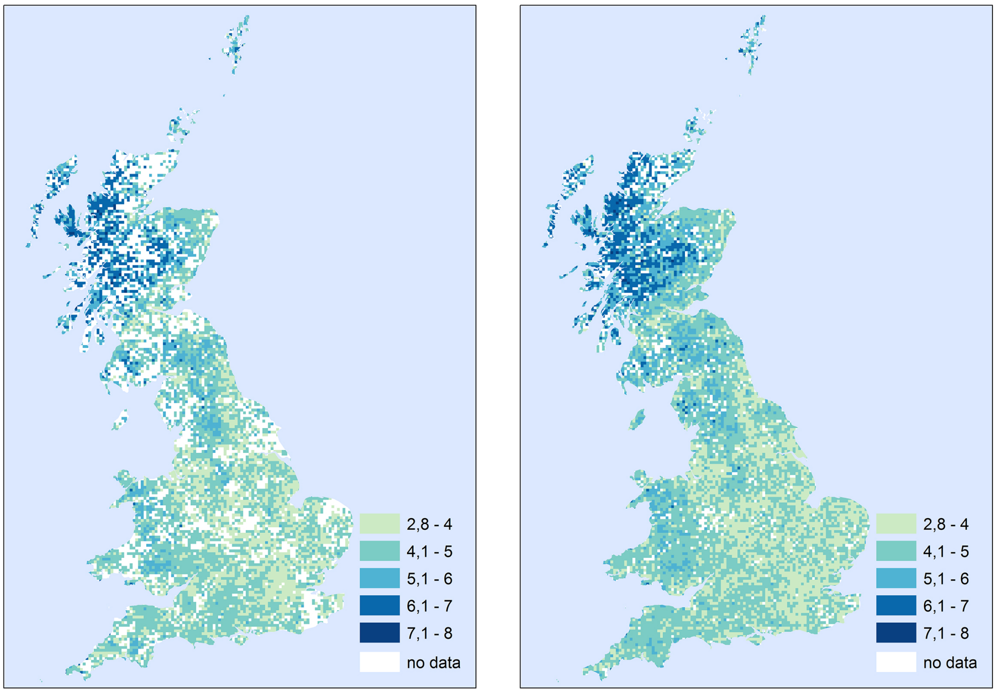
Fig. 4 shows the spatial pattern of predicted scenicness for both configurations. Particularly evident here are the larger number of grid cells for which no value could be predicted where training and test data were randomly selected according to users. Here, the effects of participation inequality result in many grid cells where the majority of images and associated descriptions were taken by a single user, and we thus cannot predict scenicness. However, given the limited variation in model goodness of fit, it appears that this restriction may be unnecessary.

A further important issue in our model is the existence of spatial autocorrelation in model residuals. Testing for Morans-I revealed values of around 0.12 according to model configuration, implying that the chances of random clustering in our model are less than 1%. A typical approach to assessing the influence of spatial autocorrelation in Random Forest regression is therefore to include grid centroids as features in the model [23]. Doing so increased goodness of fit to 56% and reduced spatial autocorrelation in the residuals to 0.05. An alternative model including spatial information by assigning county names (administrative units) to every image, resulted in a decrease of Morans's I to 0.10, with goodness of fit remaining at 52%. This approach includes local neighbourhood relationships and more natural divisions of landscape (since at least in the UK county boundaries typically are a mix of the *fiat* and *bona fide*). Since model results for a model based only on language and containing additional explicit spatial information are similar we thus conclude that our results are not biased by spatial autocorrelation [37].

4 Discussion

In this paper we explored the use of two, related datasets which were both generated by the crowd, though in very different ways, to understand how landscape, and in particular scenicness is captured in language.

Our results were generated after a typical natural language pipeline to tokenise, classify and filter image descriptions. Importantly, we also included a step to remove toponyms from image descriptions, since we were not interested in the names of scenic places, but rather in their properties. Our results demonstrate a clear transition from nouns associated with



(a) User-based division. **(b)** Random division.

Figure 4 Maps of the scenicness prediction results with 'user dependent random division' and 'fully random division'.

urban, developed scenes through more rural landscapes to natural landscapes and a long tail of nouns associated with the Highlands of Scotland. This long tail also reveals one limitation of our approach, since our natural language processing methods cannot deal with Gaelic, and some misclassified words remained in the list of nouns (e.g. *ruadh* refers to the colour red in Gaelic and is commonly used in toponyms i.e. *Sgurr Ruadh* refers to the Red Peak).

Exploration of the word clouds (Fig. 3) reveals that the scenicness of individual terms sometimes contradicts classic ideas in work on landscape preference. For example, water is commonly associated with scenic landscapes [40, 31], yet in our word clouds it has an average scenicness of only 3–5. On closer examination it becomes apparent that water lies in a word cloud containing many rural terms, and the presence of water is common in such scenes. However, at least in our data, rural as opposed to perceived natural scenes are less highly rated. Thus, treating individual nouns (or terms in general) as predictors of scenicness is difficult, and our word clouds reveal more information about the complex interplay between language and landscape. They further indicate the importance of using language, as opposed to purely data-driven approaches to exploring landscape. Approaches extracting landscape properties using intrinsic landscape qualities from standard spatial datasets and associating these with landscape preferences (e.g. [9, 38]) based on ideas of evolutionary-driven landscape perception [18] are unlikely to capture variation of the nature we observe here. Furthermore, our word clouds are potentially powerful tools for generating datasets containing imagery for use in landscape preference experiments and modelling, since they provide an empirical basis for terms used in selecting candidate images, as opposed to approaches based on introspective reasoning or intrinsic, evolutionary determined preferences (cf. [37]) to generate candidate keywords for querying.

Our model of scenicness, irrespective of training data is able to explain some 52% of the variation in scenicness. This is comparable with typical results in more traditional approaches based on interviews or participatory methods [25], approaches using land cover data [35] and work at a continental scale using social media [37]. Although the explained variance is not strongly influenced by our choice of training data, the total number of grid cells for which average scenicness value can be predicted varies by some 20% from around 7000 cells where individual users are only allowed to be present in either test or training data, to 9000 cells where image descriptions are randomly assigned to test or training data. Furthermore, this variation is strongly spatially autocorrelated, with, for example, a single user having taken some 11000 images in the Lake District National Park, of which ca. 850 were rated in the ScenicOrNot project. Such biases are a typical issue in VGI [13], whose handling requires care. Our results were also sensitive to resolution - finer granularities of model reduced model performance (e.g. at 2.5km we could explain 41% of the variation) and coarser granularities increased model performance (e.g. at 10km we could explain 67% of the variation). These results are not unexpected, since firstly the available training data is reduced as resolution becomes finer and, secondly, a coarser model smooths variation and is thus easier to predict, but conveys less fine grained information at the landscape scale.

Our best model used relatively simple features (800 most frequent unigrams, tf-idf and a dictionary of adjectives associated with landscape). Using bigrams, which might be expected to better capture noun phrases associated with scenic locations (e.g. ‘pleasant landscape’) did not in practice improve model performance, an observation which has been made in other contexts [26]. Verbs of perception appear equally likely to be used in scenic or non-scenic contexts, and were also not useful features in our model.

To our knowledge, our approach is the first attempt to use language to spatially model landscape preference, and it has obvious potential to be combined with other approaches to modelling scenicness based either on user frequentation, physical properties of landscape, or combinations thereof [36, 32, 37].

5 Conclusions and outlook

Our work took advantage of two datasets created by volunteers with very different characteristics. Key to their use in our research were firstly the size and spatial extent of both datasets, and secondly the richness of the textual descriptions associated with Geograph images. Our results demonstrate ways in which VGI and crowdsourcing can allow us to explore questions about how space, and in our case scenicness, is captured through use of language, and demonstrate the potential of such approaches. In particular, we observed:

- clear patterns in the nouns associated with scenicness, suggesting a continuum from heavily developed scenes through more rural to perceived natural scenes. Interpreting and using terms to explain scenicness in isolation is challenging, and we suggest that terms should be analysed in isolation with caution;
- a language-based model can predict some 52% of variance in scenicness, comparable with traditional approaches and state of the art statistical models based on parameters known to correlate with scenicness (e.g. terrain roughness or presence of water). Our approach allows us to capture potentially culturally varying landscape preference through the proxy of language; and
- explained variance was not strongly influenced by the way single users describe landscapes. This makes it unnecessary to restrict the appearance of descriptions of single user either in training or in test datasets.

It is important to note that the approaches we take to modelling scenicness, in contrast to our interpretation of word clouds, essentially use a *bag of words* model, where dependencies between terms are not explicitly modelled. In future research we will explore whether, and how, modelling such dependencies might contribute to our understanding of landscape aesthetics. Importantly, we do not claim that our results are universal, but rather reflect the relationship between landscape and language in a particular cultural setting.

We see this work as an example of the use of textual descriptions to explore culturally determined properties of landscape through language. We also intend to explore the transferability of our results to other user generated content (e.g. Flickr or OpenStreetMap), to other spatial regions and languages (e.g. on mainland Europe) and the impact of including additional spatial data on model performance (e.g. terrain models or land cover data). Furthermore, we see great value in attempting to use the literature to build a taxonomy of scene types, and explore their influence on our model. Such an approach could also take advantage of the "unwritten" parts of our descriptions, for example in terms of the arrangements or presence of objects in a particular image or the relationships between colours through content-based analysis of image content associated with descriptions.

Acknowledgements. We would like to thank Michele Volpi for the fruitful discussions and his and the reviewers helpful comments. We would also like to gratefully acknowledge contributors to ScenicOrNot (Open Database Licence) and Geograph British Isles (Creative Commons Attribution-ShareAlike 2.5 Licence).

References

1 J. Appleton. Prospect and Refuges Revisited. In J. L. Nasar, editor, *Environmental Aesthetics: Theory, Research and Applications*, pages 27–44. Cambridge University Press, 1988.

2 J. A. Benfield, P. A. Bell, L. J. Troup, and N. C. Soderstrom. Aesthetic and affective effects of vocal and traffic noise on natural landscape assessment. *Journal of Environmental Psychology*, 30(1):103–111, 2010. doi:10.1016/j.jenvp.2009.10.002.

3 S. Casalegno, R. Inger, C. DeSilvey, and K. J. Gaston. Spatial Covariance between Aesthetic Value & Other Ecosystem Services. *PLOS ONE*, 8(6), 2013. doi:10.1371/journal.pone.0068437.

4 A. Criminisi, J. Shotton, and E. Konukoglu. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends® in Computer Graphics and Vision*, 7(2-3):81–227, 2012.

5 A. Dunkel. Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, 142:173–186, 2015. doi:10.1016/j.landurbplan.2015.02.022.

6 A. J. Edwardes and R. S. Purves. A theoretical grounding for semantic descriptions of place. *Proceedings of the 7th international conference on Web and wireless geographical information systems*, pages 106–120, 2007. doi:10.1007/978-3-540-76925-5_8.

7 S. Elwood, M. F. Goodchild, and D. Z. Sui. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers*, 102(3):571–590, 2012. doi:10.1080/00045608.2011.595657.

8 P. Fisher and D. J. Unwin. Re-presenting Geographical Information Systems. In Fisher, Peter and Unwin, David J., editor, *Re-presenting GIS*, pages 1–14. Wiley Sons London, London, 2005.

9 S. Frank, Ch. Fürst, A. Witt, L. Koschke, and F. Makeschin. Making use of the ecosystem services concept in regional planning-trade-offs from reducing water erosion. *Landscape Ecology*, pages 1–15, 2014. doi:10.1007/s10980-014-9992-3.

10 F. Girardin, J. Blat, F. Calabrese, F. Dal Fiore, and C. Ratti. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4):36–44, 2008. doi:10.1109/MPRV.2008.71.

11 G. Gliozzo, N. Pettorelli, and M. Haklay. Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society*, 21(3), 2016. doi:10.5751/ES-08436-210306.

12 M.F. Goodchild. Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007. doi:10.1007/s10708-007-9111-y.

13 M. Haklay. Why is participation inequality important? In C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors, *European Handbook of Crowdsourced Geographic Information*, pages 35–44. Ubiquity Press, London, 2016.

14 V. Hatzivassiloglou and J.M. Wiebe. Effects of adjective orientation and gradability on sentence subjectivity. *Proceedings of the 18th conference on Computational linguistics-Volume 1*, pages 299–305, 2000. doi:10.3115/990820.990864.

15 M. Hunziker, M. Buchecker, and T. Hartig. Space and Place – Two Aspects of the Human-landscape Relationship. *Challenges for Landscape Research*, pages 47–62, 2007. doi:10.1007/978-1-4020-4436-6_5.

16 A. Jenkins, A. Croitoru, A. T. Crooks, and A. Stefanidis. Crowdsourcing a Collective Sense of Place. *Plos One*, 11(4):e0152932, 2016. doi:10.1371/journal.pone.0152932.

17 X. Junge, B. Schüpbach, T. Walter, B. Schmid, and P. Lindemann-Matthies. Aesthetic quality of agricultural landscape elements in different seasonal stages in Switzerland. *Landscape and Urban Planning*, 133:67–77, 2015. doi:10.1016/j.landurbplan.2014.09.010.

18 R. Kaplan and S. Kaplan. *The experience of nature: a psychological perspective.* Cambridge University Press, 1989.

19 A. Lothian. Landscape and the philosophy of aesthetics: Is landscape quality inherent in the landscape or in the eye of the beholder? *Landscape and Urban Planning*, 44(4):177–198, 1999. doi:10.1016/S0169-2046(99)00019-5.

20 A. Lothian. Scenic perceptions of the visual effects of wind farms on South Australian landscapes. *Geographical Research*, 46(2):196–207, 2008. doi:10.1111/j.1745-5871.2008.00510.x.

21 Ch.D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing.* The MIT Press, Cambridge, Massachusetts, 1999. doi:10.1145/601858.601867.

22 D.M. Mark, A.G. Turk, N. Burenhult, and D. Stea, editors. *Landscape in language. Transdisciplinary perspectives.* John Benjamins, Amsterdam/Philadelphia, 2011.

23 J. Mascaro, G.P. Asner, D.E. Knapp, T. Kennedy-Bowdoin, R.E. Martin, Ch. Anderson, M. Higgins, and K.D. Chadwick. A tale of two "Forests": Random Forest machine learning aids tropical Forest carbon mapping. *PLoS ONE*, 9(1):12–16, 2014. doi:10.1371/journal.pone.0085993.

24 J.L. Nasar. *Environmental Aesthetics: Theory, research and Application.* Cambridge edition, 1992.

25 J.F. Palmer. Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts. *Landscape and Urban Planning*, 69(2-3):201–218, 2004. doi:10.1016/j.landurbplan.2003.08.010.

26 B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. *Empirical Methods in Natural Language Processing (EMNLP)*, 10(July):79–86, 2002. doi:10.3115/1118693.1118704.

27 T. Plieninger, S. Dijks, E. Oteros-Rozas, and C. Bieling. Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, 33:118–129, 2013. doi:10.1016/j.landusepol.2012.12.013.

28 R. S. Purves, A. J. Edwardes, and J. Wood. Describing place through user generated content. *First Monday. Peer-reviewed journal on the internet*, 16(9), 2011. doi:10.5210/fm.v16i9.3710.

29 T. Rattenbury and M. Naaman. Methods for extracting place semantics from Flickr tags. *ACM Transactions on the Web*, 3(1):1–30, 2009. doi:10.1145/1462148.1462149.

30 D. R. Richards and D. A. Friess. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecological Indicators*, 53:187–195, 2015. doi:10.1016/j.ecolind.2015.01.034.

31 U. Schirpke, E. Tasser, and U. Tappeiner. Predicting scenic beauty of mountain regions. *Landscape and Urban Planning*, 111(1):1–12, 2013. doi:10.1016/j.landurbplan.2012.11.010.

32 C. I. Seresinhe, H. S. Moat, and T. Preis. Quantifying scenic areas using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*, page 026581351668730, 2017. doi:10.1177/0265813516687302.

33 C. I. Seresinhe, T. Preis, and H. S. Moat. Quantifying the Impact of Scenic Environments on Health. *Scientific Reports*, 5(Article number 16899):1–9, 2015. doi:10.1038/srep16899.

34 B. Smith and D. M. Mark. Geographical categories: an ontological investigation. *International Journal of Geographical Information Science*, 15(7):591–612, 2001. doi:10.1080/13658810110061199.

35 B. Stadler, R. Purves, and M. Tomko. Exploring the Relationship Between Land Cover and Subjective Evaluation of Scenic Beauty through User Generated Content. In *Proceedings of the 25th International Cartographic Conference*, Paris, 2011. doi:10.5167/uzh-52945.

36 P. Tenerelli, U. Demšar, and S. Luque. Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*, 64:237–248, 2016. doi:10.1016/j.ecolind.2015.12.042.

37 B. T. van Zanten, D. B. van Berkel, R. K. Meetemeyer, J. W. Smith, K. F. Tieskens, and P. H. Verburg. Continental scale quatification of landscape values using social media data. *Proceedings of the National Academy of Sciences*, pages 1–7, 2016. doi:10.1073/pnas.1614158113.

38 B. T. van Zanten, P. H. Verburg, M. J. Koetse, and P. J. H. van Beukering. Preferences for European agrarian landscapes: A meta-analysis of case studies. *Landscape and Urban Planning*, 132:89–101, 2014. doi:10.1016/j.landurbplan.2014.08.012.

39 A. Viberg. The verbs of perception: a typological study. *Linguistics*, 21(1):123–162, 2009. doi:10.1515/ling.1983.21.1.123.

40 B. Yang and T. J. Brown. A Cross-Cultural Comparison of Preferences for Landscape Styles and Landscape Elements. *Environment and Behavior*, 24(4):471–507, 1992. doi:10.1177/0013916592244003.

Juxtaposing Thematic Regions Derived from Spatial and Platial User-Generated Content

Grant McKenzie[1] **and Benjamin Adams**[2]

1 **Department of Geographical Sciences, University of Maryland, College Park, MD, USA**
gmck@umd.edu

2 **Department of Geography, University of Canterbury, Christchurch, New Zealand**
benjamin.adams@canterbury.ac.nz

Abstract

Typical approaches to defining regions, districts or neighborhoods within a city often focus on place instances of a similar type that are grouped together. For example, most cities have at least one bar district defined as such by the clustering of bars within a few city blocks. In reality, it is not the presence of spatial locations labeled as bars that contribute to a bar region, but rather the popularity of the bars themselves. Following the principle that places, and by extension, place-type regions exist via the people that have given space meaning, we explore user-contributed content as a way of extracting this meaning. Kernel density estimation models of place-based social check-ins are compared to spatially tagged social posts with the goal of identifying thematic regions within the city of Los Angeles, CA. Dynamic human activity patterns, represented as temporal signatures, are included in this analysis to demonstrate how regions change over time.

1998 ACM Subject Classification H.1.1 Systems and Information Theory

Keywords and phrases place type, thematic region, temporal signature, topic modeling, user-generated content

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.20

1 Introduction

Colloquially, inhabitants often refer to vernacular places within a city by their *thematic place type*, e.g., *the bar district* or *the shopping area* of the city [18]. Though each of us has a vague understanding of where these regions are (and are not) in the city, and they can be fundamental units of infrastructure for understanding urban dynamics, how we choose to delimit the boundary of these regions remains a topic of discussion for many in the spatial information science community [22, 15, 10]. This research has focused on this task from both topological and cognitive perspectives, identifying the various ways that humans choose to partition their environment. Differentiating commercial centers from residential areas, identifying a city's "downtown," and separating tourist areas from non-tourist areas have all been the subject of empirical and theoretical studies [9].

New types of data have emerged over the past few years that offer the opportunity to re-examine the concept of region delineation through a different lens. User-generated content (UGC) in the form of volunteered geographic information (VGI) and geosocial media have given inhabitants and visitors to a city a range of platforms on which to share their observations [8, 7]. While the majority of previous work in this area has focused on theoretical, simulated or small-sample surveys to gain insight into how individuals and groups

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore;
Article No. 20; pp. 20:1–20:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

think about the urban environment, these new sources of data offer a rich set of *geotagged*, heterogeneous, in situ observations. The form that these *geotags* take and their relationship to the contributed content facilitate a revised debate on the both space and place, and the role they play in identifying thematic regions.

In his early work on the concept of place, Yi-Fu Tuan focused on better understanding and defining the concept and its relation to geographic space [31]. In his writings, he describes how places are spaces instilled with meaning given to them by the people that experience them. This notion is reflected in the observations that are made by people as they move throughout the city. The spatial location of a set of geographic coordinates obtained from the GPS of a mobile device becomes a place when someone tweets about their first kiss at that location. Similarly, the location of a geotagged photograph is a place given meaning by the photographer and subjects of the photograph, and simply preserved in time through a camera. The question is then, how cohesive are these places and is there enough similarity between them to construct regions of common themes? For example, do people tend to talk about activities related to bars (e.g., cocktails, dancing) within distinct spatial areas? What is more, do the locations where people talk about bars align with actual brick-and-mortar bars? In essence, there are two ways of defining thematic regions, one focusing on the linguistic content of spatially tagged observations and another based on the clustering of place instances of a thematic type. With respect to the latter, a current approach might find that a *bar district* of a city is defined as such based on the density of bars in that part of the city. Following the logic that *people define places*, however, means that while a brick-and-mortar venue that sells liquor may be labeled as a *bar*, it is arguably not one until it has a patron. Continuing this thread, an area that has a high density of popular bars contributes more to a consensus of a *bar district* than a cluster of establishments labeled as bars that have no customers.

The times of day that people conduct activities is of particular importance when discussing regions. Kevin Lynch, in his writings on *the image of the city*, describes how one's understanding of the city changes based on time of day, season, etc. [17, p.86] Using our bar place type example, if a bar district were to exist, it would clearly be most "bar-like" at 11pm on a Friday night. Does that bar district still exist at 9am on a Tuesday morning though? We argue here that thematic regions within a city are dynamic and as the activities conducted by people change, so do the regions in which those activities are conducted. The section of the city that facilitated entertainment and alcohol related activities on Friday night ceases to socially afford these activities on Tuesday morning, instead functioning as a space for office or workplace related activities.

The effect of time on thematic regions unveiled through spatially tagged content compared to those built from place-instances remains a point of discussion. So does the influence of environmental characteristics. Some thematic regions are related to physiographic features, e.g., beaches, while other are socially constructed based on a common activity theme, e.g., bars. The regions that emerge from grouping similar spatial or platial[1] thematic instances vary depending on these characteristics. For many people, especially college students, drinking activities tend to dominate the topic of conversation meaning that observations and content related to bars present themselves under spatiotemporal conditions where no bar or drinking activity currently exist, e.g., "Really looking forward to going to the bar with friends tonight." By comparison, observations about physiographic or environmental features tend to be less influenced by time and more restricted to the spatial extent of the physiographic feature.

In this work we examine the role that place, space and time play in defining thematic

[1] We use the neologism *platial* here in reference to place, similar to how spatial refers to space.

regions within a city. Specifically, we investigate how these regions can be identified through the following tasks:

- We use a kernel density estimation model to construct thematic regions from user-generated place instances. Using attribute information associated with these place instances as a proxy for popularity, we demonstrate that regional boundaries change when comparing places that are frequented by visitors with those that are places in name only.
- Using place type-specific temporal signatures generated from millions of human activity patterns, we demonstrate that regions can be represented dynamically. Depending on the time of day, and day of the week, a region may grow or shrink in size.
- We use a topic modeling approach to extract place type linguistic patterns from spatially tagged social media. Through these topics we show how thematic regions can be exposed from natural language content. We compare and contrast thematic regions based on place type and show that there is a substantial difference between content associated with physiographic features and content related to human constructed features.
- Last, we discuss the implication of both the spatial and platial approach to defining thematic urban regions. We show how the constraints and limitations of the various content platforms have an impact on the resulting region definitions.

2 Related Work

In recent years, the explosion in new forms of user generated volunteered geographic data has rekindled a research interest in using quantitative analysis to explore the notion of place [33]. In particular, this new data is seen as an opportunity to tap more closely into the phenomenological idea of place as tied to individual human experiences of their environments [26]. Due to increased "citizen" sensing of the environment via social media and mobile device usage, people are increasingly generating data about their environment through their activities, and human sensors are able to directly collect sensory information about the environment and contextualize and communicate it in language that other humans can understand [11]. Thus, this data gives us unique insight to place identity and sentiment as well as relationships between individuals, groups, and the physical environment [27]. By and large this research has explored the use of a variety of spatial analysis techniques as well as other data science methods, such as text mining, to operationalize place in geographic information systems.

A sub-area of this work revolves around creating spatial representations for regions, which are vaguely or non-canonically defined. Examples of this work include generating spatial footprints of vaguely defined *platial* regions such as tourist areas and city centers [14, 6], and for constructing spatial footprints for digital gazetteers [16]. Much of this work is not specifically at the *place-type* level nor has it compared place-based regions with regions generated from spatial footprints.

Because many human sensor observations are stored as natural language text, a variety of research projects have used text mining and natural language processing (NLP) tools and methods to generate structured place knowledge. Textual and narrative descriptions arguably provide a unique perspective on human interpretations and conceptualizations of place, because of the opportunity to infer information about what people "think" about places. In practice, different NLP methods provide this to different degrees. Existing work has focused on extracting place semantics from user-generated spatial content and narratives [13, 28]. In addition, computational sentiment and emotion analysis has been used to infer regularities in the emotional content of place descriptions [4].

One approach to better understand the thematic contents of place narratives is topic modeling, a family of probabilistic machine learning methods commonly used to infer thematic structure in a corpus of text documents. The simplest topic model is Latent Dirichlet Allocation (LDA), which models the generation of a document set as the result of a random process [5]. First, a document is assigned a mixture of topics, then each word is randomly picked from those topics proportionately based on the importance of that word for the topic. Given this model as a starting point, the inferencing algorithm identifies the topics that would have most likely generated the existing corpus. Thus, it is an unsupervised method in that it derives the topics from the document set without any additional information or pre-defined structure. LDA is a bag-of-words model where word-order, parts of speech, and other grammatical structures are ignored. Despite this simplification, LDA is widely used as a way to quantitatively characterize the topics that are in individual documents and across an entire corpus. LDA has been used to map regions and times that are described using those themes [2]. It has also been used to discover thematic signatures of place types from text for the purpose of enriching place-based linked data [1]. In our previous work we developed a thematic search engine that uses geotagged natural language content from Wikipedia and travel blogs to cartographically present the results of topic-specific searches [3]. Additionally, topic modeling was used to extract vague cognitive regions such as *Southern California* from user-generated spatial content [10].

3 Data

A sample of 213,279 user-generated place instances were accessed via the Foursquare application programming interface[2] for the greater Los Angeles area. These place instances are categorized into one of 421 user-contributed place types (e.g., Bar, Park, Police Station) curated in a hierarchical vocabulary.[2] Of these, we selected 20 of the most dominant and unique place types restricting our place instance set to 37,302. Similarly, a random sample of 684,776, 699,113, and 642,059 geo-tagged (geographic coordinates) social media posts were collected from *Twitter*, *Instagram* and *Yik Yak* respectively over a three month time span starting January 2015. Twitter is a microposting service which restricts posts to 140 characters.[3] Yik Yak is a mobile application allowing users to post anonymous content to other users within a 5 mile radius of their location. The photo sharing platform, Instagram, geotags photographs and captions by default.[4] Note that only the text-based captions, not the photographs, were used in this analysis. These platforms range in the demographics of their user-base but most users are between the ages of 19 and 29 [23].

3.1 Temporal Signatures

Hourly temporal signatures for each of the 421 place types were constructed from check-in data collected in Los Angeles over a three month time period and aggregated to hours in a single week (see [20] for details). These normalized temporal signatures represent the *default* activity behavior at a given place type in Los Angeles. Figure 1 shows a sample of 3 place type temporal signatures. A higher value represents an increase in likelihood of finding someone at a place of that type at that time. *Mexican Restaurant* displays peaks at lunch and dinner time throughout the week while *Bar* activity is shown increasing late at

[2] `https://developer.foursquare.com/`
[3] The removal of this character limit occurred in 2016.
[4] Data was collected for this project prior to Instagram changing their location-tag settings.

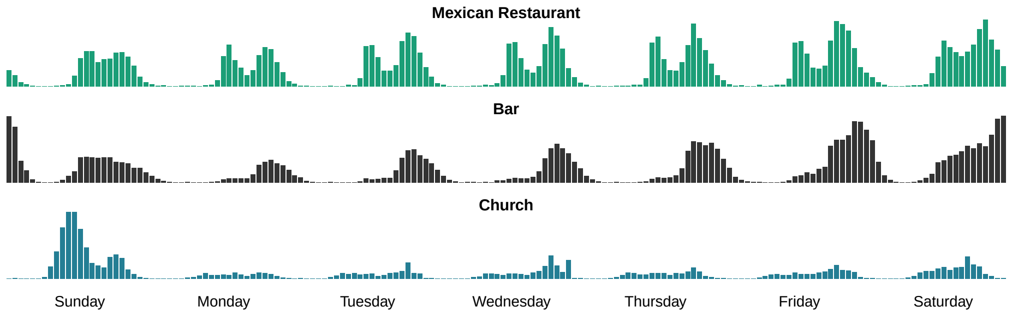


Figure 1 Hourly temporal signatures constructed from geosocial media check-ins for three place types in Los Angeles, CA.

(a) Mexican Restaurant **(b)** Bar **(c)** Park

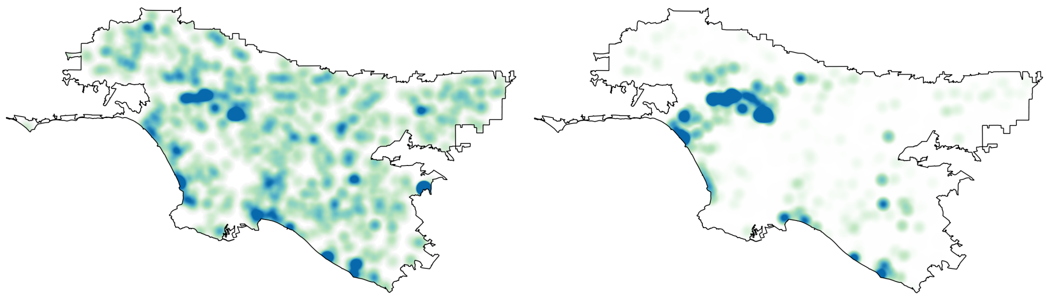
Figure 2 Three place type specific topic word clouds extracted from Foursquare tips.

night throughout the week. *Church* presents an expected peak on Sunday morning and a smaller one on Sunday afternoon with negligible activity the remainder of the week. The purpose of visually depicting these temporal signatures is to show that default activity behavior towards places does vary significantly between place types. The involvement of these temporal signatures in identifying regions will become apparent in Section 4.

3.2 Linguistic Signatures

A hexagonal grid was generated over the greater Los Angeles area with grid cells at 0.01 degrees wide in latitude (roughly 1.1 km). All geotagged tweets, Instagram captions and Yik Yak posts were intersected with the hexagonal grid and each post was assigned to a grid cell. The textual content of these posts were cleaned by removing all non-alphanumeric characters as well as removing all stop words and words less than three characters.

Topic modeling was used to extract common themes across the spatial data sources. Previous work has used topic modeling to derive thematic signatures for places from unlabeled text [2, 1]. However, in this case we had a dataset of labeled tips from Foursquare, so we could use a form of supervised topic modeling called Labeled LDA (L-LDA) to train for topics that match the 20 place types selected [25]. Similar to LDA, L-LDA models a topic as a probability distribution over words, indicating the likelihood that someone writing on that topic will use particular words. It differs from LDA in that a one-to-one relationship is maintained between the user-supplied place type labels and the topics that are learned. Figure 2 represents the topics learned from the Foursquare tips for three place types.



(a) Non-weighted KDE **(b)** KDE weighted by unique visitor check-ins

Figure 3 Two kernel density estimate representations of the *Bar* place type in Los Angeles, CA.

4 Thematic regions identified through place instances

In many gazetteers, place instances are stored as point representations. This also holds true for many geosocial media place dictionaries such as Facebook and Foursquare. One approach to constructing regions from point data is to use kernel density estimation (KDE). This has been successful for constructing spatiotemporal regions from spatial locations of photographs [12, 30] and georeferenced text [3, 24]. Here, we split the place instances in the greater Los Angeles area by their place type and construct kernel density estimations based on these points. The KDE bandwidth used in each of these was calculated using the method proposed by Sheather and Jones [29].

One approach to identifying regions for a specific place type, e.g., *Bar*, using kernel density estimation is to weight each instance of a bar equally. The assumption here being that all bars are equal in their *bar-ness* and that the presence of a bar in the city, regardless of location, size or popularity should contribute equally to the identification of one or multiple regions. While this is a reasonable approach, it does make the arguably erroneous assumption that all bars contribute equally to what one might consider a *bar region*. We argue here, that a popular bar, for example, should contribute more to defining a *bar region* than a venue that has had little to no visitors in the past year. Ascertaining the actual popularity of a venue is a monumental task however. Fortunately, new sources of user-contributed place-based data now exist that work as proxies to actual venue popularity measures. Geosocial media content such as *check-ins* offers additional information concerning both place types and place instances that were previously only accessible through cost and time-prohibitive surveys or simulated data. Interaction behavior with the Foursquare representation of a place instance is accessible via a number of attributes including *unique visitor check-ins*, *total number of check-ins* and *total number of likes*. A check-in in this case refers to the act of an individual using the Foursquare application on their mobile device to indicate that they are at the physical place represented in their application as a Foursquare *venue*. A *Like*, on the other hand, does not imply that the user is or was at the actual physical place. To account for place instances being added to the Foursquare gazetteer at different times, attribute values used in this work were restricted to the last three years.

Not surprisingly there is a high correlation ($Pearson > 0.83$, $p < 0.01$) between the three Foursquare attributes which is reflected in the regions exposed by attribute weighted kernel density estimation models. Given these high correlation values, we chose to focus on *unique visitor check-ins* in a KDE weighted model. Applying this attribute as a weight in the KDE ensures that place instances that have had a higher number of unique visitors,

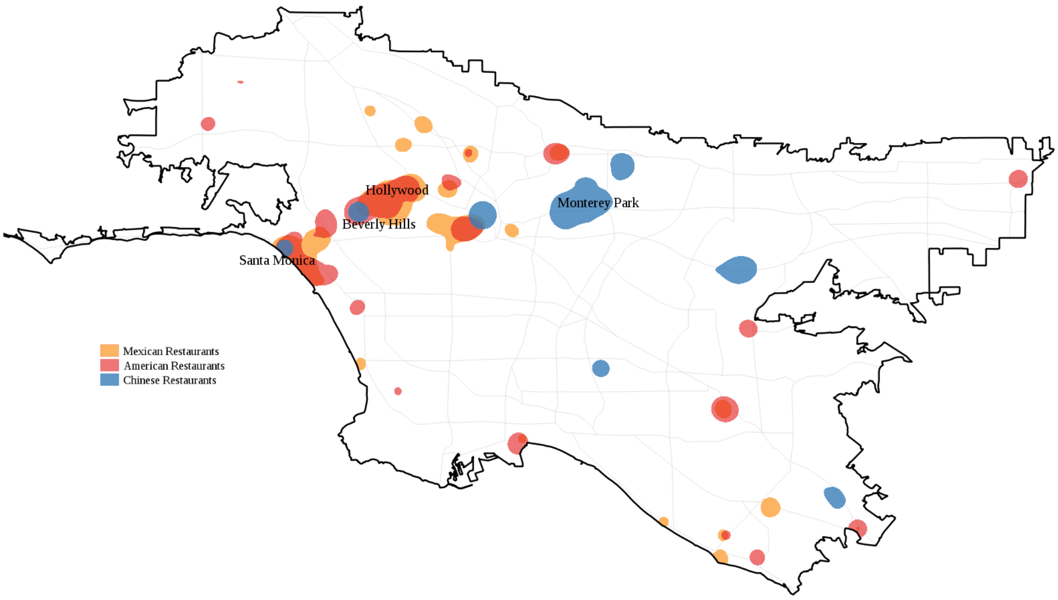


Figure 4 Three restaurant types in Los Angeles, CA.

also have a larger influence on the KDE and the discovery of thematic regions. Figure 3 shows a cartographic representation of two KDE models for *bars* using no weight (Figure 3a) and *unique visitors check-ins* (Figure 3b). The difference in these two maps highlights the influence that the popularity of a set of bars has on defining bar-type-regions. Note that over our sample set of 37302 place instances, 18% (6717) had no visitor check-ins. In our subset of *bars*, 45% had less than 10 unique visitors while 14% listed more than 3000.

4.1 Specifying region boundaries

Generating and mapping a kernel density estimation model for a place type produces a cartographic representation of a region with vague or fuzzy boundaries. As shown in Figure 3, opaque blue highlights the most *bar*-like areas while semi-transparent green indicates areas that are less *bar*-like. While this representation of regions via fuzzy boundaries is often appropriate for discussion purposes, as it reflects our cognitive perception of thematic regions, specifying a threshold on which to state that a region is either a *bar region* or not is of value in some cases [15]. For example, certain urban planning laws in the United States require that commercial land-use be specified by a hard boundary (typically streets) and restricting these boundaries or limiting place-types to a certain neighborhood or set of city blocks is often necessary for zoning purposes.

To construct these hard boundaries, we removed all raster pixel values below two standard deviations above the mean (for the given KDE raster) and assigned all other pixels a value of 1. Three types of restaurants in Los Angeles were analyzed in this way and are presented in Figure 4. There are clear similarities and differences in the regions produced through this analysis. From a qualitative perspective, the most notable similarity is that all three types of restaurants have a regional presence in the city of Santa Monica and surrounding area. Both Mexican and American restaurants are popular along the coast to Venice Beach and Marian Del Ray neighborhoods as well as inland around West Los Angeles. Similarly, Hollywood is a hot spot for both American and Mexican cuisine while Chinese Restaurants

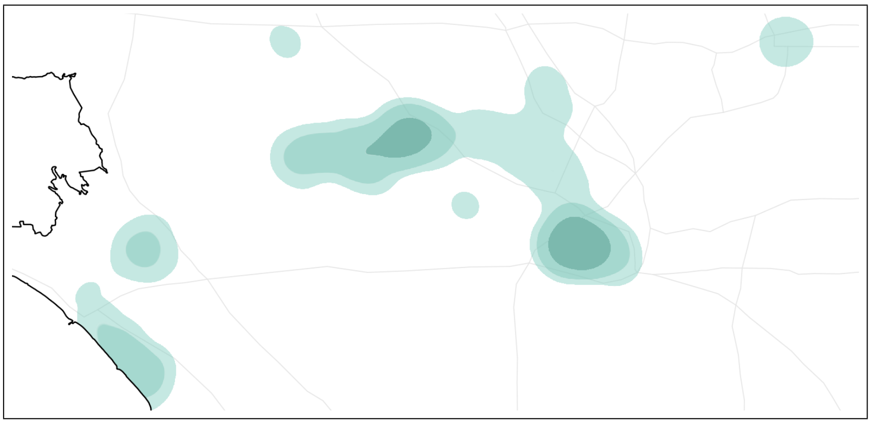


Figure 5 The *Bar* thematic place type region changing by time of day. Darkest to lightest: Friday 3pm, 7pm, 11pm.

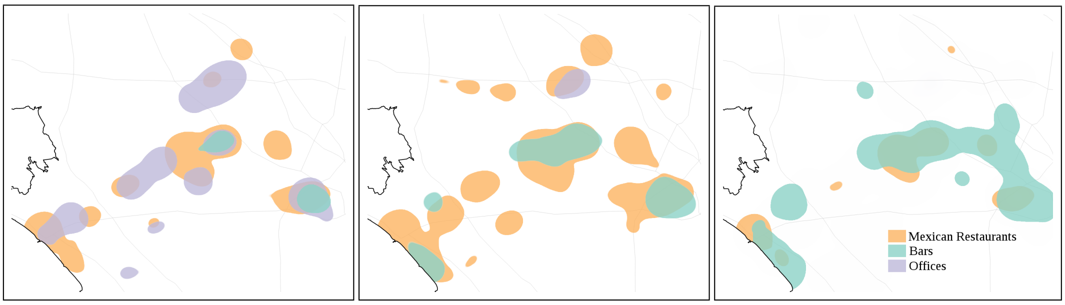
are also popular in the city of Beverly Hills. There are notable differences between these restaurant regions as well. The city of Monterey Park overlaps with the largest thematic region of Chinese Restaurants in this figure. According to the 2010 U.S. Census, Monterey Park contains a population that identifies as 66.9% Asian decent with a large concentrations of Chinese Americans [32].

4.2 Temporal dynamics of thematic regions

In our previous work on place types, we used social check-ins to generate temporal signatures of human activity behavior based on time of day [19]. This work confirms the notion that certain place types are more popular at certain times of day and days of the week. For example, people are more likely to patron restaurants during midday and evenings and employees have a high probability of being in office buildings between 9am and 5pm on weekdays. These temporal signatures offer unique insight into what place type activities happen when. Combining these temporal signatures with our thematic place type regions allows us to model the temporal dynamics of a city like Los Angeles. This work continues existing efforts in examining the *pulse* of a city, focusing on thematic regions instead of point representations [21].

Regions for three place type, namely *Mexican restaurants*, *bars* and *offices* were constructed using the unique visitor check-in weighted KDE method described in Section 4. As previously mentioned, the threshold value two standard deviations above the mean was recorded for each of these place types. The original place location data containing normalized values for unique visitor check-ins was then multiplied by the normalized temporal signature for the respective place type at three different times. This produced three new values on which to weight three new kernel density estimates (three for each place type). In this example, the times were Friday at 3pm, 7pm and 11pm. The threshold value from the original non-temporally weighted KDE was applied to each of the three new KDE maps which resulted in larger or smaller regions depending on the temporal probability value. Figure 5 shows the bar region for one section of Los Angeles as three temporal snapshots overlain on top of each other. The regions are represented temporally from darkest to lightest in this example with the smallest *bar region* around 3pm and the largest *bar region* (highest temporal probability) at 11pm.

The effect of time is different on each thematic region as shown by three examples in Figure 6. The *Office* region, shown in purple, decreases in area from 3pm to 11pm on Friday



(a) 3pm Friday **(b)** 7pm Friday **(c)** 11pm Friday

Figure 6 Merging three place-type regions with their default temporal signatures at three times on a typical Friday.

(in fact it is non existent at 11pm) while the regions representing *Mexican restaurants*, in orange, peak in size at 7pm. *Bars*, as shown in Figure 5, grow significantly from 3pm to 11pm. These visual representations reflect the idea that regions of a city are transitional [17]. While the buildings and spaces that contain the place instances exist atemporally, the places themselves and the regions that they contribute to are temporally dynamic.

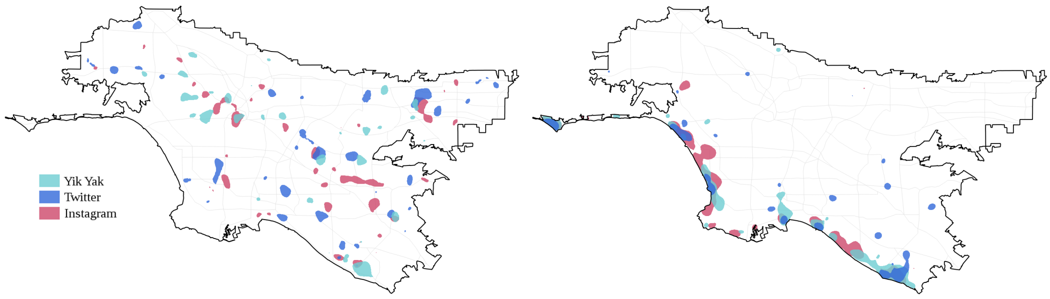
5 Thematic regions identified through spatially tagged content

The L-LDA topics extracted from geosocial spatial data were used to identify spatial regions in Los Angeles. Following a similar approach to the regions built from place-based data, kernel density estimate models were plotted from the 0.01 degree hexagonal grid using the topic values as the weight. As mentioned in Section 3.2, these topics were extracted from Foursquare tips, trained by the appropriate place type, and used to label our *spatial* data: *Twitter* tweets, *Yik Yak* posts and *Instagram* photo captions. For example, text related to *bars* in Foursquare tips were used to identify and label tweets with similar textual content. However, we found that in most cases, the thematic regions defined by the spatial datasets did not align with the place instance-identified regions. Moreover, in many cases, there was no common agreement between the different social media platforms themselves. Further investigation found that the main difference impacting agreement between datasets was the broader category of place type. More specifically, whether the thematic place type was tied to a feature in the natural or human-built environment.

5.1 Physiographic vs. human defined place types

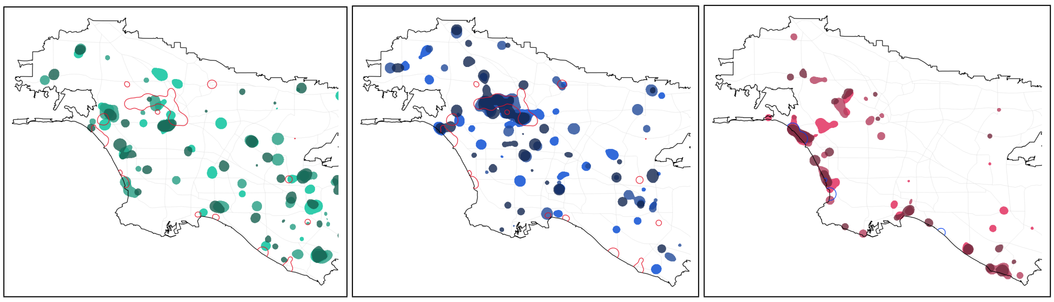
There is very little agreement between the three spatial datasets for human-built regions such as *bars* (Figure 7a), *Mexican restaurants* or *office buildings* discussed in the previous section. In fact a pixel-based Jaccard similarity coefficient (threshold at $mean + 2SD$) for these three datasets was consistently below 0.05 for all pairs of datasets.[5] By comparison, place types that mapped to physiographic features in the natural environment such as *Beach* (Figure 7b) showed much higher agreement between the three datasets with an average Jaccard coefficient of 0.28. These differences likely reflect how people refer to these different

[5] Jaccard is bounded between 0 and 1, with the latter indicating the datasets are identical.



(a) Bar **(b)** Beach

Figure 7 Regions for *Bar* and *Beach* identified from three different spatially tagged social media platforms.



(a) Bar regions via Yik Yak **(b)** Bar regions via Twitter **(c)** Beach regions via Instagram

Figure 8 Bar topics for two datasets at 3pm, 7pm, 11pm on Friday.

groups of place types as well as the demographics of the application users. This will be discussed further in Section 6.

5.2 Temporal dynamics of spatial content

Following the temporal example described in Section 4.2 of Friday at 3pm, 7pm and 11pm, we extracted spatially tagged content from each of the three social media platforms for those times periods. Though this reduced the sample size from which to generate topic signatures, Friday evening is a popular time for all social media applications meaning there was adequate data on which to run the analysis. Using the *bar* example again, we compared the regions identified for the three times of day across the three data sources. The results were inconsistent. In some cases, e.g., Yik Yak (Figure 8a), small regions were identified in different parts of Los Angeles, at different times of the day. There is little overlap between time periods and in general and the overlap that does exist does not align with the place-instance based thematic regions, outlined in red. On the other hand, Twitter (Figure 8b), again shows small, inconsistent regions outside of the city center, but there is significant overlap with the place instance-based regions over all time periods on a Friday. Notably, the number of regions do not increase nor is there a change in the size of the regions as the evening progresses.

By comparison, the *beach* regions identified via Instagram photograph captions are primarily clustered around the Los Angeles coastline and reflect a similar pattern to the one shown in Figure 7b. The overlap across hour layers is high and though there are regions identified inland, they are primarily clustered around the downtown city core where there is significant social media activity. Potential explanations for this are discussed in the next section.

6 Discussion & Conclusions

The results presented in the previous two sections deserve further discussion, specifically on the difference between regions identified through spatial data and those identified through place instances. The biases and limitations associated with these data are also discussed.

6.1 Spatial tags vs. Place instances

There are important differences between regions identified through spatial and platial sources. Spatially tagged social content reflect observations of individuals at certain locations and times. The content of an observation, however, need not reflect the affordances or activities associated with the space from which the observation was made. The bar regions we identified, for example, tend to be dispersed across all of Los Angeles at many times of day. The reason for this can be understood from an example Yik Yak post at 3pm on Thursday in North-East Los Angeles: *"Can't wait to hit up the bars tonight, it has been a long week."* What we find is that this disconnect between what place instances identify as regions and what spatially tagged posts identify as regions, varies by place type. Human constructed places tend to have a larger disconnect between the spatially and platially identified regions while physiographic regions show greater alignment.

Also of importance is the difference in the intention of the data source. For example *beaches* contributed to Foursquare tend to be officially designated public beaches. Spatially-tagged social content, however, rarely explicitly identifies a beach. The content reflects observations and language related to beaches in general. The contributor of the latter data may not actually care that she is standing on an officially designated public beach. In all likelihood her definition of a beach simply consists of a sandy area adjacent to a body of water. In this case it is not unexpected that the regions identified from the place-based data may differ from those identified from spatially tagged observations, even for physiographic features.

6.2 User-generated Content

User-generated content, of which social media is one type comes with its own set of biases. Like all data, it is influenced by the views of its creators. In the case of these geotagged data, the contributors are predominantly young adults. This demographic has biases towards certain place types and the amount and nature of the content reflects these biases. For example, young adults arguably have a much more complex relationship with *bars* than they do *beaches*. While the social capital involved with posting about beach activities is high, it pales in comparison to activities related to drinking alcohol.

The structure of the Foursquare place type vocabulary also impacts how this work identifies regions. The place type *Bar* shares a number of similarities related to entertainment and alcohol with other place types such as *nightclubs, lounges, Karaoke venue*, etc. It is likely that the words and topics extracted for bars via L-LDA are quite similar to those of these other place categories yet the regions identified may be slightly different. The language used in social content may align with these similar place types as well. Similarly, the terms identified as being most *bar*-like may potentially be used to describe social interactions with friends in a dorm room or *tailgating* outside of a stadium. Though our topic modeling approach assigned many alcohol and entertainment terms to the *bar* place type, the noise and ambiguousness common to social media posts could have lead to some mis-labeling. These are some of the known issues of working with natural language classification.

6.3 Conclusions

Understanding how thematic regions are identified within a city has been a topic of discussion in the spatial science community for many years. This work makes use of two unique types of geographic content, namely spatially tagged social media posts and thematically labeled place instances. Novel aspects of these data offer insight into how people interact with a city, allowing us to identify thematic regions through the use of weighted analysis models. The heart of this research, however, lies in a discussion of space and place. Does having access to user-contributed geographic content enhance our understanding of the relationship between space and place? Does the inclusion of new and alternative datasets change our existing cognitive and theoretical approaches to how regions are defined? These are questions that we have just scratches the surface of in this work and will continue to examine in future research.

References

1 Benjamin Adams and Krzysztof Janowicz. Thematic signatures for cleansing and enriching place-related linked data. *International Journal of Geographical Information Science*, 29(4):556–579, 2015.

2 Benjamin Adams and Grant McKenzie. Inferring thematic places from spatially referenced natural language descriptions. In Daniel Sui, Sarah Elwood, and Michael Goodchild, editors, *Crowdsourcing Geographic Knowledge*, pages 201–221. Springer, 2013.

3 Benjamin Adams, Grant McKenzie, and Mark Gahegan. Frankenplace: interactive thematic mapping for ad hoc exploratory search. In *Proceedings of the 24th International Conference on World Wide Web*, pages 12–22. ACM, 2015.

4 Andrea Ballatore and Benjamin Adams. Extracting place emotions from travel blogs. In *Proceedings of AGILE*, volume 2015, pages 1–5, 2015.

5 David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

6 Jiaoli Chen and Shih-Lung Shaw. Representing the spatial extent of places based on flickr photos with a representativeness-weighted kernel density estimation. In *International Conference on Geographic Information Science*, pages 130–144. Springer, 2016.

7 Jeremy W. Crampton, Mark Graham, Ate Poorthuis, Taylor Shelton, Monica Stephens, Matthew W. Wilson, and Matthew Zook. Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb. *Cartography and geographic information science*, 40(2):130–139, 2013.

8 Andrew Crooks, Dieter Pfoser, Andrew Jenkins, Arie Croitoru, Anthony Stefanidis, Duncan Smith, Sophia Karagiorgou, Alexandros Efentakis, and George Lamprianidis. Crowdsourcing urban form and function. *International Journal of Geographical Information Science*, 29(5):720–741, 2015.

9 Clare Davies, Ian Holt, Jenny Green, Jenny Harding, and Lucy Diamond. User needs and implications for modelling vague named places. *Spatial Cognition & Computation*, 9(3):174–194, 2009.

10 Song Gao, Krzysztof Janowicz, Daniel R. Montello, Yingjie Hu, Jiue-An Yang, Grant McKenzie, Yiting Ju, Li Gong, Benjamin Adams, and Bo Yan. A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science*, 2017, Ahead of Print.

11 Michael F. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

12 Christian Grothe and Jochen Schaab. Automated footprint generation from geotags with kernel density estimation and support vector machines. *Spatial Cognition & Computation*, 9(3):195–211, 2009.

13 Heidelinde Hobel, Paolo Fogliaroni, and Andrew U Frank. Deriving the geographic footprint of cognitive regions. In *Geospatial Data in a Changing World*, pages 67–84. Springer, 2016.

14 Livia Hollenstein and Ross Purves. Exploring place through user-generated content: Using flickr tags to describe city cores. *Journal of Spatial Information Science*, (1):21–48, 2010.

15 Christopher B. Jones, Ross S. Purves, Paul D. Clough, and Hideo Joho. Modelling vague places with knowledge from the web. *International Journal of Geographical Information Science*, 22(10):1045–1065, 2008.

16 Carsten Keßler, Patrick Maué, Jan Torben Heuer, and Thomas Bartoschek. Bottom-up gazetteers: Learning from the implicit semantics of geotags. In *International Conference on GeoSpatial Sematics*, pages 83–102. Springer, 2009.

17 Kevin Lynch. *The image of the city*, volume 11. MIT press, 1960.

18 G. Mattson. Bar districts as subcultural amenities. *City, Culture, Society*, 6(1):1–8, 2015.

19 Grant McKenzie and Krzysztof Janowicz. Where is also about time: A location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures. *Computers, Environment and Urban Systems*, 54:1–13, 2015.

20 Grant McKenzie, Krzysztof Janowicz, Song Gao, and Li Gong. How where is when? on the regional variability and resolution of geosocial temporal signatures for points of interest. *Computers, Environment and Urban Systems*, 54:336–346, 2015.

21 Grant McKenzie, Krzysztof Janowicz, Song Gao, Jiue-An Yang, and Yingjie Hu. POI pulse: A multi-granular, semantic signature–based information observatory for the interactive visualization of big geosocial data. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 50(2):71–85, 2015.

22 Daniel R. Montello, Michael F. Goodchild, Jonathon Gottsegen, and Peter Fohl. Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition & Computation*, 3(2-3):185–204, 2003.

23 Pew Research Center. Demographics of key social networking platforms, 2015. Accessed: 2017-02-27. URL: `http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/`.

24 Ross S. Purves and Curdin Derungs. From space to place: Place-based explorations of text. *International Journal of Humanities and Arts Computing*, 9(1):74–94, 2015.

25 Daniel Ramage, David Hall, Ramesh Nallapati, and Christopher D. Manning. Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, EMNLP'09, pages 248–256, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

26 Edward Relph. *Place and placelessness*, volume 67. Pion London, 1976.

27 Stéphane Roche. Geographic information science II Less space, more places in smart cities. *Progress in Human Geography*, 40(4):565–573, 2016.

28 Simon Scheider and Ross Purves. Semantic place localization from narratives. In *COMP@ SIGSPATIAL*, pages 16–19, 2013.

29 Simon J. Sheather and Michael C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 683–690, 1991.

30 Yeran Sun, Hongchao Fan, Marco Helbich, and Alexander Zipf. Analyzing human activities through volunteered geographic information:Using Flickr to analyze spatial and temporal pattern of tourist accommodation. In *Progress in location-based services*, pages 57–69. 2013.

31 Yi-Fu Tuan. *Space and place: The perspective of experience.* Univ. of Minnesota, 1977.

32 United States Census. United States Census Quick Facts: Monterey Park, CA, 2010. Accessed: 2017-02-27. URL: `https://www.census.gov/quickfacts/`.

33 Stephan Winter and Christian Freksa. Approaching the notion of place by contrast. *Journal of Spatial Information Science*, 2012(5):31–50, 2012.

Using Flickr for Characterizing the Environment: An Exploratory Analysis*

Shelan S. Jeawak[1], Christopher B. Jones[2], and Steven Schockaert[3]

1 **Cardiff University, School of Computer Science and Informatics, Cardiff, UK**
JeawakSS@cardiff.ac.uk
2 **Cardiff University, School of Computer Science and Informatics, Cardiff, UK**
JonesCB2@cardiff.ac.uk
3 **Cardiff University, School of Computer Science and Informatics, Cardiff, UK**
SchockaertS1@cardiff.ac.uk

Abstract

The photo-sharing website Flickr has become a valuable informal information source in disciplines such as geography and ecology. Some ecologists, for instance, have been manually analysing Flickr to obtain information that is more up-to-date than what is found in traditional sources. While several previous works have shown the potential of Flickr tags for characterizing places, it remains unclear to what extent such tags can be used to derive scientifically useful information for ecologists in an automated way. To obtain a clearer picture about the kinds of environmental features that can be modelled using Flickr tags, we consider the problem of predicting scenicness, species distribution, land cover, and several climate related features. Our focus is on comparing the predictive power of Flickr tags with that of structured data from more traditional sources. We find that, broadly speaking, Flickr tags perform comparably to the considered structured data sources, being sometimes better and sometimes worse. Most importantly, we find that combining Flickr tags with structured data sources consistently, and sometimes substantially, improves the results. This suggests that Flickr indeed provides information that is complementary to traditional sources.

1998 ACM Subject Classification I.2.6 Learning

Keywords and phrases Social media, Volunteered Geographic Information, Ecology

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.21

1 Introduction

Social media websites such as Flickr[1], Twitter[2] and Facebook[3] have become a popular platform to share and find information. Flickr, for instance, hosts more than 10 billion photographs[4], many of which are associated with textual data in the form of a title, description and a set of tags that express what is in the photo. The availability of GPS systems in current electronic devices such as smartphones enables latitude and longitude coordinates to be recorded as meta-data. For a large number of photos on Flickr, these coordinates have been made publicly available[5]. Together with their textual descriptions, such photos can be

* Shelan Jeawak has been sponsored by HCED Iraq. Steven Schockaert has been supported by ERC Starting Grant 637277.
[1] http://www.flickr.com
[2] http://www.twitter.com
[3] http://www.facebook.com
[4] http://expandedramblings.com/index.php/flickr-stats
[5] We were able to crawl around 350M georeferenced Flickr photos in September 2015.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore;
Article No. 21; pp. 21:1–21:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

regarded as Volunteered Geographic Information (VGI [11]). The coordinates and textual meta-data associated with Flickr photos have already proven valuable in many disciplines such as health [18], geography [4, 12], and ecology [1, 5].

Although there are many organizations that serve environmental data, the information they provide is far from complete [1]. The idea of using Flickr as a supplementary source of environmental data is appealing for several reasons. For example, due to the fact that photos are often uploaded directly after they have been taken, Flickr can provide us with more up-to-date information than traditional citizen science datasets, which is important e.g. for monitoring the spread of invasive species and migration patterns of pollinators. Moreover, the information that is captured by Flickr tags is broader than what is normally recorded, and includes e.g. subjective assessments about the scenicness of a landscape. In fact, Flickr has already proven valuable as a resource for ecological analysis. However, ecologists are currently mostly analysing Flickr data manually. For example [5, 20] manually analysed the content of social media websites to assess its usefulness for ecology. Manually analysing Flickr is clearly limited and time-consuming. Moreover, both the structure and the volume of the data present practical challenges [5], compared to formal or semi-formal citizen science monitoring data [23]. Nonetheless, these studies prove that Flickr contains valuable information which could be used to support the available sources [1, 5]. All this highlights the need for automated methods for extracting environmental information from Flickr.

The main research questions we consider in this paper are whether it is possible to extract large amounts of high-quality environmental information from Flickr, and if so, how complementary this information is to publicly available environmental data sets. In particular, we test the usefulness of Flickr for predicting a broad set of environmental features: scenicness, species distribution, land cover, and climate data. Our main aim is to provide a clearer picture of the kinds of environmental characteristics for which scientifically useful information can be derived from Flickr. Our analysis in this paper will focus on features that we can ascribe to locations (e.g. there is a coniferous forest at this location) rather than to individual photos (e.g. this is a photo of a 7-spot ladybird). While the latter is also important, it requires solving a particular set of challenges beyond the scope of this paper (e.g. distinguishing photos of sightings from photos that are tagged with the name of a species for other reasons) and is difficult to evaluate given the lack of ground truth data.

The remainder of this paper is organised as follows. Section 2 gives an overview of related work. Section 3 presents our methodology for making predictions from Flickr tags and from traditional data. In Section 4 we then provide a detailed discussion about our experimental results. Finally, Section 5 summarises our conclusions and plans for future work.

2 Related work

2.1 Citizen science related research

Considerable progress has been made in recent years in citizen science projects in the environmental sciences that recruit participants to contribute actively to particular campaigns such as in land cover mapping [9], hydrological surveys [17], ornithology and many forms of ecological study [6]. In parallel with these initiatives there is growing interest in the potential of "passive" survey methods that exploit social media to provide additional useful data. For instance, [27] analysed the visual features of the photographs on Flickr automatically to observe natural world features such as snow cover and a particular species of flower. In [28] photos from Flickr were used to estimate snow cover and vegetation cover, and compare these estimations with fine-grained ground truth collected by earth-observing satellites and ground stations. Both the text associated with Flickr photographs and their visual features were used in [16] to perform land-use classification. The approach was evaluated on two university campuses and three land-use classes were considered: Academic, Residential, and Sports. In [7] and [8], they classified a sample of georeferenced Flickr photos according to CORINE land cover classes. They also evaluated the use of Flickr photos in supporting Land Use/Land Cover (LULC) classification for the city of Coimbra in Portugal and for comparison with Corine Land Cover (CLC) level 1 and level 2 classes. The results of this approach were evaluated manually by experts. Their results suggest that Flickr photos cannot be used as a single source to achieve this purpose but they could be helpful if combined with other sources of data.

The authors of [24] explored the relationship between CORINE land cover classes and the valuation of natural scenery, namely scenicness, scenic beauty, landscape beauty, aesthetics, or cultural ecosystem services (CES), through user evaluated georeferenced photos from the ScenicOrNot[6] website. They employed the user's rating of a photo in a specific area as an evaluation of the land cover of that area. The results of this study showed that the highest rated areas belong to the forest and semi natural areas, and water bodies classes. Measures of scenicness are important since they reflect human well-being and can be taken into consideration in land planning and decision-making processes. Nonetheless, people's perceptions of landscapes are subjective and cannot easily be quantified [24]. Some authors have assessed the beauty of the landscape through groups of evaluators using images, videos and/or questionnaires [24, 19], while others used geographic information system (GIS) data such as elevation together with visual assessments and/or questionnaires to predict the scenicness [2, 21]. Another group of works, such as [3], [10], and [25], quantify landscape aesthetics according to the number of photos taken near a given location [3] or the number of people who published photos [10] in photo-sharing websites such as Flickr and Panoramio. Considering popularity on social media as a surrogate for the level of appreciation of a place might work with some types of landscapes, but the results might be liable to be biased towards more accessible places.

Another growing area of interest is in the use of social media data for ecological monitoring. For example, [1] examined Flickr biodiversity data quality by analysing its metadata and comparing it with ground-truth data, using Snowy owls and Monarch butterflies as a case study. They concluded that Flickr data has potential to add to knowledge of these species in terms of geographic, taxonomic, and temporal dimensions, which tends to be complementary to the information contained in other available sources. In another similar work, based on a manual analysis of Twitter posts, [5] confirm that social media mining for ecological analysis is as important as traditional monitoring and the features derived from Twitter could be integrated with and hence improve the value of existing sources of such information. In [20] the content of the Flickr photos was analysed manually to assess the quality of cultural ecosystem services and derive useful information to manage Singapore's mangroves.

Despite all the efforts that have been made to explore the role of social media in obtaining or supplementing environmental information, research in this area is fairly new and research questions about its effectiveness still remain open.

2.2 Geo-spatial analysis of social media

Many recent studies have focused on analysing social media data, with the aim of extracting useful information in domains such as geography (e.g. [13]). In particular, there is a large

[6] http://scenic.mysociety.org/

number of studies that derive such information from georeferenced Flickr photos. For example, [12] described two methods for the automatic delineation of imprecise regions based on geotagged photos. The first one is a method based on kernel density estimation (KDE) and the second is based on one class support vector machines (SVMs). Similarly, [4] presents an approach for automatically defining the geographic boundaries of vague regions by using one class support vector machines (SVMs) and learning multiple kernels. To describe regions, they rely on a combination of the Flickr tags of the photos that were tagged with the region's name, and external features such as the land cover data, population count, elevation and the geographical coordinates (latitude and longitude) of Flickr photos that are tagged with the region's name. They showed that their method performs better than the simpler methods described by [12]. Our work is analogous to these approaches, in applying support vector machine learning methods to Flickr tags in combination with other geo-spatial data, but we are concerned with characterizing and predicting information about the environment.

The authors of [22] presented and evaluated methods for automatically geo-referencing Flickr photos using the textual annotations of photos to predict the single most probable location where the image was taken. They showed that location-specific language models, based on sets of distinctive tags, can be estimated effectively by analysing the terms people use to describe images taken at particular locations. They demonstrated how to incorporate the GeoNames database and they defined extensions to improve their language models using tag-based smoothing and cell-based smoothing, and by leveraging spatial ambiguity. In [26], a language modelling approach was used to discover and characterize places of interest (POIs). They experimented with both Flickr data and Twitter data, finding that Flickr data on its own is more useful than Twitter data for this task, while combining both sources led to the best results. Similar to this latter work, we explore the possibility that sets of tags cannot just distinguish one location from another, but can contribute to classifying aspects of the environment.

3 Methodology

In the next section, we will consider a number of classification and regression problems that are aimed at assessing the value of Flickr as a source for environmental information. Here we first explain how locations are represented in these experiments. In particular, Section 3.1 explains how feature vectors describing locations can be obtained from the tags associated with georeferenced Flickr photos. In Section 3.2 we then describe what structured information sources will additionally be considered.

3.1 Modelling locations using Flickr tags

Many of the tags associated with Flickr photos tell us something about the locations where these photos were taken. For example, tags might refer to toponyms (e.g. United Kingdom, England, London), landmarks (e.g. London Eye, Westminster Abbey, Hyde Park) or land cover types (e.g. forest, beach, airport). Using the Flickr API, we collected the metadata of all geo-referenced Flickr photos that were uploaded before the end of September 2015, leading to a total of over 70 million photos with coordinates in Europe (which is the region our experiments will focus on).

Let $L = \{l_1, ..., l_m\}$ be a set of (point) locations, each characterized by latitude and longitude coordinates. Our aim is to associate with each of these locations a weighted bag of tags, intuitively encoding for each tag how often it occurs in photos near that location. To this end, we first use a BallTree[7] to retrieve the set F_l of all Flickr photos whose distance to the considered location l is at most D. Let us write $U_{t,c}$ for the set of users who have assigned tag t to a photo with coordinates c. Then we can define $n(t,l) = \sum_{d(c,l)\leq D} |U_{t,c}|$, with d the Haversine distance. Intuitively, $n(t,l)$ is the number of times tag t appears among the photos in F_l. However, to reduce the impact of bulk uploading, following [26], we count a tag occurrence only once for all photos by the same user at the same location.

One problem with using $n(t,l)$ to measure the importance of tag t for location l is that it gives equal weight to all photos, whereas intuitively we want photos which are closer to l to influence our characterization of l more than photos which are further away. To this end, following [26], we use a Gaussian kernel to weight the tag occurrences:

$$w(t,l) = \sum_{d(c,l)\leq D} |U_{t,c}| \cdot \exp\Big(-\frac{d^2(l,c)}{2\sigma^2}\Big)$$

where σ is a bandwidth parameter.

The weight $w(t,l)$ still has the problem that common words (e.g. *iphone*) are given the same importance as more specific words. Intuitively, we want the weight of tag t to reflect how strongly it is associated with location l. A standard way of measuring this in bag-of-words models is to use Positive Pointwise Mutual Information (PPMI), which essentially compares the actual number of occurrences with the expected number of occurrences (given how many tags occur overall near l and how common the tag t is). Specifically, the weight of tag t in our bag-of-words representation of l is then given by:

$$\mathit{PPMI}(t,l) = \max\left(0, \log\left(\frac{P(t,l)}{P(t)P(l)}\right)\right)$$

where:

$$P(t,l) = \frac{w(t,l)}{N} \quad P(t) = \frac{\sum_{l'\in L} w(t,l')}{N} \quad P(l) = \frac{\sum_{t'\in T} w(t',l)}{N} \quad N = \sum_{t'\in T}\sum_{l'\in L} w(t',l')$$

with T the set of all tags that appear in the collection. Finally, each location l is represented as a sparse vector, encoding the weights $\mathit{PPMI}(t,l)$ for all the tags in T.

3.2 Modelling locations using structured data

There is a wide variety of structured data that can be used to describe places. The most obvious type of structured data are the coordinates of the photo itself. Clearly, latitude and longitude degrees can be helpful for predicting a range of environmental phenomena (e.g. Southern areas of Europe tend to be warmer than Northern areas). In addition to geographic coordinates, we will consider the following sources of structured scientific data:

- CORINE Land Cover 2006[8] is a European dataset which describes land cover with a 100-meter spatial resolution. CORINE uses three levels of description: a top level with 5 classes, an intermediate level with 15 classes and a detailed level with 44 classes.
- SoilGrids[9] is a global raster dataset, which classifies locations into 116 types of soil, using a 250 meter spatial resolution

[7] http://scikit-learn.org/stable/modules/generated/sklearn.neighbors.BallTree.htm
[8] http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-2
[9] https://www.soilgrids.org

- The Digital Elevation Model over Europe (EU-DEM)[10] is a Europe-wide digital surface model, encoding elevation with a spatial resolution of about 30 meter.
- European Population Map 2006[11] is a digital raster grid that reports the number of residents (night-time population) with a 100-meter spatial resolution.
- WorldClim[12] is a global raster dataset, containing average monthly recordings of the following climate features, over the period 1970-2000, using a 1 km spatial resolution: temperature, precipitation, solar radiation, wind speed and water vapor pressure. In this work, we convert the monthly averages reported in the dataset to a single annual average.

To encode locations, we consider a feature vector that contains one binary feature for each CORINE land cover class (being 1 if the location belongs to that class and 0 otherwise), one binary feature for each SoilGrids class, and 9 real-valued features (encoding latitude, longitude, elevation, population, and the 5 climate features). The real-valued features have been normalised using the standard z-score. In experiments where both Flickr data and structured data are used, we simply concatenate the two corresponding feature vectors.

4 Experiments

In the following experiments, we evaluate how well we can predict a number of environmental features using Flickr tags and the considered structured data. For all experiments, we have set the maximum Haversine distance D (cluster radius) to 1 kilometre and the bandwidth σ to D/3. The choice of D represents a trade-off, where larger values can potentially lead to better results but also lead to a higher computational cost. The choice of $\sigma = D/3$ was found to be reasonable in a small set of initial experiments. To make predictions, we use Support Vector Machines (SVMs) for classification problems and Support Vector Regression (SVR) for regression problems. In both cases, we used the SVM^{light} implementation[13][15]. For each experiment, the set of locations L was split into two-thirds for training, one-sixth for tuning the parameters of the SVM/SVR models, and one-sixth for testing.

4.1 Predicting the scenicness of a place

In this first experiment, we consider the problem of predicting people's opinions of landscape beauty, using the UGC dataset from the ScenicOrNot website[14] as ground truth. This website allows people to evaluate places in Britain by rating photos collected from Geograph[15]. The dataset contains 217,000 photos (at distinct locations), each of which has been rated by at least three people on a scale from 1 (not scenic) to 10 (very scenic). For 25,395 of the photos in this dataset, our Flickr collection did not contain any georeferenced photos within a 1 km radius. Therefore, we only report results for the remaining 191,605 photos (i.e. 88.3% of the full dataset). The number of Flickr photos within a 1km radius of these locations varies between 1 and 397982.

For this experiment, L thus contains the locations of these 191,605 photos. Table 1 shows the results for three different variants: only using structured data, only using Flickr data, and combining both. Based on the tuning data, for the SVR model, we found a Gaussian

[10] `http://www.eea.europa.eu/data-and-maps/data/eu-dem`
[11] `http://data.europa.eu/89h/jrc-luisa-europopmap06`
[12] `http://worldclim.org`
[13] `http://www.cs.cornell.edu/people/tj/svm_light/`
[14] `http://scenic.mysociety.org/`
[15] `http://www.geograph.org.uk/`

Table 1 Results for predicting scenicness.

Dataset	Mean Absolute Error	Spearman ρ
Structured	1.031	0.556
Flickr	1.013	0.570
Structured + Flickr	1.006	0.581

kernel to be optimal when only structured data is used, and a linear kernel to be optimal otherwise. The results in Table 1 show the mean absolute error between the predicted and actual scenicness scores, as well as the Spearman ρ correlation between the rankings induced by both sets of scores. Note that the mean value of this data set is 4.372 and the standard deviations is around 1.6. While the differences are small, we find that using Flickr outperforms using structured data, and that combining both leads to the best results overall. Looking at what tags most influence the regression model, among the highest weighted tags we find terms relating to natural and open-country landscape such as *scotland*, *highlands*, *mountains* and *sea*, while among the lowest weighted tags we find names of artificial and urban phenomena such as *station*, *bus*, *pub* and *railway*. This reinforces the finding from [24] that land cover categories are strongly correlated with scenicness scores.

We also tested whether the number of photos (or users) could be used to predict scenicness, as was suggested in [3, 25, 10] for particular restricted settings. However, we actually found a negative correlation of around -0.12 (resp. -0.1) between scenicness and the number of photos (resp. users who have posted photos) near a given location.

4.2 Predicting species distribution

The next experiment we considered was to predict the distribution of species across Europe, using as ground truth the dataset of the European network of nature protected sites Natura 2000[16]. This dataset contains information about 35,600 species from 7 classes: Amphibians, Birds, Fish, Invertebrates, Mammals, Plants and Reptilia. In particular, it specifies which species occur at 26,425 different sites across Europe. For this experiment, L is defined as the set of these sites.

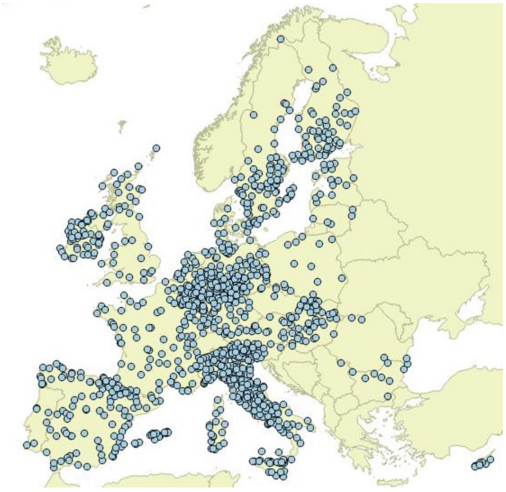
For species that only occur at a few of the sites in L, it is clearly not possible to estimate a reliable distribution model. Therefore, we focused our evaluation on 100 species which occur at more than 500 sites. For each of these species, we consider a binary classification problem, i.e. predicting at which of the sites the species occurs. Note that as in all analyses we use all Flickr tags, some of which might include the species name. The results are reported in Table 2, showing that combining structured data with Flickr data leads to substantially better results than either structured data alone or Flickr data alone. Comparing Flickr with structured data directly is more difficult, as Flickr data led to a much higher precision whereas the structured data led to a much higher recall.

As an example, Figure 1 compares the predictions that were made by the different models with the ground truth for a particular species: the black woodpecker (dryocopus martius). For this species, the F1 scores were 0.594, 0.648 and 0.927 for structured data, Flickr data, and the combined data, respectively. This example shows that highly accurate distribution models can be learned for species that occur in sufficiently many sites. Interestingly, while the number of occurrences is overestimated in e.g. Spain and the UK when only Flickr data

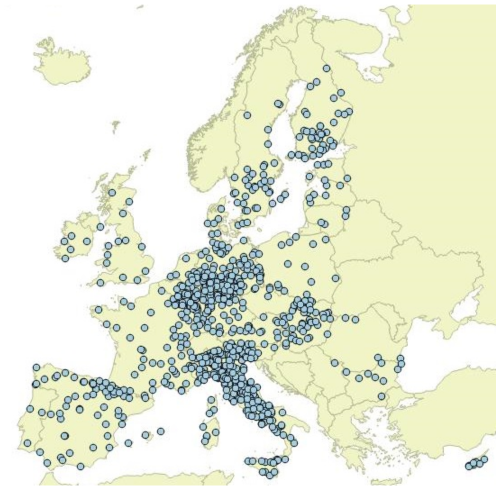
[16] `http://ec.europa.eu/environment/nature/natura2000/index_en.htm`

Table 2 Results for predicting species distribution.

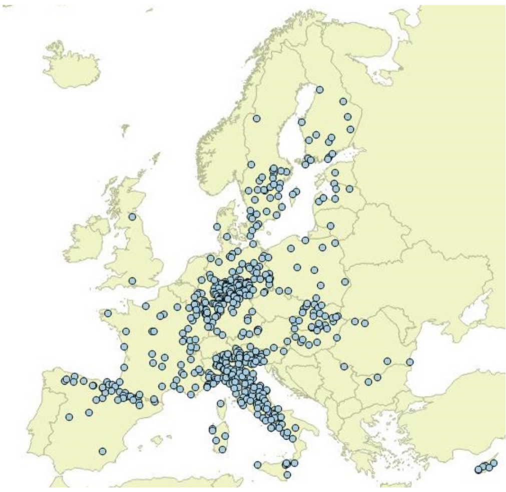
Dataset	Precision	Recall	F1 Score
Structured	0.241	0.568	0.338
Flickr	0.577	0.112	0.188
Structured + Flickr	0.650	0.506	0.569



(a) Structured data



(b) Flickr



(c) Combination of structured data and Flickr

(d) Ground truth data

Figure 1 Prediction of the black woodpecker distribution across Europe.

or only structured data is used, much more accurate predictions are made for these countries using the combined model. For species that have a more restricted geographic scope (in terms of number of sites), it is likely that better results can be obtained by looking at a wider radius (i.e. choosing $D \gg 1$ km) and by specifically counting photos that mention the name of the species, as a separate feature. This is left as an issue for future work, however, as in this paper our focus is on assessing the overall usefulness of Flickr.

Table 3 Results for predicting CORINE land cover classes, at levels 1, 2 and 3.

	Level 1			Level 2			Level 3		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Structured	0.437	0.363	0.397	0.346	0.160	0.219	0.207	0.070	0.105
Flickr	0.499	0.457	0.477	0.205	0.139	0.166	0.145	0.086	0.108
Structured + Flickr	0.523	0.514	0.518	0.270	0.199	0.229	0.184	0.112	0.139

Table 4 Top 5 Flickr tags for CORINE level 1 classes in the SVM models.

Artificial surfaces	Agricult. areas	Forest & semi nat. areas	Wetlands	Water bodies
Babenhausen	field	wald	bog	lake
Ceskedrahy	grass	forest	moor	island
Meppen	horse	mountains	marsh	sea
Tuplice	vineyard	woods	swamp	boat
Deutsche Reichsbahn	meadow	mountain	saline	sailing

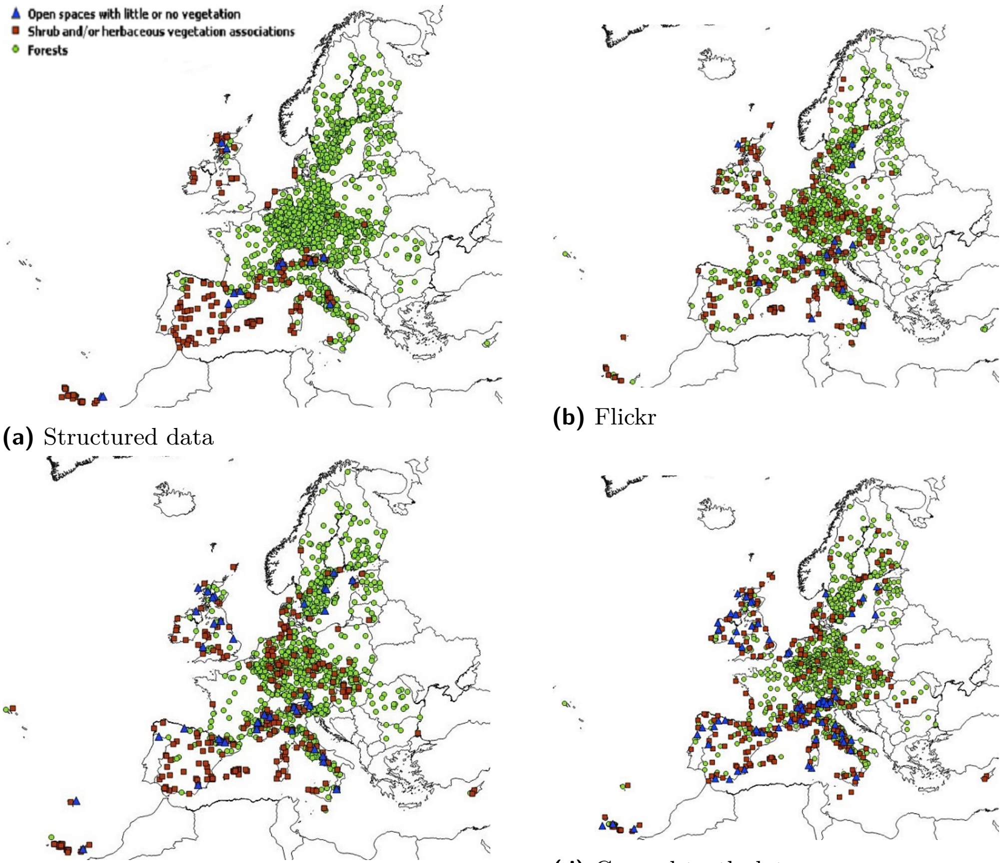
4.3 Predicting CORINE land cover classes

In this Section we consider the task of predicting CORINE land cover classes. For this experiment, we have used the same set L of sites as for species distribution. Since the task is about predicting CORINE land cover classes, for the results reported in this section we do not consider any CORINE features in the representations of the locations (as the CORINE data serve here as ground truth). We experimented with predicting CORINE land cover classification at levels 1, 2 and level 3, each time treating the task as a binary classification problem. The results are presented in Table 3, showing again that combining structured data and Flickr data clearly leads to the best results. The difference in performance between structured data alone and Flickr data alone is mixed, with e.g. Flickr data performing better at level 1 but worse at level 2. For level 1, we found that Flickr outperformed structured data in 4 out of the 5 classes, with the 'artificial surfaces' class being the only exception. This seems related to the small number of sites for this particular class (e.g. only 4% of the training data sites belong to this class). To illustrate how Flickr tags are used to predict CORINE classes, Table 4 shows the 5 tags with highest weight in the SVM classifier for each of the classes at level 1.

By far the largest CORINE class at level 1 is 'Forest & semi natural areas'. At level 2 this class has three subclasses. The predictions of the three models for these three subclasses are compared with the ground truth in Figure 2. Clearly, in this case, the structured data has resulted in a model that is too simplistic, essentially segmenting Europe into forest areas and 'Shrub and/or herbaceous vegetation'. Flickr data alone leads to more faithful predictions in these subclasses, but instances of 'open spaces with little or no vegetation' are underreported. This issue is alleviated in the combined model.

4.4 Predicting climate data

In the last experiment, we assess the usefulness of Flickr tags in the task of predicting climate data. We again use the same set of sites L as in the species distribution experiment. In this case, we omit all the climate related features from the feature vector representations as they constitute the ground truth. We consider 5 different regression problems: predicting average temperature, average precipitation, average solar radiation, average wind speed, and



(a) Structured data

(b) Flickr

(c) Combination of structured data and Flickr

(d) Ground truth data

Figure 2 Prediction of subclasses of the CORINE class "Forest & semi natural areas".

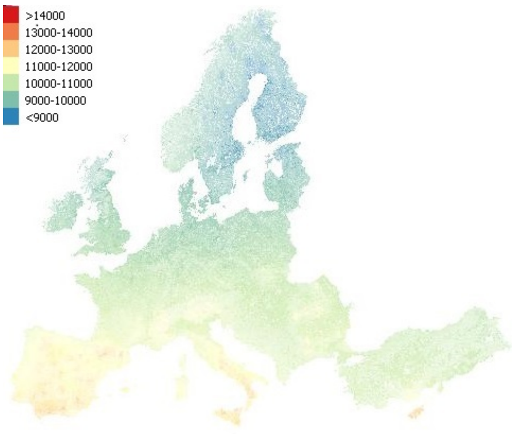
average water vapor pressure. The results are reported in Table 5, in terms of mean absolute error (MAE) and Spearman ρ. Overall, structured data and Flickr data perform comparably. However, by far the best results are obtained when combining both types of data, showing again that the information we obtain from Flickr is complementary to what is available as structured data. As an example of how Flickr tags are used by the regression model, the tag 'sea' has a very high weight in the model for predicting water vapor pressure, while the tag 'mountain' has a very low weight in this model. In Figure 3, we illustrate the predictions made by the different models for solar radiation. Clearly, the model based on structured data is too simplistic, mostly capturing the impact of latitude.

5 Conclusions and future work

In this paper, we have analysed how Flickr tags can be used to supplement structured scientific data in tasks that rely on characterizing the environment. To this end, we have considered four different evaluation tasks. The first experiment aimed to predict the scenicness of a place, as assessed subjectively by humans on the ScenicOrNot website. In the second experiment, we focused on modelling the distribution of species across Europe, using observations from the Natura 2000 dataset as ground truth. The third experiment consisted in predicting CORINE land cover categories. Finally, we looked at predicting five climate related properties. Each

Table 5 Results for predicting average climate data.

	Mean value	STDEV	Structured MAE	Structured ρ	Flickr MAE	Flickr ρ	Struct+ Flickr MAE	Struct+ Flickr ρ
Temperature(°C)	9.268	3.490	0.789	0.938	1.623	0.814	0.728	0.940
Precipitation(mm)	66.625	24.827	13.173	0.709	11.660	0.689	10.523	0.755
Solar Rad(kJ $m^{-2}day^{-1}$)	11478	2388	1726.5	0.747	926.3	0.832	484.8	0.939
Wind Speed(m s^{-1})	3.605	1.126	0.508	0.791	0.545	0.756	0.429	0.846
Water Vapor Press(kPa)	0.958	0.186	0.060	0.903	0.083	0.719	0.053	0.914



(a) Structured data

(b) Flickr

(c) Combination of structured data and Flickr

(d) Ground truth data

Figure 3 Prediction of solar radiation (in kJ $m^{-2}day^{-1}$).

time, we compared three different setups. In a first setup, we used features that were derived from a number of structured scientific datasets. In the second setup, we used a bag of words representation, capturing how strongly each tag is associated with photos that appear near a considered location. In the final setup, we combined both data sources, concatenating the corresponding feature vectors.

Our main finding is that the combined model substantially and consistently outperformed the model that only relied on structured data sources. This strongly suggests that Flickr can indeed be valuable, as a supplement to more traditional datasets in environmental analyses. While it may be possible to reduce some of the performance gap by considering additional scientific datasets, we found the versatility of Flickr data that was displayed in the four experiments to be remarkable.

There are a number of directions for future work. First, it may be possible to improve the way we have combined structured features with bag-of-words features by learning a low-dimensional vector space embedding that captures both kinds of data, similar to how embeddings of Wikipedia entities were learned in [14] by combining bag-of-words representations with semantic descriptions from WikiData. Second, many of the considered features are strongly spatially autocorrelated. As such, we can expect more accurate predictions by formulating some of the considered tasks as collective classification problems, where we would intuitively take into account the predictions for neighbouring sites when making a prediction. Finally, it remains unclear to what extent Flickr can be used for more fine-grained ecological analyses, e.g. at the level of individual sightings.

References

1 Vijay V. Barve. *Discovering and developing primary biodiversity data from social networking sites*. PhD thesis, University of Kansas, 2015.

2 Ian D. Bishop and David W. Hulse. Prediction of scenic beauty using mapped data and geographic information systems. *Landscape and urban planning*, 30(1-2):59–70, 1994.

3 Stefano Casalegno, Richard Inger, Caitlin DeSilvey, and Kevin J. Gaston. Spatial covariance between aesthetic value & other ecosystem services. *PloS one*, 8(6):e68437, 2013.

4 Eduardo Cunha and Bruno Martins. Using one-class classifiers and multiple kernel learning for defining imprecise geographic regions. *International Journal of Geographical Information Science*, 28(11):2220–2241, 2014.

5 Stefan Daume. Mining Twitter to monitor Invasive Alien Species – An analytical framework and sample information topologies. *Ecological Informatics*, 31:70–82, 2016.

6 Janis L. Dickinson, Benjamin Zuckerberg, and David N. Bonter. Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, 41:149–172, Jan-12-2010 2010.

7 Jacinto Estima, Cidália C Fonte, and Marco Painho. Comparative study of Land Use/-Cover classification using Flickr photos, satellite imagery and Corine Land Cover database. In *Proceedings of the 17th AGILE International Conference on Geographic Information Science, Castellon, Spain*, pages 1–6, 2014.

8 Jacinto Estima and Marco Painho. Photo based Volunteered Geographic Information initiatives: A comparative study of their suitability for helping quality control of Corine Land Cover. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, 5(3):73–89, 2014.

9 Steffen Fritz, Ian McCallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. van der Velde, F. Kraxner, and M. Obersteiner". Geo-wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31:110–123, 2012.

10 Gianfranco Gliozzo, Nathalie Pettorelli, and Mordechai Muki Haklay. Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society*, 21(3), 2016.

11 Michael F. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

12 Christian Grothe and Jochen Schaab. Automated footprint generation from geotags with kernel density estimation and support vector machines. *Spatial Cognition & Computation*, 9(3):195–211, 2009.

13 Livia Hollenstein and Ross Purves. Exploring place through user-generated content: Using flickr tags to describe city cores. *Journal of Spatial Information Science*, 1, 2010.
14 Shoaib Jameel and Steven Schockaert. Entity embeddings with conceptual subspaces as a basis for plausible reasoning. In *Proceedings of the 22nd European Conference on Artificial Intelligence*, pages 1353–1361, 2016.
15 Thorsten Joachims. Making large-scale SVM learning practical. Technical report, SFB 475: Komplexitätsreduktion in Multivariaten Datenstrukturen, Universität Dortmund, 1998.
16 Daniel Leung and Shawn Newsam. Exploring geotagged images for land-use classification. In *Proceedings of the ACM multimedia 2012 workshop on Geotagging and its applications in multimedia*, pages 3–8, 2012.
17 Christopher S. Lowry and Michael N. Fienen. Crowdhydrology: Crowdsourcing hydrologic data and engaging citizen scientists. *Ground Water*, 51(1):151–156, 2013.
18 Michael J. Paul, Abeed Sarker, John S. Brownstein, Azadeh Nikfarjam, Matthew Scotch, Karen L. Smith, and Graciela Gonzalez. Social media mining for public health monitoring and surveillance. In *Pacific Symposium on Biocomputing*, pages 468–79, 2016.
19 Chad D. Pierskalla, Jinyang Deng, and Jason M. Siniscalchi. Examining the product and process of scenic beauty evaluations using moment-to-moment data and GIS: The case of Savannah, GA. *Urban Forestry & Urban Greening*, 19:212–222, 2016.
20 Daniel R. Richards and Daniel A. Friess. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: content analysis of social media photographs. *Ecological Indicators*, 53:187–195, 2015.
21 Uta Schirpke, Erich Tasser, and Ulrike Tappeiner. Predicting scenic beauty of mountain regions. *Landscape and Urban Planning*, 111:1–12, 2013.
22 Pavel Serdyukov, Vanessa Murdock, and Roelof Van Zwol. Placing Flickr photos on a map. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 484–491, 2009.
23 S. Andrew Sheppard, Andrea Wiggins, and Loren Terveen. Capturing quality: retaining provenance for curated volunteer monitoring data. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1234–1245, 2014.
24 B. Stadler, R. Purves, and M. Tomko. Exploring the relationship between land cover and subjective evaluation of scenic beauty through user generated content. In *Proceedings of the 25th International Cartographic Conference*, 2011.
25 Patrizia Tenerelli, Urška Demšar, and Sandra Luque. Crowdsourcing indicators for cultural ecosystem services: a geographically weighted approach for mountain landscapes. *Ecological Indicators*, 64:237–248, 2016.
26 Steven Van Canneyt, Steven Schockaert, and Bart Dhoedt. Discovering and characterizing places of interest using Flickr and Twitter. *Hospitality, Travel, and Tourism: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, page 393, 2014.
27 Jingya Wang, Mohammed Korayem, and David Crandall. Observing the natural world with Flickr. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 452–459, 2013.
28 Haipeng Zhang, Mohammed Korayem, David J. Crandall, and Gretchen LeBuhn. Mining photo-sharing websites to study ecological phenomena. In *Proceedings of the 21st international conference on World Wide Web*, pages 749–758, 2012.

Defining Local Experts: Geographical Expertise as a Basis for Geographic Information Quality

Colin Robertson[1] and Rob Feick[2]

1 **Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, ON, Canada**
`crobertson@wlu.ca`

2 **School of Planning, University of Waterloo, Waterloo, ON, Canada**
`robert.feick@uwaterloo.ca`

Abstract

As more data are produced by location sensors, mobile devices, and online participatory processes, the field of GIScience has grappled with issues of information quality, context, and appropriate analytical approaches for data with heterogeneous and/or unknown provenance. Data quality has often been viewed through a bifurcated lens of experts and amateurs, but consideration of what the nature of geographical expertise is reveals a much more more nuanced situation. We consider how adapting frameworks from the field of studies of experience and expertise may provide a conceptual basis and methodological framework for evaluating the quality of geographic information. For contributed geographic information, quality is typically derived from a data user's trust in and/or perception of the reputation of the data producer. Trust and reputation of producers of geographic information has typically been derived from the presence or absence of professional qualifications and training. However this framework applies exclusively to 'crisp' notions of data quality, and has limited utility for more subjective contributions associated with volunteered geographic information which may provide a rich source of geographic information for many applications. We hypothesize that a conceptual framework for geographical expertise may be used as the basis for assessing information quality in both formal and informal sources of geospatial data. Two case studies are used to highlight the new concepts of geographical expertise introduced in the paper.

1998 ACM Subject Classification E.0 [Data] General

Keywords and phrases data quality, expertise, geographic information, conceptual framework

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.22

1 Introduction

What does it mean to be a 'local expert'? Local knowledge and expertise has long been a valued input into geographical inquiry, and geographic information systems and geoweb technologies are instrumental in acquiring, structuring, representing and disseminating georeferenced local knowledge in many application areas. The notion of local expertise rests on ideas of place [34], familiarity and personal experiences with locales [24]. These are distinct from and related to more formal types of geographical expertise (GE) such as metric or topological relations between and among places on the earth's surface. As geographical knowledge of all forms is increasingly codified into a variety of information products, a reconsideration of GE itself, what defines GE and how is it obtained, may provide a framework for evaluating and integrating heterogeneous sources of geospatial data.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 22; pp. 22:1–22:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2 Related Work

2.1 Geographic Data Quality

Interest in approaches to conceptualize, measure, and improve spatial data quality has been a longstanding concern to GIScience research and practice. Devillers and Jeansolin [12] note that spatial data quality is often viewed from two interrelated perspectives - internal (producer) and external (user). Internal spatial data quality focuses on the characteristics of spatial data as a product and is concerned with measuring and documenting quality relative to known specifications that a dataset must satisfy. Specifications have a dual role since they provide data producers guidance for ensuring internal quality throughout data capture, compilation, and documentation procedures and they provide benchmarks for validating the quality of data products. The widely used ISO reference standard (ISO 19157:2013) decomposes internal data quality across six main dimensions (positional accuracy, thematic accuracy, completeness, logical consistency, temporal quality, usability) and outlines a corresponding set of quantifiable data quality measures.

External data quality, by contrast, focuses on the degree of correspondence between a user's needs in a given application environment and a dataset's internal characteristics [12]. Chrisman [5, 6] popularized the concept of viewing external quality in terms of 'fitness for use' to recognize that quality requirements vary among users and across different tasks. While intuitively appealing and conceptually elegant, broad-based implementation of the fitness for use concept has proven to be challenging in practice. User needs, even within a well-defined application context, are often diverse and dynamic as people gain a better understanding of the problem at hand or as decision making processes progress through successive stages (e.g. problem scoping, evaluation of options, etc. in land planning). Progress is being made within specific subfields and defined task arenas, however [27] note that additional research is required to develop domain ontologies and easy-to-use tools that enable users to define and assess fitness for use.

The past decade has seen our understanding of fitness for use and spatial data quality more generally challenged by increasing variety in what constitutes geographic data, who creates these data and how they are produced. The growth of personal computing and location aware devices has fuelled crowdsourcing, citizen science and volunteered geographic information (VGI) activities that have resulted in new sources and types of geographic data pertinent to many existing and novel applications. However, since the data sets that emerge from these processes are not products of a single entity with known and reliable data quality procedures (e.g. national mapping agency), traditional measures of data quality are not applicable. Heterogeneity is perhaps the common characteristics of these data sets. A single data set can contain wiki-like edits from many individuals who differ in terms of skills, motivations, and knowledge, data capture methods (e.g. screen digitizing, GPS), and geographic areas of interest [35]. Some types of VGI, such as OpenStreetMap, have a loosely-defined schema that provides a foundation for consistency that also allows for local adaptations [26], while others that are by-products of social processes (e.g. Twitter posts, geotagged photos, GPS trail mapping) are largely unstructured. In this context, quality measures such as accuracy, lineage and consistency can vary with contributor, area and on a feature-by-feature basis.

Since the heterogeneity within and across VGI and citizen science data sets precludes the use of formal internal data quality measures, quality assessment shifts from evaluating data products toward more inferred and producer-centric foci [32]. For example, the quality of OpenStreetMap data in a locale is often inferred from the number of local contributors or with reference to deviations in coverage or positional accuracy relative to expert-generated comparators [21, 15]. Many VGI sources centre on phenomena that are not collected by authoritative agencies (e.g. bird sightings) or contain more subjective observations (e.g. geotagged photos and comments about camping experiences) that cannot be ground truthed, but may be considered to potentially exhibit a degree of 'local expertise'. In both of these instances, the quality of data elements is often judged based on contributors' attributes such as formal qualifications, trustworthiness, reputation, or credibility [32, 28]. Bishr and Kuhn [1] refer to this transitivity between trust in an individual and trust in the spatial data they create as informational trust, which they demonstrate can vary spatially and temporally and Keßler and De Groot [23] extend to the feature level in OSM.

These more nuanced approaches to understanding the expertise of individuals (i.e. persons or single entities) to create spatial data that are fit for specific uses and locales have interesting parallels with what Golledge [18] referred to as the changing nature of geographic knowledge. The dominant change in geographic knowledge [18] identified was "a change from inventory dominated activity to the creation of knowledge generated by emphasizing cognitive demands" such as processes of logical reasoning, deduction, and geographic association. Our current data-rich environment is seemingly more focused on inventorying than searching for understanding, in part because the former is technologically and socially accessible. To address current data quality challenges, Goodchild and Li [20] advocate for a knowledge-based approach where geographic concepts are used to assess data quality. Operationalizing this notion of geographic knowledge is challenging and, we suggest, requires consideration of individuals' differing expertise to expertise across a range of activities such as inventorying ubiquitous features and facts (e.g. trace building outline) or contributing more specialized and locale-specific information (e.g. document biodiversity in local wetland).

2.2 Studies of Experience and Expertise

Academic study of expertise and experience has been the traditional domain of the sociology of science. In Collins and Evans' seminal paper on SEE [11], the distinction between contributory and interactional expertise within a scientific field was introduced. Embedded in this juxtaposition are two elements of expertise in science: the knowledge and capability to make contributions to the field, and the ability to participate and interact with other actors in the field. Contributory expertise is generally acquired through formal training and education and working within the domain of interest, while interactional expertise is gained through socialization and exposure to tacit knowledge. The traditional view of expertise is a one-dimensional construct based on accumulation of ability and experience, which leads to phase or stage-based models of expertise development [13]. Critically, this expertise is typically only recognized if it is acquired through educational programs that initiate socialization and immersion into the society of experts (i.e., contributory expertise can only be recognized through interactional expertise).

Some recent work in citizen science has attempted to classify aspects of citizen science in such one-dimensional models. Haklay's [22] typology of citizen scientists is an recent example of this, presenting a typology of citizen science projects from passive crowdsourcing to collaborative science. Coleman's [8] typology focuses on individual motivations, extending from neophyte through to expert amateur. Goodchild [19] is a rare paper explicitly examining GE itself, yet fails to provide any framework or methodology for evaluating or characterizing it, while highlighting the critical need for theorizing GE in the geoweb and neogeography era.

The motivating rationale for many VGI and citizen science projects is that many individuals may possess expertise that could be valuable as input into pressing environmental and societal issues. Collins and Evans [11] describe Wynne's [37] study, which contrasted the expertise of sheep farmers without qualifications with government scientists in the aftermath of a nuclear contamination incident. The study showed that sheep farmers had specialized contributory expertise relevant to the ecology of sheep in the region. More recently, Maderson and Wynne-Jones [25] describe the role of beekeeper knowledge in understanding causes of and solutions to colony collapse disorder. In many cases, the criterion for expertise relevant to an environmental problem should be *experience* rather than normative qualifications.

Collins [9] derives, based on long-term studies into expertise, three dimensions of expertise: that which is attained through accumulation of experience and enable ability to contribute to a specific domain (i.e. the traditional form), the tacit knowledge of a domain that can only be acquired through socialization within that domain, and the 'esotericity', or the degree to which the domain is esoteric (e.g., gravitational physics vs. car driving). This results in a three dimensional model of expertise, termed the Expertise-Space-Diagram (ESD) based on its graphical representation, providing a conceptual tool to investigate expertise (Figure 1a). In this paper, we explore the ESD for deconstructing and analyzing GE in a variety of contexts and forms.

2.3 Towards Geographical Expertise

The nature of GE in the context of information quality has not been explored in great detail. The data quality literature in GIScience is dominated by a paradigm borrowed from transactional data architectures where data models are defined *a priori* and discrepancy metrics can be formulated easily. Unfortunately, this approach has limited utility for messy, heterogeneous information sources, where often the question being asked is 'what can I do with these data?' rather than 'does this data meet the requirements of this application?'.

As discussed above, expertise can be defined along three dimensions; contributory, interactional, and esotericity. The forms of expertise commonly represented in geographic information vary widely, yet these concepts have not been formally incorporated into approaches for evaluating information quality. Creators of geographic information may have any combination of levels of expertise as it relates to a given type of geographic information. GE that is place-based, contextual, and general in nature today tends to be derived from experience and less frequently from training in regional specializations in academic geography. Goodchild [19] argues that locale familiarity was a cornerstone of GE in the regional tradition, and this has been greatly democratized by increasing travel, allowing more people to become familiar with more places. Many studies that employ analysis of place-based social media data attempt to capture local expertise [3, 4]. As well, many forms of indigenous knowledge is place-based, grounded in narratives and experiences, and conferred through traditions and oral histories [31]. Locale-familiarity, or place-based expertise, is a dimension of GE which has to do with experience of a particular place. This type of naïve geographical knowledge tends to be approximate, more topological than metric, and often prone to biases [14]. From their inception, GIS has had a difficult time representing these forms of knowledge [18], despite concerted research efforts to develop their capabilities. Spatial operations designed for crisp data models also have difficulty handling fuzzier spatial information sources representing these forms of expertise and new 'platial' analogues is currently an active research area [29, 16].

A related, but distinct form of GE is that which relates to knowledge of geographical archetypes. With the shift to understanding of geographic processes rather than regions, geographers have developed expertise in the outcome of spatial processes and forms thereof. We refer to groupings of these geographical expressions as place-types. A component of GE that relates to knowledge of place-types is a core aspect of many academic geographers. For example, experts in grassland ecosystems or suburban sprawl, may not have any experience with a particular locale, but will be able to use knowledge of underlying processes to make expert judgements and understanding of places. Valuable geographic information could therefore be contributed by individuals with high locale-familiarity or high place-type knowledge.

Geographical expertise that is technical, pertaining to the tools and practices of producing accurate geographic data has, until recently, been almost exclusively held by professionals in surveying and mapping sciences. However, many forms of data production that used to require such technical expertise no longer do, as new technologies simplify many tasks of geographic data production [19]. As some forms of expertise are attained by more people via their use of simple GIS or web-based editing tools, according to the Collins [9] model of expertise described above, this represents a change in the type of expertise required for this task (via reduced 'esotericity'), from specialist knowledge to ubiquitous expertise. This is a critical contribution of the ESD model, in that expertise itself is deconstructed relative to its ubiquity, such that expertise moves from being only possible for a select few (i.e., those that participate in an esoteric domain) to almost everyone. This model of expertise underlies more inclusive science-society relationships in general, as exemplified by the transition in citizen science from citizens-as-sensors (i.e., data collectors) to higher levels of citizen participation in design, scoping, analysis, and interpretation [2].

As GE pertains to the processes and patterns at or near the earth's surface, we provide a conceptual model of GE based on a translation of the ESD model in Figure 1. The key difference here is replacing the notion of specialist tacit knowledge in ESD with locale-familiarity and place-type knowledge on the X-axis in our model of GE. The tacit knowledge in SEE 'can be acquired only by immersion in the society of those who already possess it'. Locale-familiarity/place-type tacit knowledge can be attained by immersion within the locale of interest or, through immersion and study of locales of a similar place-type. While interactional expertise with a community of experts maps directly to the model of SEE, a prolonged immersion within a geographic locale can alone provide sufficient experience to enable locale-familiarity knowledge. Note also that locale-familiarity knowledge can indeed always be enhanced through socialization within a community of locale-familiarity experts, according a locale-familiarity-type of interactional expertise. Many individuals will have both locale-familiarity and place-type knowledge, as place-type is the generalization of locale-familiarity expertise into archetypes. In short, place-type expertise emphasizes the common characteristics across distinct locales, while locale-familiarity emphasizes the local uniqueness of places. The Y-axis pertains to the traditional view of expertise, as ability to make contributions to the field and is typically attained and/or recognized through formal education and training and accomplishments such as publications, expert testimony, posts on editorial boards, professional reputation etc. Finally, the Z-dimensions we have renamed to 'thematic specificity' which is how general or specific a geographically-based topic or issue is. For example, an individual with knowledge of major landmarks in a city would have lower thematic specificity than one with deep understanding of its road network evolution or immigrant social services network. As with the ESD, both specialist and ubiquitous theme specificity constitute valid forms of GE.

In order to link GE at the individual level to a conceptual framework for evaluating information quality, we need to consider how expertise manifests in geographic information and data products. Single-producer data can be assessed relative to the GE of the contributor. Note that single-producer in this context includes entities like national mapping agencies

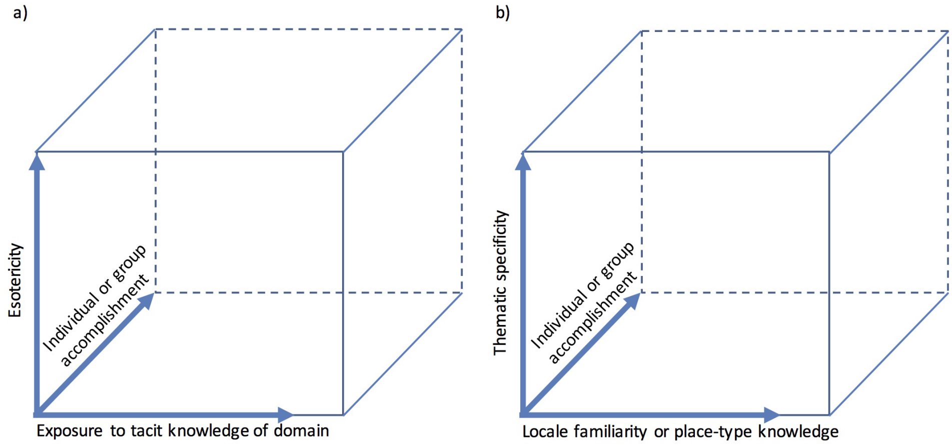


Figure 1 The dimensions of expertise according to the a) expertise-space diagram (ESD) and b) geographical expertise-space diagram (GESD).

that apply uniform data quality practices internally across their staff. Multi-authored data where contributors are at best loosely coordinated can be considered as a composite source of individual levels of GE.

Building on the Imitation Game methodology of SEE [10], if we consider two data products for an area such as two representations of a road centreline dataset in a city; one produced by professional mapping surveyors and another produced by several volunteers via OpenStreetMap. If both datasets are provided to quality control technicians for evaluation, and deemed to be of similar quality or potential utility, we can confer a level of expertise to the volunteer group based on their ability to produce a data product of equal value as traditional expert data. Such comparisons have been made in the literature many times, in efforts to demonstrate the quality of VGI or citizen-produced data. However the key point here is not the search for specific metric values but rather the interpretation of an information product by another member of the expert group. A related approach for local contribution data common of VGI, could be developed whereby community members judge each others' GE through collaborative experimental methods such as the imitation game.

3 A Framework for Geographical Expertise and Information Quality

Conceptualizing GE as a multi-dimensional surface provides a tool to position geographical information products according to the relative expertise of their authors in terms of their knowledge of locales, GI production, and the thematic specificity of the topic being represented. Such a framing enables deeper questions about information quality than what is typical for the fitness-for-use paradigm. As well, given the 'neogeographical' trend and shift towards heterogeneous data collectors and global scale data projects, explicit consideration of the components of GE articulates a more precise definition of quality in a given context, irrespective of potential application. The ESD for GI production is outlined in Figure 2, with the only difference from Figure 1 being the definition of Y-axis according to knowledge of geographical information production, ranging from those with no training and/or professional experience to GI professionals (e.g., surveyors, GI researchers).

To illustrate, four potential GI contributors are positioned within the GESD model, signified by letters. Person A might be a soil scientist collecting soil samples data in an

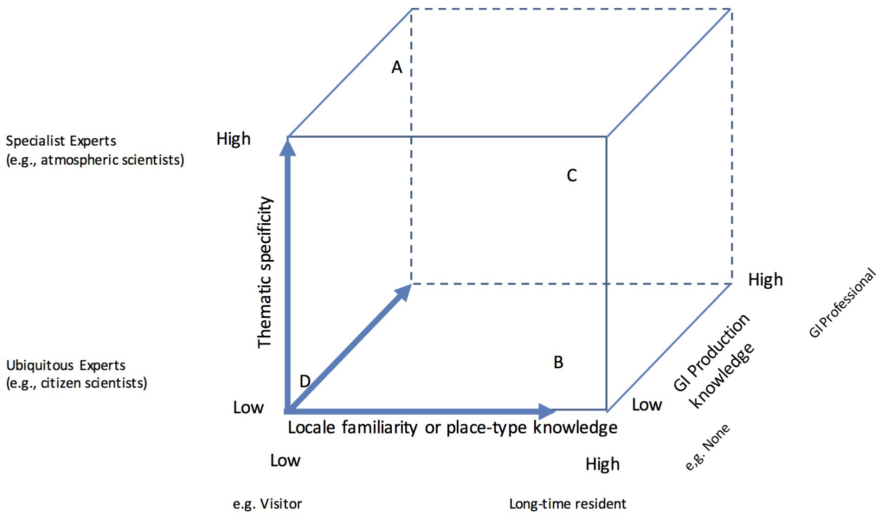


Figure 2 Geographical expertise diagram for geographical information production.

unfamiliar ecotype, exhibiting high thematic-specificity, low locale-familiarity/place-type, and high geographic-information-knowledge. Due to the high thematic-specificity of the GI production task, high geographic-information-knowledge is particularly important, whereas locale-familiarity/place-type is less so. Person B in Figure 2 could be a long-term resident inventorying heritage buildings without guidance, a task of low thematic-specificity, with high locale-familiarity/place-type and low geographic-information-knowledge associated with collecting georeferenced photographs of heritage buildings, basic data structuring, etc. Person C could be a local biologist creating a map of at-risk habitat based on a mix of input data, exhibiting high thematic-specificity, high locale-familiarity/place-type and relatively moderate geographic-information-knowledge. Such data may be of high quality depending on the technologies used for data collection. Finally, person D may be sharing personal photos or social media posts with vernacular place references or unconscious geotags; low in all dimensions, but still creating potentially useful GI.

Further examples of the GESD are provided in Figure 3 which compares two GI production tasks of different levels of thematic-specificity. Note that filled planes in Figures 3 and 4 refer to inclusion zones for producers of high quality GI, and dotted planes are for visual aid only. In landmark mapping, knowledge of city landmark locations is fairly ubiquitous, opening up the domain to participation by people with little training in geography. As well, the task of identifying features on a web-map or marking locations with a mobile app can be done fairly easily by individuals with little-to-no knowledge of GI production. As such, the plane of potential participants in terms of locale-familiarity/place-type and geographic-information-knowledge is large. Alternatively, a GI task higher in thematic-specificity such as developing a map of health risks due to air pollution would be possible for a much smaller subset of participants – those with higher geographic-information-knowledge and varying degrees of locale-familiarity/place-type. In this case, geographic-information-knowledge might be offset by their degree of local knowledge (e.g., experts in point-source interpolation may require less local knowledge). The plane of potential producers of high quality GI for this task is much smaller, and those falling outside of this plane might be more likely to produce erroneous and/or lower quality information products.

Thus far we have considered the GESD from the perspective of creators of GI. An alternate perspective is to frame minimal expertise required to contribute to existent GI sources within

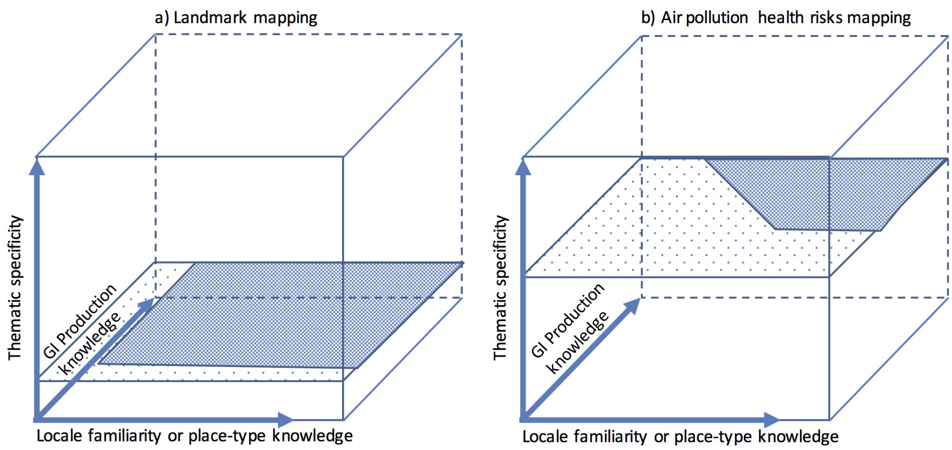


Figure 3 Geographical information production tasks in the geographical expertise diagram.

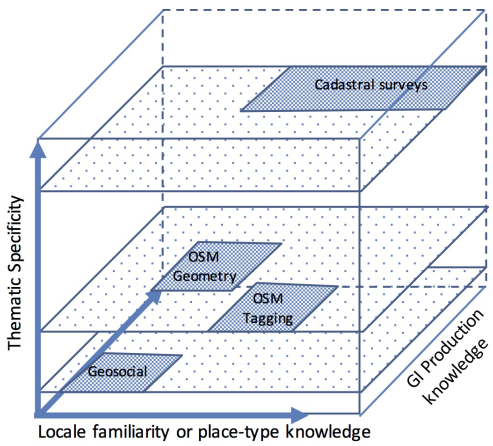


Figure 4 Geographical information products in the geographical expertise diagram.

the GESD, as is outlined in Figure 4. For cadastral surveys, a very high thematic-specificity application, high geographic-information-knowledge in the form of extensive training and professional experience, and a moderate level of locale-familiarity/place-type is required. VGI contributions to OSM constitute a more moderate thematic-specificity application, with different levels of GI and locale-familiarity/place-type required for tagging local features (i.e., attribution) and geometry editing. Geosocial data sources such as geotagged Twitter posts or Flickr photos require very little locale-familiarity/place-type, little geographic-information-knowledge, and are open to almost anyone with Internet access (i.e., low thematic-specificity).

While we have been considering locale-familiarity/place-type as a general category for inherent GE, there is an important scale dimension to all geographic knowledge. In general, geographers tend to develop expertise in one or several spatial scales, and GE as expressed in locale-familiarity/place-type is scale-dependent. We deconstruct spatial scale into two constituent components; grain and extent, and plot their relation to locale-familiarity/place-type in Figure 5. When spatial extent is very large and spatial grain is very small, the plane of potential expertise is limited, as heterogeneity dominates (i.e., one cannot be an expert in all individual areas over a large region). For small spatial extents, GE exists on the plane from

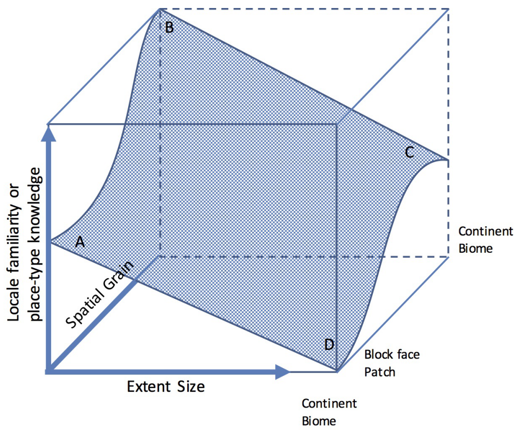


Figure 5 Spatial scale and geographical expertise.

moderate to high locale-familiarity/place-type as spatial grain gets larger (i.e., measurements are aggregated over large areas). For large extents, locale-familiarity/place-type varies from low to moderate as grain size increases.

Example individuals may further illustrate this relation between spatial scale and locale-familiarity/place-type illustrated in Figure 5. Individual A, with GE over a small study area, and a small unit size could represent a researcher studying how small irregularities in sidewalk conditions (e.g. cross-slope, missing and raised sidewalk sections) on a streets hinder mobility of persons using walking aids or wheelchairs. Individual B with GE for a small study area, large unit size might be an expert on history of urban development for a particular street in a major city. Individual C, with GE pertaining to a large study area and large grain size, such as global climate will have a limit on expertise when study area is huge, even if grain size is large, due to spatial heterogeneity. Finally, person D, were GE pertains to a large study area and small grain size, such as neighbourhood socioeconomic status in North American cities or wolf den habitat in the Boreal forest, has severe constraints on GE, as individuals cannot be an expert for all places over huge areas. Note also that the degree of spatial heterogeneity and/or autocorrelation impacts the potential plane for locale-familiarity and place-type.

4 Applied Examples

4.1 Geosocial Data Analysis – Mapping Emotional Affect

One of the ways that GE is implicitly considered in many studies that employ geosocial media data, is computationally distinguishing between ‘locals’ and ‘tourists’ [17]. Often this is done as a filtering step in data processing prior to more in-depth content or spatial analysis. The justification for this filtering stems from a desire to use these sources as a spatially explicit listening post in communities, to sample (albeit from a very unrepresentative sample frame) attitudes or activities in a community of interest. In our model of GE presented here, the filtering step would be estimating inclusion criterion for the geosocial potential plane mapped in Figure 4.

As part of an ongoing research project into geosocial media and urban stress, we investigated peoples’ emotional affect at the time and place of contributing posts on Twitter. The rationale for this study was to validate sentiment analysis metrics for big data using

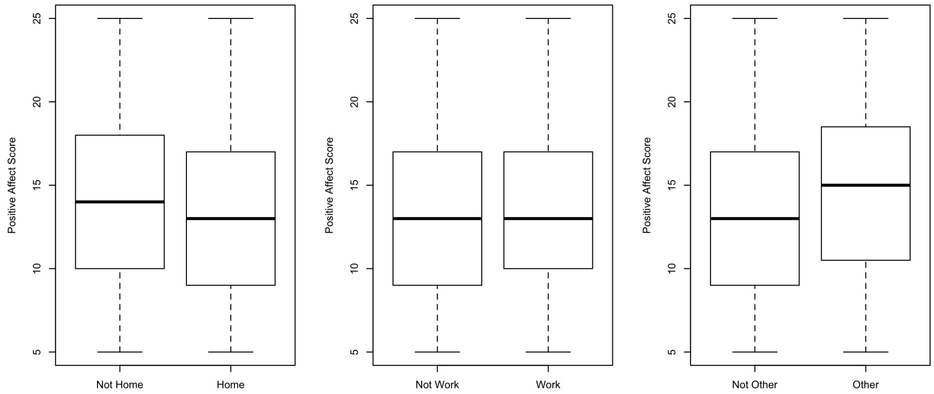


Figure 6 Positive affect scores for Twitter users at the time and place of Tweeting (significant differences in mean positive affect scores for home, not-home and work, not-work based on T tests, $\alpha = .05$).

in-situ psychometrically-valid survey questions [36]. Details of the study design are described elsewhere [33]. A total of 34 Twitter users participated in the study and contributed at least ten posts to their social media accounts during the study period selected here. Each user received short surveys upon entering and exiting the study, as well in response to their social media activity. In [30] we showed that participants were less likely to post messages that related to their immediate surroundings if they were at home or at work. Here we investigate whether participants emotional affect differed for participants based on Tweeting locale. Such information might be of interest for health planners hoping to gauge social media data for analysis of emotional expressions over space and time. From the perspective of GE, individuals' familiarity with the area would likely directly impact the degree to which mood is impacted by the environment.

We show in Figure 6 that positive affect scores differed when participants were Tweeting from home, work, or other locations. In terms of GE, we might be most interested in spatial patterns of emotional affect for residents (with greater experience and locale-familiarity) than with users Tweeting during work or leisure activities. Naïve mining of geosocial data streams may ignore these personal place-based variations in emotional affect.

4.2 Environmental Citizen Science – GrassLander Project

Data quality is a persistent issue in environmental citizen science research and practice, as assumptions about the expertise inherent in the categories of 'citizen' and 'scientist' dominate perceptions of information value. As Cinnamon [7] illustrates, such dichotomies are far from the norm in most VGI projects, where participants in projects cover a range of skill levels, experience, and training. We initiated a citizen science project eliciting input from farmers about agriculture activities and bird observations seen on their farms. This web-based application (www.grasslander.org) required farmers to go through a farm-set up phase in which they selected land parcel polygons for their farm boundaries and digitized their farm fields using web-based spatial data editing tools. We assume that farmers, whose livelihoods depend on their land, have high locale-familiarity/place-type, and as such can adequately provide geospatial data on their properties, regardless of their

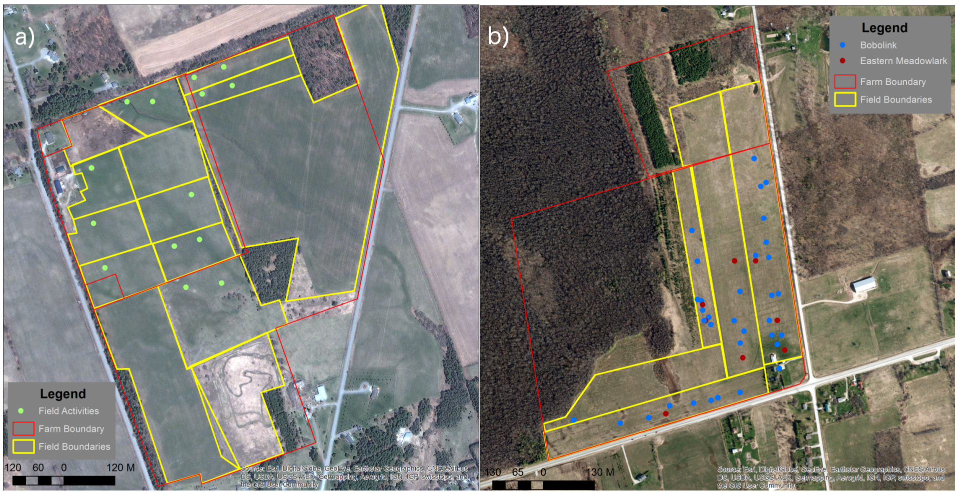


Figure 7 Geographical information contributed by 'citizen scientists' on farm geometry, a) field activities, and b) bird observations in southwest Ontario, Canada).

geographic-information-knowledge if provided with simple tools for creating geographic data. As well, farmers were asked to report on observations of grassland birds (bobolinks and eastern meadowlarks) sighted on their properties.

Geographical information produced by two participants are provided in Figure 7. In both examples, field geometry was found to coincide well with visible fields from aerial imagery. Digitized geometry evident from the participant in Figure 7a demonstrates high attention to spatial details, excluding between-field areas and careful digitizing around their home, however a section of field extends beyond the farm boundary. The participant data in Figure 7b shows field geometry that would not be known from existing aerial imagery. In the context of our GE model, the provision of intuitive web-based digitizing tools lowered the required geographic-information-knowledge needed for producing high quality data in this application. As well, the participant contributed many bird sighting records throughout their property. In future work, we aim to evaluate the degree of locale-familiarity vs. place-type expertise by having participants comment on and characterize farms other than their own, having individuals serve as 'experts' on their own local properties. Such a system can then be incorporated into statistical measures of GE using the imitation game methodology commonly deployed in SEE [10].

5 Conclusions

The model for GE presented here provides three dimensions of GE based on the ESD model of Collins [9]. We demonstrate how concepts of geographic knowledge can be embedded within the ESD model and concepts from SEE can be adapted to geographical information. Our framework provides the tools to deconstruct both data contributors and geographical information products along three dimensions of GE. Such a deconstruction may serve as a basis for developing information quality metrics robust to heterogeneity in the contributions common, but not limited to, many forms of volunteered and/or 'neo' geographic information.

Acknowledgements. The authors would like to thank the participants that contributed to both case studies presented in this research.

References

1 M. Bishr and W. Kuhn. Trust and reputation models for quality assessment of human sensor observations. In *International Conference on Spatial Information Theory*, pages 53–73. Springer, 2013.

2 Rick Bonney, Jennifer L. Shirk, Tina B. Phillips, Andrea Wiggins, Heidi L. Ballard, Abraham J. Miller-Rushing, and Julia K. Parrish. Next Steps for Citizen Science. *Science*, 343(6178):1436–1437, March 2014. doi:10.1126/science.1251554.

3 Z. Cheng, J. Caverlee, H. Barthwal, and V. Bachani. Finding Local Experts on Twitter. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, pages 241–242, New York, NY, USA, 2014. ACM.

4 Z. Cheng, J. Caverlee, H. Barthwal, and V. Bachani. Who is the Barbecue King of Texas?: A Geo-spatial Approach to Finding Local Experts on Twitter. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '14, pages 335–344, New York, NY, USA, 2014. ACM.

5 N. Chrisman. Part 2: Issues and Problems Relating to Cartographic Data Use, Exchange and Transfer: The Role Of Quality Information In The Long-Term Functioning Of A Geographic Information System. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 21(2-3):79–88, July 1984.

6 N. Chrisman. Development in the treatment of spatial data quality. *Fundamentals of Spatial Data Quality*, pages 21–30, 2006.

7 J. Cinnamon. Deconstructing the binaries of spatial data production: Towards hybridity. *The Canadian Geographer / Le Géographe canadien*, 59(1):35–51, March 2015.

8 D. Coleman, Y. Gerogiadou, and J. Labonte. Volunteered Geographic Information: the nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research*, 4:332–358, 2009.

9 H. Collins. Three dimensions of expertise. *Phenomenology and the Cognitive Sciences*, 12(2):253–273, April 2011.

10 H. Collins, R. Evans, M. Weinel, J. Lyttleton-Smith, A. Bartlett, and M. Hall. The Imitation Game and the Nature of Mixed Methods. *Journal of Mixed Methods Research*, page 1558689815619824, December 2015.

11 H. M. Collins and R. Evans. The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science*, 32(2):235–296, April 2002.

12 R. Devillers and R. Jeansoulin, editors. *Fundamentals of spatial data quality*. ISTE, 2006.

13 H. L. Dreyfus and S. E. Dreyfus. From Socrates to Expert Systems: The Limits of Calculative Rationality. In C. Mitcham and A. Huning, editors, *Philosophy and Technology II*, number 90 in Boston Studies in the Philosophy of Science, pages 111–130. Springer Netherlands, 1986.

14 M. J. Egenhofer and D. M. Mark. Naive Geography. In *Spatial Information Theory A Theoretical Basis for GIS*, pages 1–15. Springer, Berlin, Heidelberg, September 1995.

15 Hongchao Fan, Alexander Zipf, Qing Fu, and Pascal Neis. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4):700–719, 2014.

16 S. Gao, K. Janowicz, and G. McKenzie. Towards Platial Joins and Buffers in Place-Based GIS. In *Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place*, COMP '13, pages 42:42–42:49, New York, NY, USA, 2013. ACM.

17 F. Girardin, F. Calabrese, F. D. Fiore, C. Ratti, and J. Blat. Digital Footprinting: Uncovering Tourists with User-Generated Content. *IEEE Pervasive Computing*, 7(4):36–43, October 2008.

18 R. G. Golledge. The Nature of Geographic Knowledge. *Annals of the Association of American Geographers*, 92(1):1–14, March 2002.

19 M. Goodchild. NeoGeography and the Nature of Geographic Expertise. *J. Locat. Based Serv.*, 3(2):82–96, June 2009.

20 M. F. Goodchild and L. Li. Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120, 2012.

21 M. Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4):682–703, 2010.

22 M. Haklay. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In Daniel Sui, Sarah Elwood, and Michael Goodchild, editors, *Crowdsourcing Geographic Knowledge*, pages 105–122. Springer Netherlands, January 2013.

23 Carsten Keßler and René Theodore Anton De Groot. Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. In *Geographic information science at the heart of Europe*, pages 21–37. Springer, 2013.

24 M. Lewicka. Place attachment: How far have we come in the last 40 years? *Journal of Environmental Psychology*, 31(3):207–230, September 2011.

25 S. Maderson and S. Wynne-Jones. Beekeepers' knowledges and participation in pollinator conservation policy. *Journal of Rural Studies*, 45:88–98, June 2016.

26 S. Maguire and M. Tomko. Ripe for the picking? Dataset maturity assessment based on temporal dynamics of feature definitions. *International Journal of Geographical Information Science*, 0(0):1–25, February 2017.

27 I. Poças, J. Gonçalves, B. Marcos, J. Alonso, P. Castro, and J. P. Honrado. Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring. *International Journal of Geographical Information Science*, 28(11):2356–2371, 2014.

28 J. Preece. Citizen Science: New Research Challenges for Human–Computer Interaction. *International Journal of Human-Computer Interaction*, 32(8):585–612, 2016.

29 T. Quesnot and S. Roche. Platial or Locational Data? Toward the Characterization of Social Location Sharing. In *2015 48th Hawaii International Conference on System Sciences*, pages 1973–1982, January 2015.

30 C. Robertson, R. Feick, M. Sykora, K. Shankardass, and K. Shaughnessy. Personal Activity Centres and Geosocial Data Analysis: Combining big data with small data. In *AGILE 2017*, Wageningen, Netherlands, May 2017. Springer-Verlag.

31 R. A. Rundstrom. GIS, Indigenous Peoples, and Epistemological Diversity. *Cartography and Geographic Information Systems*, 22(1):45–57, January 1995. `doi:10.1559/152304095782540564`.

32 H. Senaratne, A. Mobasheri, A. L. Ali, C. Capineri, and M. Haklay. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1):139–167, 2017.

33 M. D. Sykora, C. Robertson, K. Shankardass, R. Feick, K. Shaughnessy, B. Coates, H. Lawrence, and T. Jackson. Stresscapes: validating linkages between place and stress expression on social media. In *Proceedings of the 2nd International Workshop on Mining Urban Data*, pages 80–84, Lille, France, 2015.

34 Y.-F. Tuan. Place: An Experiential Perspective. *Geographical Review*, 65(2):151–165, April 1975. ArticleType: research-article / Full publication date: Apr., 1975 / Copyright © 1975 American Geographical Society.

35 M. Van Exel, E. Dias, and S. Fruijtier. The impact of crowdsourcing on spatial data quality indicators. *Proceedings of GiScience 2011*, 2010.

36 D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070, 1988.

37 B. Wynne. Sheepfarming after Chernobyl: A Case Study in Communicating Scientific Information. *Environment: Science and Policy for Sustainable Development*, 31(2):10–39, March 1989.