14th International Conference on Spatial Information Theory

COSIT 2019, September 9–13, 2019, Regensburg, Germany

Edited by

Sabine Timpf
Christoph Schlieder
Markus Kattenbeck
Bernd Ludwig
Kathleen Stewart

Editors

Sabine Timpf (ID)
University of Augsburg, Germany
sabine.timpf@geo.uni-augsburg.de

Christoph Schlieder (ID)
University of Bamberg, Germany
christoph.schlieder@uni-bamberg.de

Markus Kattenbeck (ID)
TU Wien, Austria
markus.kattenbeck@geo.tuwien.ac.at

Bernd Ludwig (ID)
University of Regensburg, Germany
bernd.ludwig@ur.de

Kathleen Stewart (ID)
University of Maryland, U.S.A.
stewartk@umd.edu

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.

To our colleagues who fostered and shaped this community.

Contents

Invited Talk

Modeling and Interaction

Reasoning and Reproducibility

Spatial language

Invited Talk

Ontological Modeling

Diagrams and visualization

Invited Talk

Invited Talk

Cognitive Models of Wayfinding

Spatial knowledge and learning

Preface

With this COSIT conference being held in Regensburg (Germany) in September, 2019, this conference series on spatial information theory has lasted for more than a quarter of a century. For a community that is made up of researchers from many different disciplines, this is a remarkable achievement - one, which will be given due consideration during the conference week of September 09–13, 2019. With the selection of Regensburg as the conference location, the series also returns to Germany 20 years after the last COSIT held in a German location, that of Stade, near Hamburg.

One of the hallmarks of COSIT (see the conference series website www.cosit.org) is the desire to hold the conference in a remote but still accessible location – accessibility being determined as at most 2 hours travel time from an international airport – in order to promote the building of a cohesive community. Towards this end, we have also been restricting COSIT to be a single track conference with (personal) presentations, trying to create a scientific event where participants share the same experiences (scientifically and socially) and are able to discuss these experiences in person. This is, we believe, one of the prerequisites for building such a tightly-knit community. Continuing in this vein, we eschew the use of Skype presentations or similar virtualization efforts, focusing on personal contact.

Putting together a single-track conference means that, unfortunately, there are fewer opportunities for presentation and a reduced number of paper acceptances. For this newest installment of the series, we received 30 full paper submissions, 48 short paper submissions and 7 vision paper submissions. Vision papers are a new category in which not only evidence-based ideas may be presented but also a broader view of COSIT topics as a whole may be expounded.

It is a remarkable fact about the early conferences in the COSIT series that they were successful at producing numerous seminal papers, which spelled out a research program or provided a synthesis of findings from an interdisciplinary perspective. To revisit visions that were suggested in the past and to stimulate the discussion of new research directions, we decided to solicit vision papers as a new type of submission for COSIT 2019. These are rigorously argued papers that identify emerging problems or questions for the geospatial science community as a whole to address. Vision papers were coordinated by Maria Vasardani.

All COSIT 2019 papers were reviewed by at least three members of the program committee and the program chairs selected 8 full (27% acceptance rate), 12 short (25% acceptance rate) and 4 vision (57% acceptance rate) papers for publication and oral presentation at the conference. Those paper authors that could not get a spot in the tight program were given an invitation to present a poster during a specially arranged poster session.

The 23 accepted papers show both breadth and diversity with respect to topics, chief among these *modelling and interaction* (4 papers), *spatial reasoning* (2), *reproducibility* (1), *spatial language* (4), *ontological modeling* (3), *diagrams* (2), *visualization* (2), *cognitive models of wayfinding* (3), *spatial knowledge* (2), and *learning* (1).

Looking back on more than 25 years of COSIT we can observe the re-occurrence of several by now established research areas such as spatial reasoning, ontologies, and wayfinding. More recent topics such as urban networks and reproducible research in geoinformatics are also represented in the program. The different disciplines involved at COSIT have learned from each other, which is an achievement in itself. As a result we see papers integrating

methodological approaches from disciplines present at COSIT, e.g., a geography and computer science or psychology and artificial intelligence. This year we observe an increased interest in spatial language as evidenced by two sessions as well as a workshop on the topic.

An important part of the COSIT conferences are the accompanying satellite events such as the Doctoral Colloquium (this year organized by Ioannis Giannopoulos and Hedda R. Schmidtke) and four workshops and a tutorial (coordinated by Toru Ishikawa and Johannes Scholz). These events serve as additional presentation platforms, and they also showcase the currently active interest of the research community. The topics of this year's workshops and tutorial range from Spatial Cognition and AI, quality aspects in the context of localization, computing techniques for spatio-temporal data in archaeology and cultural heritage, communicating about space, and the goodness of space (related to Alexander's notion of wholeness). In addition to these events focused on content, COSIT provides further opportunities for social interaction with the poster session and reception, the afternoon excursion and the conference dinner.

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 and timely reviews they provided.

We would like to thank the University of Regensburg for supporting the conference by offering the location as well as technical and logistic support free of charge. We gratefully acknowledge the support of our sponsors, the Universitätsstiftung Hans Vielberth (Regensburg, Germany) which generously covered the travel expenses of the COSIT 2019 keynote speakers, as well as the financial support by Krones AG (Neutraubling, Germany) and Number42 (Regensburg, Germany).

Finally, we would like to thank all who will attend COSIT 2019 to present their work, to discuss the work showcased at the conference, and to advance the state of the art in the field of spatial information theory, thus keeping this community vibrantly alive.

<div align="right">

July 2019
Sabine Timpf, Christoph Schlieder,
Markus Kattenbeck, Bernd Ludwig, and Kathleen Stewart

</div>

COSIT Steering Committee

Christophe Claramunt
Naval Academy Research Institute, France

Eliseo Clementini
University of L'Aquila, Italy

Anthony Cohn
University of Leeds, UK

Claire Davies
University of Winchester, UK

Matt Duckham
RMIT University, Australia

Max Egenhofer
University of Maine, U.S.A.

Sara Fabrikant
University of Zurich, Switzerland

Paolo Fogliaroni
TU Wien, Austria

Andrew Frank
TU Wien, Austria

Christian Freksa
University of Bremen, Germany

Nicholas Giudice
University of Maine, U.S.A.

Stephen Hirtle
University of Pittsburgh (Chair), U.S.A.

Werner Kuhn
UCSB, U.S.A.

David Mark
SUNY Buffalo, U.S.A.

Dan Montello
UCSB, U.S.A.

Reinhardt Moratz
University of Münster, Germany

John Stell
University of Leeds, UK

Kathleen Stewart
University of Maryland, U.S.A.

Thora Tenbrink
Bangor University, UK

Sabine Timpf
University of Augsburg, Germany

Stephan Winter
University of Melbourne, Australia

Zena Wood
University of Greenwich, UK

Michael Worboys
University of Greenwich, UK

Program Committee

Benjamin Adams
University of Canterbury, New Zealand

Jochen Albrecht
City University of New York, U.S.A.

Thomas Barkowsky
University of Bremen, Germany

Joerg Baus
University of Saarbrücken, Germany

Alberto Belussi
University of Verona, Italy

Itzhak Benenson
Tel Aviv University, Israel

Sven Bertel
Flensburg University of Applied Sciences,
Germany

Michela Bertolotto
University College Dublin, Ireland

Mehul Bhatt
Örebro University, Sweden

Thomas Bittner
SUNY Buffalo, U.S.A.

Stefano Borgo
ISTC, Italy

Alan Both
RMIT University, Australia

Boyan Brodaric
Geological Survey of Canada

David Brosset
Naval Academy Research Institute, France

Tad Brunye
Tufts University, U.S.A.

Guoray Cai
The Pennsylvania State University, U.S.A.

Malumbo Chipofya
University of Münster, Germany

Christophe Claramunt
Naval Academy Research Institute, France

Eliseo Clementini
University of L'Aquila, Italy

Richard Dapoigny
LISTIC, France

Auriol Degbelo
University of Osnabrück, Germany

Géraldine Del Mondo
LITIS, France

Paolino Di Felice
University of L'Aquila, Italy

Heshan Du
University of Nottingham, U.K.

Matt Duckham
RMIT University, Australia

Max Egenhofer
University of Maine, U.S.A.

Sara Fabrikant
University of Zurich, Switzerland

Zoe Falomir
University of Bremen, Germany

Paolo Fogliaroni
TU Wien, Austria

Fred Fonseca
The Pennsylvania State University, U.S.A.

Christian Freksa
University of Bremen, Germany

Scott Freundschuh
University of New Mexico, U.S.A.

Antony Galton
University of Exeter, U.S.A.

Ioannis Giannopoulos
TU Wien, Austria

Michael Goodchild
UCSB, U.S.A.

Björn Gottfried
University of Bremen, Germany

Amy Griffin
RMIT University, Australia

Torsten Hahmann
University of Maine, U.S.A.

Francis Harvey
Leibniz Institute for Regional Geography

Stephen Hirtle
University of Pittsburgh, U.S.A.

Hartwig Hochmair
University of Florida, U.S.A.

Barbara Hofer
University of Salzburg, Austria

Konstantin Hopf
University of Bamberg, Germany

Yingjie Hu
SUNY Buffalo, U.S.A.

Toru Ishikawa
The University of Tokyo, Japan

Bin Jiang
University of Gävle, Sweden

David Jonietz
HERE Technologies, Switzerland

Tomi Kauppinen
Aalto University School of Science, Finland

Marinos Kavouras
National Technical University of Athens,
Greece

Carsten Kessler
Aalborg University Copenhagen, Denmark

Peter Kiefer
ETH Zurich, Switzerland

Alexander Klippel
The Pennsylvania State University, U.S.A.

Christian Kray
University of Münster, Germany

Jakub Krukar
University of Münster, Germany

Werner Kuhn
UCSB, U.S.A.

Longin Latecki
Temple University, U.S.A.

Sang Ah Lee
University of Trento, Italy

Ickjai Lee
James Cook University, Australia

Didier Leibovici
The University of Sheffield, UK

Sanjiang Li
University of Technology Sydney, Australia

Alan Maceachren
The Pennsylvania State University, U.S.A.

Claudio Masolo
ISTC-CNR, Trento, Italy

Grant McKenzie
McGill University, Canada

Sara Migliorini
Università degli Studi di Verona, Italy

Franz-Benjamin Mocnik
Heidelberg University, Germany

Dan Montello
UCSB, U.S.A.

Reinhard Moratz
University of Münster, Germany

Till Mossakowski
University of Magdeburg, Germany

Bernard Moulin
Laval University, Canada

Stephan Mäs
TU Dresden, Germany

Ralf Möller
University of Lübeck, Germany

Gerhard Navratil
TU Wien, Austria

Marta Olivetti Belardinelli
Sapienza University of Rome, Italy

Ana-Maria Olteteanu
Freie Universität Berlin, Germany

Dimitris Papadias
The Hong Kong University of Science and Technology

Avi Parush
Technion, Israel

Edzer Pebesma
University of Münster, Germany

David Peebles
University of Huddersfield, UK

Brandon Plewe
Brigham Young University, U.S.A.

Ross Purves
University of Zurich, Switzerland

Martin Raubal
ETH Zurich, Switzerland

Tumasch Reichenbacher
University of Zurich, Switzerland

Kai-Florian Richter
Umea University, Sweden

Simon Scheider
University Utrecht, The Netherlands

Hedda Schmidtke
University of Oregon, U.S.A.

Angela Schwering
University of Münster, Germany

Linda See
International Institute for Applied Systems Analysis, Austria

Takeshi Shirabe
KTH Royal Institute of Technology, Sweden

Gaurav Sinha
Ohio University, U.S.A.

Spiros Skiadopoulos
University of Peloponnese, Greece

Barry Smith
SUNY Buffalo, U.S.A.

Emmanuel Stefanakis
University of Calgary, Canada

Klaus Stein
University of Bamberg, Germany

John Stell
University of Leeds, UK

Kristin Stock
Massey University, New Zealand

Kazuko Takahashi
Kwansei Gakuin University, Japan

Thora Tenbrink
Bangor University, UK

Martin Tomko
University of Zurich, Switzerland

Nico Van de Weghe
Ghent University, Belgium

Antonio Vieira Monteiro
INPE, Brazil

Jan Oliver Wallgrün
The Pennsylvania State University, U.S.A.

Stephan Winter
The University of Melbourne, Australia

Diedrich Wolter
University of Bamberg, Germany

Wai Yeap
Auckland University of Technology, New Zealand

May Yuan
The University of Texas at Dallas, U.S.A.

Sponsors

University of Regensburg
Regensburg, Germany

Universitätsstiftung Hans Vielberth
Regensburg, Germany

University of Augsburg
Augsburg, Germany

Krones AG
Neutraubling, Germany

Number42
Regensburg, Germany

Human Vision at a Glance

Ruth Rosenholtz[1]

Massachusetts Institute of Technology, Department of Brain & Cognitive Sciences, CSAIL, USA
http://persci.mit.edu/people/rosenholtz
rruth@mit.edu

Dian Yu

Massachusetts Institute of Technology, Department of Brain & Cognitive Sciences, CSAIL, USA
dyumit17@gmail.com

── Abstract ──

Recent advances in human vision research have pointed toward a theory that unifies many aspects of vision relevant to information visualization. According to this theory, loss of information in peripheral vision determines performance on many visual tasks. This theory subsumes old concepts such as visual saliency, selective attention, and change blindness. It predicts the rich details we have access to at a glance. Furthermore, it provides insight into tasks not commonly studied in human vision, such as ability to comprehend connections in a network diagram, or to compare information in one part of a display with that in another.

1 Relevance of human vision to design

Designs, whether of information visualizations or of user interfaces, can be good or bad for both cognitive and perceptual reasons. At the bare minimum, designers strive to make relevant information easily perceptible. If a legend is not easily discriminable from the rest of the display, a user will not easily notice it. If icons with distinct meaning do not also have a sufficiently distinct appearance, a user will need to search for the icons of interest, an often slow process, and they will have trouble getting the "gist" of the display, e.g. the spatial distribution of the two types of icons. In addition, even if information is readily available to perception, a confusing design can make tasks cognitively difficult, for example when icons are clearly identifiable but have unclear meaning. Here we focus on some important perceptual issues:

1. How easy will it be for users to find information in a display?
2. Will users notice important and unexpected information, such as an alert?
3. In dynamic displays, will users notice and make sense of information that changes?
4. How difficult is it to connect related information in multi-view displays?
5. What information can the user get at a glance at a display, i.e. what "gist" can they easily extract?

---

[1] Corresponding author

6. What distinguishes details that users can perceive at a glance from those they are more likely to miss?
7. Many visualizations like maps and graphs contain complex connections between nodes. Can the user easily make sense of these connections?

Classic vision science has studied these issues as essentially six separate problems: search, and saliency of "alerts" (items 1, 2); change or difference detection (3, 4); scene gist (5, 6); gist of an ensemble of items (5, 6); what details require "attention" to perceive veridically (3, 6); and visual cognition problems such as graph and maze connectedness (7). While vision science has elucidated a number of important phenomena, the results have often been unsatisfying in terms of translation to applications such as information visualization.

First, in the interest of pinpointing visual mechanisms, vision scientists have quite reasonably studied simple controlled displays. For instance, experiments have asked observers to locate a red item amongst a large number of homogeneous green items. Such experiments aimed to determine what feature differences (color, size, etc.) lead unusual items to "pop out" and "draw the observer's attention"; a question related to whether a user will notice an alert. These experiments have been interesting for probing vision, but such homogeneous displays have limited applicability to more complex real-world designs.

Second, much of the relevant vision science has largely remained at the level of descriptive enumeration of behavioral results, rather than transitioning to predictive models. For instance, suppose you have two alternative designs for a subway map, and want to know which makes it easier for observers to find a route between two stations. Vision science could provide some good rules of thumb such as "high contrast lines are better". But it cannot provide a direct answer to the original question: which design is better, and by how much? This reliance on descriptive rather than predictive models is particularly problematic since vision science has produced few behavioral results related to perception of complex applied displays.

Third, to the extent that vision science has unified these 7 key perceptual issues for designs, it has done so by relating them all to visual attention: many details are difficult to rapidly perceive because they require focused attention [18]; gist is easy because it does not require focused attention [17, 19]; change detection is difficult because it does require attention to the change [5, 9, 13]; tracing a path through a graph is difficult because attention needs time to spread along each route [6]. The problem with these theories, even if correct, is that attention is nearly impossible to measure using behavioral methods alone. Attempts to fix this problem by using eye movements as a proxy for attention have been modestly successful only in very limited situations [16].

New developments in vision science, over the last decade, have instead suggested that many perceptual issues of relevance to design largely depend upon a single factor: the information lost and maintained in peripheral vision (see [10] for a review). Human performance of visual tasks relies greatly on peripheral vision, the region outside the rod-free fovea, comprising more than 99% of the visual field. We mainly use our foveas for work that requires precision and ultra-fine details, such as reading a paragraph of text or making careful measurements. Peripheral vision, on the other hand, is critical for efficient processing of large portions of a display or scene, e.g. to get the gist of a scene, to notice an alert or find a target, or to compare two conditions in a plot or two views of the same data. In fact, one might often want to design for maximum information gathering with a minimal number of eye movements, making use of efficient peripheral processing to understand a display at a glance (or anyhow, a small number of glances).

Understanding of peripheral vision in fact appears to subsume understanding of many important visual phenomena, from saliency and search, through change detection and gist perception. Some items are salient because they can easily be seen in the periphery, due to

distinct visual features or a lack of clutter [1, 8, 12, 14, 20]. The gist, or information readily available at a glance, derives from the information that survives peripheral vision. The details a user might miss are those less readily available in the periphery, sometimes leading to a failure of the user to detect changes to a dynamic display [15]. Peripheral vision can also inform us about the usability of network graphs and maps – topics of relevance to a number of applications, but little studied in vision science.

Human vision depends upon peripheral vision in spite of the fact that peripheral vision loses a great deal of information. The big loss of information has to do with phenomena known as "crowding": the degradation in visual performance in the presence of clutter. These losses are complex, stimulus-specific, and it is hard to get good intuitions about what information is lost and what is preserved.

Our lab has spent the last decade studying the effect of peripheral vision on many visual tasks. We have developed and extensively tested the state-of-the-art model of peripheral vision [2–4, 7, 11, 12, 20]. We call this the Texture Tiling Model because peripheral vision appears to compress and summarize large regions of the visual world by treating them as texture. This model makes predictions about what information is preserved and lost in peripheral vision, and does so in an intuitive way; it outputs visualizations of the information available to peripheral vision. To date, we have demonstrated that this model does well at predicting performance on over 70 visual tasks.

From the point of view of understanding one's users, it is good news that peripheral vision has proven a more powerful explanation than attention. Whereas attention is difficult to measure it is relatively easy to know where a user is pointing their eyes and therefore what portion of the visual field lands in the periphery. On the other hand, the importance of peripheral vision means that we must interpret eye-tracking results with care; users will often be processing visual input across large portions of the visual field, not just where they point their eyes.

To get intuitions about what tasks users can successfully complete with peripheral vision, one can use one's own visual system. Rather than the old advice to "squint" at a design, point your eyes at locations the user is likely to fixate (for example, on a button they must click), and introspect on what information is clear or confusing in the periphery. Alternatively, one can inspect the visualizations output by our Texture Tiling Model – with no need to fixate – to get intuitions about what information the model predicts is and is not preserved in peripheral vision. One need not even entirely trust the model, as again one can verify any predictions by using one's own visual system to introspect on the information available.

References

1 J. J. Andriessen and H. Bouma. Eccentric vision: Adverse interactions between line segments. *Vision Research*, 16(1):71–78, 1976.

2 B. J. Balas, L. Nakano, and R. Rosenholtz. A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision*, 9(12):13, 2009.

3 H. Chang and R. Rosenholtz. Search performance is better predicted by tileability than by the presence of a unique basic feature. *Journal of Vision*, 16(10):13, 2016.

4 K. A. Ehinger and R. Rosenholtz. A general account of peripheral encoding also predicts scene perception performance. *Journal of Vision*, 16(2):13, 2016.

5 A. Hollingworth and J. M. Henderson. Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, 28(1):113–136, 2002.

6 R. Houtkamp, H. Spekreijse, and P.R. Roelfsema. A gradual spread of attention. *Perception & Psychophysics*, 65(7):1136–1144, 2003. `doi:10.3758/BF03194840`.

7   S. Keshvari and R. Rosenholtz. Pooling of continuous feature provides a unifying account of crowding. *Journal of Vision*, 16(3):39, 2016.

8   F. L. Kooi, A. Toet, S. P. Tripathy, and D. M. Levi. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, 8(2):255–279, 1994.

9   R. A. Rensink, J. K. O'Regan, and J. J. Clark. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8:368–373, 1997.

10   R. Rosenholtz. Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, 2(1):437–457, 2016. `doi:10.1146/annurev-vision-082114-035733`.

11   R. Rosenholtz, J. Huang, and K. A. Ehinger. Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology*, 3:13, 2012. `doi:10.3389/fpsyg.2012.00013`.

12   R. Rosenholtz, J. Huang, A. Raj, B. Balas, and L. Ilie. A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, 12(14):1–17, 2012. `doi:10.1167/12.4.14`.

13   B. J. Scholl. Attenuated change blindness for exogenously attended items in a flicker paradigm. *Visual Cognition*, 7(1-3):377–396, 2000. `doi:10.1080/135062800394856`.

14   M. Scolari, A. Kohnen, B. Barton, and E. Awh. Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision*, 7(2):7, 2007.

15   M. E. Smith, L. Sharan, E. Park, L. C. Loschky, and R. Rosenholtz. Difficulty detecting changes in complex scenes depends in part upon the strengths and limitations of peripheral vision. *Journal of Vision*, under revision.

16   B. W. Tatler, M. M. Hayhoe, M. F. Land, and D. H. Ballard. Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, 11(5):5, 2011. `doi:10.1167/11.5.5`.

17   A. Treisman. How the deployment of attention determines what we see. *Visual Cognition*, 14:411–443, 2006.

18   A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, 1980.

19   J. M. Wolfe, M. L.-H. Vo, K. K. Evans, and M. R. Greene. Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Sciences*, 15(2):77–84, 2011.

20   X. Zhang, J. Huang, S. Yigit-Elliot, and R. Rosenholtz. Cube search, revisited. *Journal of Vision*, 15(3):9, 2015.

Smartphone Usability for Emergency Evacuation Applications

David Amores 🄳
University of Melbourne, Melbourne 3010, Australia
damores@student.unimelb.edu.au

Maria Vasardani 🄳
RMIT University, Melbourne 3000, Australia
mvasardani@rmit.edu.au

Egemen Tanin 🄳
University of Melbourne, Melbourne 3010, Australia
etanin@unimelb.edu.au

—— **Abstract** ——

Mobile phone ubiquity has allowed the implementation of a number of emergency-related evacuation aids. Yet, these applications still face a number of challenges in human-mobile interaction, namely: (1) lack of widely accepted mobile usability guidelines, (2) people's limited cognitive capacity when using mobile phones under stress, and (3) difficulty recreating emergency scenarios as experiments for usability testing. This study is intended as an initial view into smartphone usability under emergency evacuations by compiling a list of experimental observations and setting the ground for future research in cognitively-informed spatial algorithms and app design.

1 Introduction

Wide adoption of smartphones has motivated creating technological aids for emergencies. There is extensive research in disaster management, specifically for indoor emergency evacuations. Still, emergency apps have yet to take off and enjoy broad user acceptance. This setback has been attributed to the lack of robust and widely deployed technologies such as indoor positioning or building mapping [16]. Although these factors are indeed limiting, we show in this paper that there is room for improvement in areas related to usability, as follows:

- **Use of design guidelines.** Previously proposed emergency applications often do not abide by best practices in mobile application design or navigation services design.
- **Consideration of cognition demands.** Related literature seldom performs usability testing and does not assess an application's cognitive demands, even though novel technologies, such as Augmented Reality (AR), have proven to be distracting [4].
- **Experiment validity (ecological validity).** A variety of experiments are done in literature but rarely do they recreate lifelike emergencies. Given that people may act differently during emergencies [17], experiments should resemble actual scenarios.

Thus, we propose that *beyond indoor technological setup issues, smartphone app usability is a key factor in evacuation app adoption.* This research in-progress constitutes an initial step towards building cognitively-ergonomic applications and spatial algorithms, thus, enabling the creation of advanced evacuation systems. As such, this paper intends to analyse the

aforementioned issues by first collecting previous works that propose emergency applications (Section 2) and highlighting their impediments and impact in the lack of subsequent app adoption. In parallel, an experiment (Section 3) is performed where we develop an emergency application and set up a scenario resembling common characteristics of previous studies to take a closer look at the issues. Finally, a set of observations that consolidates the findings is presented in Section 4 and constitutes the main contribution of this ongoing work.

2   Related work

2.1   Smartphone usability and emergencies

General guidelines for best practices of mobile interface development have been proposed before [14]. In relation to navigation, usability guidelines for digital maps in mobile phones are given in [11] and mental load for different digital map representations is assessed in [7]. Complementary to these studies, our research aims to obtain insights from empirical data strictly related to applications for emergency scenarios. The context in these cases – e.g. the emergency – may play a bigger role than other scenarios, ultimately suggesting that alternative design and interaction paradigms can be used where cognition is a central factor.

Harrison *et al* [8] propose a high-level usability framework tailored for smartphones where cognitive load is added as an attribute besides more established ones such as effectiveness or efficiency. *Cognitive load* – the amount of mental effort needed to perform a task – is pertinent to mobile phones as users commonly engage in other activities in parallel such as driving or evacuating [8]. We argue that cognitive load is not only an additional usability attribute but a central one in emergency scenarios. Consequently, we give experimental observations evidencing the importance of cognitive load in usability testing. We believe evaluating cognitive load is especially relevant during emergencies where the main task is evacuating, and using a mobile phone should be a secondary and supporting task.

When proposing an application for emergencies, a variety of experiments can be done such as computer simulations, controlled lab experiments or drills. Choosing one of these experiments entails a trade-off between experimental control and *ecological validity* – i.e. the degree of correspondence between experimental and real-life settings [10]. People's behaviours in simulated emergencies differ to real-life emergencies [17], where people tend to exhibit non-adaptive behaviours [2]. A general framework for ecological validity in usability is given in [10]. However, there is a lack of more measurable and objective ways of quantifying ecological validity. Our research intends to advance this by providing observations on ecological validity in emergency experiments leading to more specific measures in future research.

2.2   Emergency applications

Table 1 lists representative studies that have previously proposed an application as an emergency evacuation aid. Column *DG* refers to whether the paper roughly mentions using mobile or map design guidelines for the app. Column *CL* indicates whether cognitive load is assessed within app testing. Column *EV* states the amount of ecological validity where ✗, ✓ and ✓ are used to depict the amount of realism, and "–" when there is no user experiment.

It is unclear whether the listed works follow design guidelines or not, mainly because usability design is not the main focus of those studies. Based on app descriptions or screenshots, design guidelines are seen to be roughly followed but with a number of shortcomings. A system in [3] describes an app for delivering good routing instructions. However, temperature and distance are shown with text and no assessment is made as to whether users are able to

Table 1 Studies proposing the implementation of an emergency evacuation application.

Previous work	DG	CL	EV
Smartescape: A mobile smart individual fire evacuation system ... (2018) [3]	✓	X	X
Publish-subscribe smartphone sensing platform for the acute ... (2014) [13]	✓	✓	–
An indoor augmented-reality evacuation system for the ... (2012) [1]	X	X	–
Presenting evacuation instructions on mobile devices by means ... (2008) [5]	✓	X	✓
Handheld augmented reality indoor navigation with activity ... (2011) [12]	✓	X	–
An integrated building fire evacuation system with rfid and ... (2011) [6]	X	X	–
Integrating geographical information and augmented reality ... (2012) [15]	✓	X	✓
Indoor emergency evacuation service on autonomous navigation ... (2008) [9]	✓	X	–

grasp this info while moving. The problem with long texts is present in other applications as well [9, 1]. Good practices offered by underlying maps such as Google Maps are present in some works (e.g. [13]) but not in others where maps are built from scratch [15, 12].

Applications proposed in [1, 15, 12] convey route instructions with AR using the phone's camera, prompting users to hold their phones upright while walking. However, no assessment is done to check if users found this distracting. Indeed, AR has been proven to be diverting and hazardous for mobile users [4]. Other applications propose 3D [3, 5] or 2D maps [9] but cognitive load testing is lacking. Only one work [13] mentions doing it as a future direction. In fact, usability testing is rarely performed in these works. Because they are built for emergencies, we believe these studies must assess how distracting novel features may be.

A subset of earlier studies in emergency management proposes applications but do not experiment with them [6, 9] and others use computer simulations [1]. Some studies perform experiments in controlled settings, providing participants specific tasks [3, 5, 15]. Low ecological validity could lead to unauthentic behaviours, jeopardising usability studies.

3 Experiment setup

The purpose of this preliminary experiment is to take a closer look at the aforementioned issues with smartphone usability during emergencies: design guidelines, cognitive load and ecological validity. The experiment consisted in equipping building occupants with a mobile application that would deliver evacuation information. The application was built to resemble a conventional smartphone app for emergency instruction conveyance, with two main features (Figure 1): **Messaging** – Push notifications, notifications history page – and **Map** – Multi-storey building map, personalised and dynamic exit routes, indoor landmarks, outdoor route to assembly area. A scheduled drill in a university staff building was chosen as the scenario. Seventy four (74) participants registered in our application prior to the drill.

To understand people's perception of the *application's design*, we gave participants a post-drill survey to fill. To assess the role *cognitive load* played, each participant was given access either to map-related features only, to notifications only, or to every feature, and compared phone interactions. We also compared the number and type of interactions (e.g. zooming, tapping) made with the app at different times of the evacuation. During the drill, participants had the liberty to use or disregard the application. To promote the use of the app, an exit was purposefully blocked and messages were sent throughout the drill. *Ecological validity* was assessed by observers during the drill. We collected three types of data:

- **Survey.** Answers to a survey including a general feedback question.
- **Interactions.** Recorded interactions with the app: taps, map zooming/panning.
- **Observations.** Experimental observations by the researchers during the drills.

(a) List of notifications.   **(b)** Indoor exit route.   **(c)** Outdoor route to assembly area.
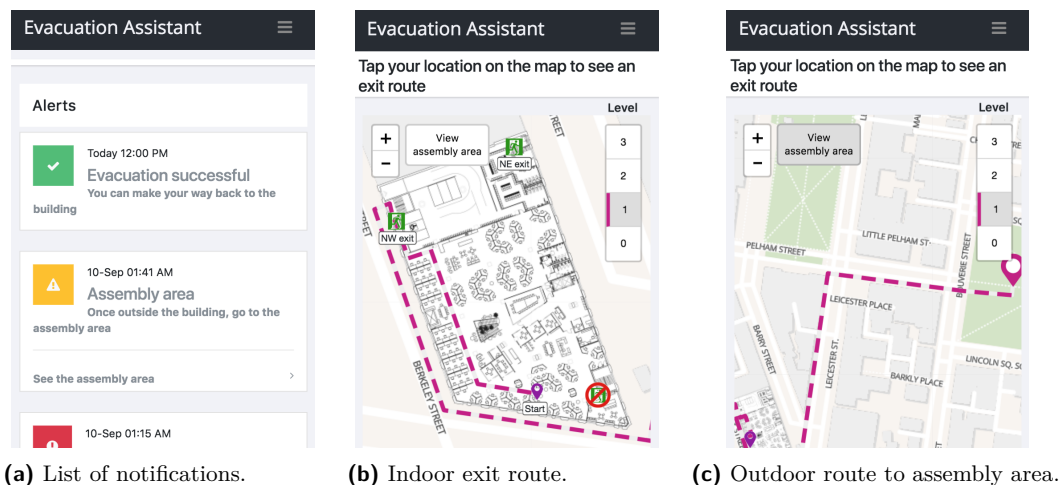
Figure 1 Screenshots of the web mobile application used for the experiment.

4   Result analysis and discussion

The following subsections provide observations from the experiment in regards to the challenges mentioned in the introduction. These observations serve as a starting point to provide more specific and measurable guidelines and propositions for future research. Additionally, given that the main findings relate to cognitive load, we expect that these observations pave the way for visualisation and algorithms centered on human cognition.

4.1   Digital maps and mobile design

Observation 1: Users deem established apps as baselines.   Feedback from the experiment showed that users regard popular applications as reference for newer apps, posing a challenge when building novel technologies (e.g. 3D models, AR). One participant expected to see his position on the map while another wished to have directions at each decision point and to be notified if a wrong way was taken. Both mentioned Google Maps as a reference point suggesting that established applications have placed expectations on users for future technologies. Users are also aware of general mobile design practices. A prevalent comment from participants who did not receive push notifications was the need to have them for quick and easy access to the application contents. This observation encourages developers to view established apps as a baseline when proposing new applications and developing new visualisation, communication, or design paradigms. During evacuation scenarios, adaptation of new methods would rather be difficult, at best.

4.2   Cognitive load and usability

Observation 2: Cognitive load is important and noticed by users.   Participants realised the need to minimise the cognitive load the application takes from their working memory. Some participants show concern in their feedback regarding cognitive load: "you need to (...) mov(e) out of the building and PAYING ATTENTION to your surrounds, not stuck with your head in your device". Similar feedback was present in 6 out of 16 texts. Figure 2 reveals interaction patterns during the evacuation where local peaks are visible. The highest peak happens after the first alarm goes off, hinting that people were able to interact with the device before starting to move. The other peaks coincide with people getting out of the

building and in the assembly area. That is, people were able to interact more with the device when they stopped moving. Thus, cognitively-intensive tasks should ideally happen at the start and end of evacuations and in between movement periods, and should be minimised otherwise. This contradicts AR, 3D or similar technologies that are being proposed.
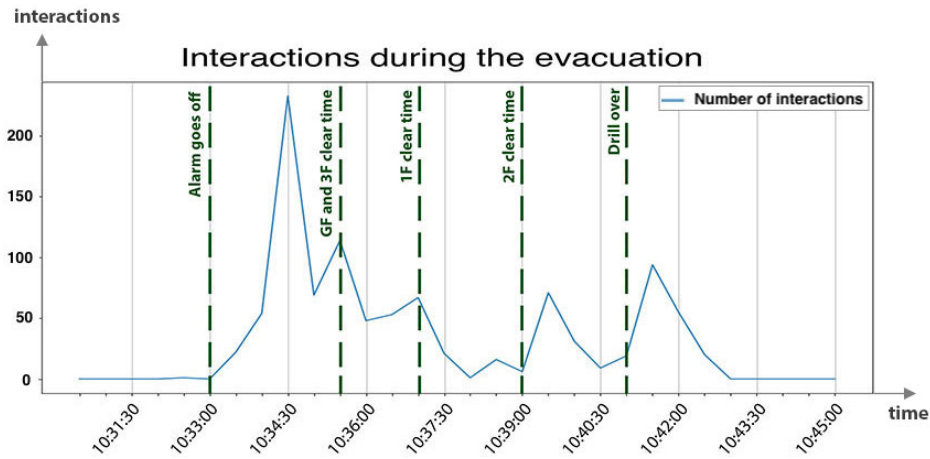
Figure 2 Number of interactions during the complete evacuation period.

Furthermore, Figure 3 shows the responses to a survey question regarding the app's usefulness. The first chart (Figure 3a) shows responses from every participant, revealing "somewhat useful" as the most common answer. Figure 3b shows the responses from people that made use of "shortcut" functions in the app such as push notifications or the assembly area button. These shortcuts allowed users rapid access to functionality, thus, requiring less cognitive load. None of these participants thought the app was not useful at all and, in fact, the proportion of "very useful" responses increases. On the other hand, Figure 3c shows the responses from people who made the most map interactions such as panning or zooming, using much of their cognitive load. The figure shows that no participant thought the app was very useful and the "not at all useful" response reappears. These results suggest that cognitive load due to app interactions impact the overall application's perceived usability.

(a) Total participants.   **(b)** Used shortcut functions.   **(c)** Interacted with map.
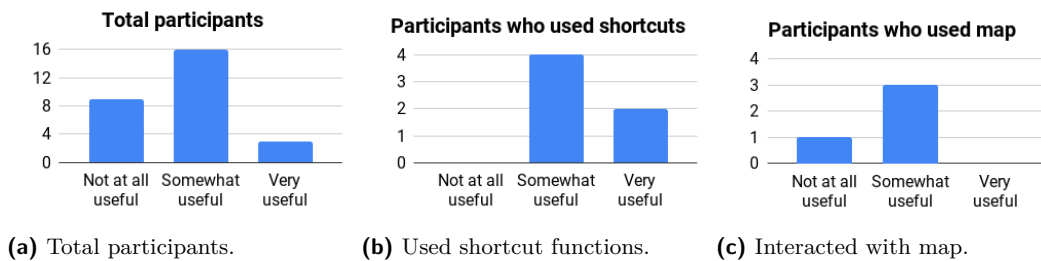
Figure 3 Segmented survey responses to question regarding usefulness.

4.3   Experiments' ecological validity

Observation 3: Ecological validity influences user behaviours.   Column "EV" in Table 1 shows no checkmark (✓) value, highlighting that no experiment is realistic. Consequences of not having a realistic setting were exhibited in our experiment. People did not show signs of

stress and used the mobile phone casually as if no emergency was ongoing, some explicitly confirmed they knew it was a drill. Others were seen coming back for sweaters, phones and even cigarettes. We suggest, then, adding urgency by replicating lifelike emergency attributes or using alternative methods such as giving a reward to the first few people who exit the building. Additionally, the lack of hazard cues was also a contributing factor. A group was observed figuring out if the emergency was real but as they did not see smoke, or other hint, they concluded it was a drill and exited casually. Adding fake smoke or loud noises can be ways to elicit more realistic behaviours, provided ethics approval. Moreover, some participants were more aware about the proximity of a drill. People who registered a week before the drill interacted with the app 25% less (in average) than people who registered the previous day. That is, late registrants were more aware about the coming drill so were more willing to participate, thus affecting genuine app usage behaviours.

5   Conclusion and future work

Our experiment allowed us to get first-hand experience and insights from smartphone use in emergency evacuations. This evidence suggests that usability considerations such as digital map design, cognitive load assessment and ecological validity do play a role in emergency evacuation app adoption, and designers should keep them in mind. Our literature review revealed common problems in the development of mobile applications for emergency evacuations. Data from the experiment ascertained these problems with emergency apps: (1) mobile map design is not properly implemented, (2) cognitive load is not accounted for, especially given the app's nature, and (3) ecological validity is not considered for testing. The resulting observations lay groundwork for the future: building smarter and personalised emergency evacuation systems. Future research aiming to achieve this goal will be oriented towards emergency aid design, cognitively-appropriate algorithms, and experiment design.

References

1 Junho Ahn and Richard Han. An indoor augmented-reality evacuation system for the Smartphone using personalized Pedometry. *Human-Centric Computing and Information Sciences*, 2(1):18, 2012.

2 David Amores, Maria Vasardani, and Egemen Tanin. Early Detection of Herding Behaviour during Emergency Evacuations. In *10th International Conference on Geographic Information Science (GIScience 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

3 Umit Atila, Yasin Ortakci, Kasim Ozacar, Emrullah Demiral, and Ismail Karas. SmartEscape: A Mobile Smart Individual Fire Evacuation System Based on 3D Spatial Model. *ISPRS International Journal of Geo-Information*, 7(6):223, 2018.

4 John W Ayers, Eric C Leas, Mark Dredze, Jon-Patrick Allem, Jurek G Grabowski, and Linda Hill. Pokémon GO—a new distraction for drivers and pedestrians. *JAMA internal medicine*, 176(12):1865–1866, 2016.

5 Luca Chittaro and Daniele Nadalutti. Presenting evacuation instructions on mobile devices by means of location-aware 3D virtual environments. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 395–398. ACM, 2008.

6 Liou Chu and Shih-Jung Wu. An integrated building fire evacuation system with RFID and cloud computing. In *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2011 Seventh International Conference on*, pages 17–20. IEEE, 2011.

7 Arzu Çöltekin, Benedikt Heil, Simone Garlandini, and Sara Irina Fabrikant. Evaluating the effectiveness of interactive map interface designs: a case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science*, 36(1):5–17, 2009.

8   Rachel Harrison, Derek Flood, and David Duce. Usability of mobile applications: literature review and rationale for a new usability model. *Journal of Interaction Science*, 1(1):1, 2013.

9   Yutaka Inoue, Akio Sashima, Takeshi Ikeda, and Koichi Kurumatani. Indoor emergency evacuation service on autonomous navigation system using mobile phone. In *Universal Communication, 2008. ISUC'08. Second International Symposium on*, pages 79–85. IEEE, 2008.

10 Suzanne Kieffer, Ugo Braga Sangiorgi, and Jean Vanderdonckt. ECOVAL: A Framework for Increasing the Ecological Validity in Usability Testing. In *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, pages 452–461. IEEE, 2015.

11 Liisa Kuparinen, Johanna Silvennoinen, and Hannakaisa Isomäki. Introducing usability heuristics for mobile map applications. In *Proceedings of the 26th International Cartographic Conference, August 25 30, 2013, Dresden, Germany, ISBN 978-1-907075-06-3*. International Cartographic Association, 2013.

12 Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. Handheld augmented reality indoor navigation with activity-based instructions. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, pages 211–220. ACM, 2011.

13 Jaziar Radianti, Jose J Gonzalez, and Ole-Christoffer Granmo. Publish-subscribe smartphone sensing platform for the acute phase of a disaster: A framework for emergency management support. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, pages 285–290. IEEE, 2014.

14 Maria Shitkova, Justus Holler, Tobias Heide, Nico Clever, and Jörg Becker. Towards Usability Guidelines for Mobile Websites and Applications. In *Wirtschaftsinformatik*, pages 1603–1617, 2015.

15 Ming-Kuan Tsai, Yung-Ching Lee, Chung-Hsin Lu, Mei-Hsin Chen, Tien-Yin Chou, and Nie-Jia Yau. Integrating geographical information and augmented reality techniques for mobile escape guidelines on nuclear accident sites. *Journal of environmental radioactivity*, 109:36–44, 2012.

16 Lauri Wirola, Tommi A Laine, and Jari Syrjärinne. Mass-market requirements for indoor positioning and indoor navigation. In *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on*, pages 1–7. IEEE, 2010.

17 Xiaolin Yang, Zhongliang Wu, and Yingchun Li. Difference between real-life escape panic and mimic exercises in simulated situation with implications to the statistical physics models of emergency evacuation: The 2008 Wenchuan earthquake. *Physica A: Statistical Mechanics and its Applications*, 390(12):2375–2380, 2011.

Functional Scales in Assisted Wayfinding

Heinrich Löwen ⓘD
Institute for Geoinformatics, University of Münster, Germany
loewen.heinrich@uni-muenster.de

Jakub Krukar ⓘD
Institute for Geoinformatics, University of Münster, Germany
krukar@uni-muenster.de

Angela Schwering ⓘD
Institute for Geoinformatics, University of Münster, Germany
schwering@uni-muenster.de

—— **Abstract** ——

GPS-based navigation systems are widely used to get wayfinding assistance. Current navigation systems incorporate different map scales for presenting wayfinding instructions, however, the selection of scale is not supported by psychological findings. Different tasks of the users such as the identification of the next decision point or the orientation within the environment might be supported best at particular scales. We propose a new conceptual distinction of *functional* scales with respect to their role in supporting wayfinding and orientation. We suggest that these *functional* scales can have a benefit for supporting wayfinding and orientation if used for providing wayfinding instructions. This we aim to empirically evaluate in future work.

1 Introduction

Current navigation systems incorporate different map scales for presenting wayfinding instructions to the users. Typically users may see an overview map of the whole route at the start of the travel. During the travel, the navigation system dynamically scales the map with respect to the speed of travel and the distance to the next decision point. These scale changes are very useful to support the task at hand, which is the interpretation of instruction and the identification of decision points. However, wayfinding support systems might target different task such as spatial learning and the orientation within the local or global context of the route. While the scale changes support the identification of decision points, they are not systematically chosen based on their benefit in spatial knowledge acquisition and orientation. Current research is ignoring that the relevance of environmental features for wayfinding and orientation support might depend on different goals users have during wayfinding, that are best supported at particular scales. To our knowledge, there is no work existing that conceptually distinguishes scales with respect to different functions in assisted wayfinding scenarios. In this paper we propose a conceptual distinction of *functional* scales with respect to their role in supporting wayfinding and orientation. The distinction consists of five scales that vary in the extent of navigationally-relevant space they represent.

2 Background

2.1 Environmental Features supporting Spatial Knowledge Acquisition during Assisted Wayfinding

Landmarks are important features in wayfinding and navigation, because they structure human mental representations of space [18, 2, 12]. There is empirical evidence from the analysis of human route instructions showing that these contain a significant amount of orientation information, including local and global landmarks, which support the acquisition of survey knowledge [1, 10, 9]. The feature selection, which is natural for humans, is not trivial from the computational perspective. Different approaches have been developed to automatically select environmental features such as landmarks for wayfinding and orientation support (e.g. [16, 15, 5, 3]), however empirical evidence with respect to spatial learning is rarely presented.

Others investigated users' spatial knowledge acquisition during assisted wayfinding. Different wayfinding aids such as traditional paper maps and GPS-based navigation systems were compared (e.g. [14, 7, 4]), all showing negative consequences of digital navigation systems on the formation of mental spatial representations. Navigation systems seem to change the way users attend to the environment by providing a sequential set of turn instructions that can be passively followed with little attention to the environment [6, 20, 21]. New types of instructions have been presented to support users' spatial knowledge acquisition and orientation, e.g., *spatial chunking* where elementary wayfinding actions are merged into higher order chunks that convey information about meaningful parts of the route [8]. Schwering et al. [20] suggested to provide instructions not in a turn-by-turn manner but in a holistic way in order to support spatial learning of the route as well as the surrounding environment.

Recent research has shown that the selection and accentuation of map features has a significant influence on users' spatial learning and orientation [11]. The authors described a semi-automatic process of selecting environmental features based on a classification scheme that distinguishes orientation information as landmarks, network structures, and structural regions. They showed that the accentuation of local features supported the acquisition of route knowledge, whereas the accentuation of global features supported the acquisition of survey knowledge. This research, however, neglects that the relevance of environmental features for wayfinding and orientation support might depend on the representation at different scales. Moreover, the suitability of a particular scale depends on the current situation of the driver and the task to be supported by the map visualization.

2.2 Scale

The term *scale* is used for different concepts. Cartographers use it for describing the ratio of real world distance and map distance. They specify how environmental features are represented at particular map scales. Psychologists make a qualitative distinction of *scale* with respect to the perception of space. The dominant distinction was presented by Montello who classifies psychological space into multiple classes based on the projective size of space relative to the human body [13]. He distinguished the classes *figural, vista, environmental,* and *geographical* space. Few works have looked at scales in similar contexts: Richter et al. classified the granularity of environmental features in place descriptions, distinguishing the levels *furniture, room, building, street, district, city,* and *country* [17]. Schmid et al. distinguished three levels of detail in You-Are-Here maps with respect to Montello's psychological spaces and Worboys' nearness relations [22], which they refer to as *immediate neighborhood, larger neighborhood,* and *beyond that horizon* [19].

In the following we propose a conceptual distinction of *functional* scales with respect to their role in supporting wayfinding and orientation. Our classification is derived from the interaction of cartographic and psychological scales: Any map presented on the screen of a wayfinding support system would classify as a *figural* space in Montello's terms; however, for the purpose of supporting wayfinding, this map might represent the extent of space equivalent to either *vista, environmental*, or *geographical* psychological space (such as a turn at the current junction, or a route passing through an entire country). Seeing so different extents of space during distinct phases of an assisted wayfinding scenario is likely to affect the users' ability to spontaneously learn and orient within the environment. While the existing navigation systems utilise this principle in its most simplistic form, e.g., by displaying the route's overview at the beginning of the journey, and zooming in near junctions, this system behaviour is not optimised to continuously support spatial knowledge acquisition.

3   Functional Scales

We suggest five major categories of functional scales of wayfinding maps: *intersection, neighborhood, city, region*, and *route overview* scale. As opposed to cartographic map scales that are expressed in the ratios between real world distance and the corresponding map distance, the functional scales are defined by the containment of features relevant for different aspects of navigation. For example, one of the listed functional scales is required to contain the entire route, no matter of its euclidean length. We relate the functional scale classes to previous definitions of *map scales* and *psychological scales*. The full categorization is shown in Table 1.

The *intersection* scale depicts a particular decision point at a large scale facilitating local orientation and decision making. At this functional scale, maps contain detailed information about local features at the intersection including local landmarks and full layout of the street network. Direction instructions at decision points of contemporary navigation systems can be categorized into this scale, however, only prototypes incorporate landmarks, yet (e.g. Natural Guidance by HERE, Garmin Real Directions). We relate this category to the *vista* space (see [13]) and consider the map scale as fixed with respect to the required screen size and resolution.

The *neighborhood* scale depicts information about the local context of the route in order to support the understanding of the local route context and surrounding connections. Relevant information are considered to be local and global landmarks, the full street network, and structural regions at a size that does not exceed the particular neighborhood. Global landmarks might not be located at the route, but support the overall understanding of the neighborhood. We consider the *neighborhood* scale to be projectively larger than the *intersection* scale. It exceeds the *vista* space, thus can be related to the *environmental* space. The related map scale is considered to be relative to the size of the neighborhood.

The *city* scale depicts information about the global context of the city in order to support the understanding of the global city context and the main city structure. It provides an overview of a whole city or, in case of a drive between cities, the area between two cities. Relevant information at this scale are global landmarks, the main street network, and structural regions at a size that does not exceed the size of the particular city. The *city* scale is considered to be projectively larger than the *neighborhood* scale at a map scale that is relative to the size of the city or the area between two cities. Although a city might only be directly apprehended with a considerable amount of time, it can still be related to the *environmental* space.

▮ **Table 1** Functional Scales in Wayfinding Support.

Functional Scale	Information Content	Function in Wayfinding Support	Related Cartographic Scale	Related Psychological Scale
intersection	– detailed information about DP – building information – local landmarks – full street network	– identification of DP – local orientation – understanding of visual information at DP	fixed large scale *	vista space
neighborhood	– information about local context – local & global landmarks – full street network – structural regions ≤ neighborhood	– understanding of local route context – understanding of surrounding connections	relative to size of neighborhood	environmental space
city	– information about global context of city – global landmarks – main street network – structural regions ≤ city	– understanding of global city context – understanding of city structure and main connections	relative to size of city **	environmental space
region	– information about global context of the region – main street network – structural regions ≥ city	– understanding global region context – understanding of main connections through region	relative to size of cities and regions	geographical space
route overview	– combined information from neighborhood, city and region scale	– understanding of the global route context – getting overview of whole route	relative to length of route	

* e.g., for an average 5 inch smartphone screen this relates to a map scale of 1:1.000 − 1:3000.

** e.g. for the city of Münster, western Germany, this relates to a map scale of 1:100.000 − 1:200.000.

The *region* scale depicts information about the global context of the region in order to support the orientation within the region and support the understanding of the overall structure of the region. It only highlights the main street network and structural regions at a size of the cities or larger; more detailed information such as separate instances of landmarks are not considered as relevant for this scale. We consider the *region* scale to be projectively larger than the *city* scale at a map scale that is relative to the size of the particular region. The environmental spaces represented in the *region* scales are too large to be apprehended directly through locomotion, although the related wayfinding scenario involves locomotion through the region; thus the *region* scale is related to the *geographical* space as defined by Montello.

The *route overview* scale depicts information about the whole route in a single map in order to provide overview of the whole route and surrounding environment and to support the understanding of the global route context. While the previous scale categories are considered as not overlapping and ordered from the *intersection* scale to the *region* scale, the *route overview* scale might overlap with the other scale categories. The related map scale of the *route overview* scale is relative to the length of the route such as to contain the whole route in a single map. We consider environmental features relevant for the *route overview* scale to be composed of information from the *neighborhood* scale, the *city* scale, and the *region* scale; this relates to the structure of the particular route. It was shown that routes have a typical structure, which was divided in three parts: a detailed beginning, a coarse middle, and a detailed end [23]. The route structure is considered in the *route overview* scale, e.g., detailed information about the local route context are depicted around the beginning and the end of the route (see *neighborhood* scale); coarse information about the global context of the city (*city* scale) or even the region (*region* scale) are depicted for route parts consisting of higher order streets such as secondary roads, primary roads or highways. Depending on the length of the particular route, only a subset of the functional scale categories might be relevant; e.g. for a route that lies entirely within a city, the *region* scale is redundant.

Although the functional scales are defined to selectively represent environmental features, this does not solve the problem of small-display cartography to visualize geographic information on small displays with sufficient level of detail. The functional scales are related to map scales, which are relative to the size of the neighborhood, city, or region, or the length of the route. Depending on the actual size of the related features it might not be possible to visualize the defined information content of the functional scales on small displays in a legible way. To cope with this, we refer to ongoing research on the selection of environmental features to support orientation and spatial knowledge acquisition (see [11]).

4   Conclusion

Current navigation systems incorporate different map scales for presenting wayfinding instructions, however, the selection of scale is not supported by psychological findings. We suggest a categorization of *functional* scales of wayfinding maps, which are distinguished by the containment of features relevant for different aspects of navigation. As described above, we suggest that these *functional* scales can have a benefit for supporting wayfinding and orientation if used for providing wayfinding instructions.

In future work, we aim to empirically evaluate the categorization of functional scales in two aspects. On the one hand users' preferences in assisted wayfinding scenarios with respect to the functional scales will be investigated. This aims to get a first insights into and explore the relevance of the functional scales with respect to different route contexts. On the other hand the relevance of environmental features with respect to the functional scales and the effect on spatial knowledge acquisition will be investigated. We thereby target the question what scale is most suited with respect to different functions and contexts in wayfinding and orientation support. Our work contribute to the general understanding of spatial knowledge acquisition in assisted wayfinding scenarios.

References

1   Vanessa Joy Anacta, Mohammed Imaduddin Humayun, Angela Schwering, and Jakub Krukar. Investigating Representations of Places with Unclear Spatial Extent in Sketch Maps. In Arnold Bregt, Tapani Sarjakoski, Ron van Lammeren, and Frans Rip, editors, *Societal Geo-innovation: Selected papers of the 20th AGILE conference on Geographic Information Science*, pages 3–17. Springer, Cham, 2017. `doi:10.1007/978-3-319-56759-4_1`.

2   Gary Burnett. 'Turn right at the Traffic Lights': The Requirement for Landmarks in Vehicle Navigation Systems. *The Journal of Navigation*, 53(3):499–510, 2000.

3   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`.

4   Frank Dickmann. City Maps Versus Map-Based Navigation Systems – An Empirical Approach to Building Mental Representations. *The Cartographic Journal*, 49(1):62–69, 2012. `doi:10.1179/1743277411Y.0000000018`.

5   Matt Duckham, Stephan Winter, and Michelle Robinson. Including landmarks in routing instructions. *Journal of Location Based Services*, 4(1):28–52, March 2010. `doi:10.1080/17489721003785602`.

6   Aaron L Gardony, Tad T Brunyé, Caroline R Mahoney, and Holly A Taylor. How Navigational Aids Impair Spatial Memory: Evidence for Divided Attention. *Spatial Cognition & Computation*, 13(4):319–350, 2013. `doi:10.1080/13875868.2013.792821`.

7   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.2007.09.002`.

8   Alexander Klippel, Heike Tappe, and Christopher Habel. Pictorial Representations of Routes: Chunking Route Segments during Comprehension. In Christian Freksa, Wilfried Brauer, Christopher Habel, and Karl F. Wender, editors, *Spatial Cognition III: Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Learning*, pages 11–33. Springer, Berlin, Heidelberg, 2003.

9   Rui Li, Amichi Korda, Maurin Radtke, and Angela Schwering. Visualising distant off-screen landmarks on mobile devices to support spatial orientation. *Journal of Location Based Services*, 8(3):166–178, 2014. `doi:10.1080/17489725.2014.978825`.

10   Kristin L. Lovelace, Mary. Hegarty, and Daniel R. Montello. Elements of Good Route Directions in Familiar and Unfamiliar Environments. In Christian Freksa and David M Mark, editors, *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, pages 65–82. Springer, Berlin, Heidelberg, 1999. `doi:10.1007/3-540-48384-5_5`.

11   Heinrich Löwen, Jakub Krukar, and Angela Schwering. Spatial Learning with Orientation Maps: The Influence of Different Environmental Features on Spatial Knowledge Acquisition. *ISPRS International Journal of Geo-Information*, 8(3), March 2019. `doi:10.3390/ijgi8030149`.

12   Andrew J May and Tracy Ross. Presence and Quality of Navigational Landmarks: Effect on Driver Performance and Implications for Design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 48(2):346–361, June 2006. `doi:10.1518/001872006777724453`.

13   Daniel R. Montello. Scale and Multiple Psychologies of Space. In Andrew U. Frank and Irene Campari, editors, *Spatial Information Theory: A Theoretical Basis for GIS, Proceedings COSIT '93*, Lecture Notes in Computer Science, pages 312–321. Springer, 1993.

14   Stefan Münzer, Hubert D. Zimmer, Maximilian Schwalm, Jörg Baus, and Ilhan Aslan. Computer-assisted navigation and the acquisition of route and survey knowledge. *Journal of Environmental Psychology*, 26(4):300–308, December 2006. `doi:10.1016/j.jenvp.2006.08.001`.

15   Clemens Nothegger, Stephan Winter, and Martin Raubal. Selection of Salient Features for Route Directions. *Spatial Cognition & Computation*, 4(2):113–136, June 2004. `doi:10.1207/s15427633scc0402_1`.

16   Martin Raubal and Stephan Winter. Enriching Wayfinding Instructions with Local Landmarks. In Max J. Egenhofer and David M. Mark, editors, *Geographic Information Science: Second International Conference, GIScience 2002*, Lecture Notes in Computer Science, pages 243–259, Boulder, CO, USA, 2002. Springer. `doi:10.1007/3-540-45799-2_17`.

17   Daniela Richter, Maria Vasardani, Lesley Stirling, Kai-Florian Richter, and Stephan Winter. Zooming In–Zooming Out: Hierarchies in Place Descriptions. In Jukka M Krisp, editor, *Progress in Location-Based Services*, Lecture Notes in Geoinformation and Cartography, pages 339–355. Springer, 2013. `doi:10.1007/978-3-642-34203-5_19`.

18   Kai-Florian Richter and Stephan Winter. *Landmarks: GIScience for Intelligent Services.* Springer, Cham, Switzerland, 2014.

19   Falko Schmid, Colin Kuntzsch, Stephan Winter, Aisan Kazerani, and Benjamin Preisig. Situated Local and Global Orientation in Mobile You-are-here Maps. In Marco de Sá, Luís Carriço, and Nuno Correia, editors, *Proceedings of the 12th Conference on Human Computer Interaction with Mobile Devices and Services*, Mobile {HCI} '10, pages 83–92, Lisbon, Portugal, 2010. ACM.

20   Angela Schwering, Jakub Krukar, Rui Li, Vanessa Joy Anacta, and Stefan Fuest. Wayfinding Through Orientation. *Spatial Cognition & Computation*, 17(4):273–303, October 2017. `doi:10.1080/13875868.2017.1322597`.

21   Katharine S Willis, Christoph Hölscher, Gregor Wilbertz, and Chao Li. A comparison of spatial knowledge acquisition with maps and mobile maps. *Computers, Environment and Urban Systems*, 33(2):100–110, March 2009. `doi:10.1016/j.compenvurbsys.2009.01.004`.

22   Michael F Worboys. Nearness relations in environmental space. *International Journal of Geographical Information Science*, 15(7):633–651, October 2001. `doi:10.1080/13658810110061162`.

23   Dieter Wunderlich and Rudolph Reinelt. How to Get There From Here. In R. J. Jarvella and W. Klein, editors, *Speech, Place, and Action*, pages 183–201. John Wiley & Sons Ltd., 1982.

Representation of Interdependencies Between Urban Networks by a Multi-Layer Graph

Laura Pinson
Normandie Univ, INSA Rouen, UNIROUEN, UNIHAVRE, LITIS, 76000 Rouen, France
laura.pinson@insa-rouen.fr

Géraldine Del Mondo
Normandie Univ, INSA Rouen, UNIROUEN, UNIHAVRE, LITIS, 76000 Rouen, France
geraldine.del_mondo@insa-rouen.fr

Pierrick Tranouez
Normandie Univ, Univ Rouen, UNIROUEN, UNIHAVRE, LITIS, 76000 Rouen, France
pierrick.Tranouez@univ-rouen.fr

—— **Abstract** ——

The RGC4 (Urban resilience and Crisis Management in a Context of Slow Flood to Slow Kinetics) project aims to develop tools to help manage critical technical networks as part of the management process of crisis in a context of slow kinetic flooding in Paris. This project focuses on cascading models to identify a number of inter-dependencies between networks and to define tools capable of coordinating the actions of managers before and during the crisis. This paper revisits the conceptual and methodological bases of networks approach to study the inter-dependencies between networks. Research that studies the return to service of infrastructure networks often angle it from the perspective of operational research. The article proposes a graph theory perspective based on a multi-layer network approach and shows how to characterize the inter-dependencies between networks at three process levels (macro, meso, micro)

1 Introduction

The ever-increasing city services are based on the growing complexity of urban technical networks (electricity, water supply, transport, telecommunications, etc.). However, these networks, generally interdependent, are highly vulnerable to hazards and extreme weather events. The localized failure of a network component can impact several services over large areas, sometimes well beyond the areas directly subject to the trigger hazard. Recent disasters, such as Hurricane Katrina in New Orleans, have contributed to the development of the concept of urban resilience. Urban resilience is a key focus of current approaches to flood management. The notion of resilience encompasses pre-disaster planning and warning systems, emergency handling procedures and post-disaster reconstruction. The concept of resilience leads in particular to an interest in the post-disaster period and consequently in the phenomenon of "reconstruction" and "return to normal" [9]. Research that studies the return to service of urban networks often responds to this from the perspective of operational

research. The objective is to optimize the return to service of these networks under the constraints of availability of resources [13], [10].

This research proposes to identify a number of inter-dependencies between networks related to *resilience*. While we may know enough inter-dependencies between networks to be able to simulate the impacts of flood on networks and their cascading failures, it is not clear how to schedule infrastructure return-to-service / troubleshooting, with the additional difficulty of recognizing that managers may have conflicting interests. This analysis of inter-dependencies requires to model cascading effects[7]. These cascades significantly increase the vulnerability of the urban system and makes recovery and reconstruction processes more difficult and slower after a disruption. We described our networks as multi-layer-graphs upon which we modeled this "inverted domino effect" by topological operations.

In this paper we chose to describe in depth the topological structure of our model, but not go into details of the involved algorithms.

2   Modeling urban services with a multi-layer graph

We use graphs to model networks and our resilience issue because graph algorithms and metrology on large graphs highlight possible structural and functional properties related to interactions. To avoid a terminological confusion, further on we will employ "Network" when we mean the real world organization and "Graph" when we mean our model of the Network.

Choice of modeling by graph

To understand the resilience of networks during flood periods, it is important to model their failure dynamics. This modeling, through a graph, requires identifying the entities (vertices) and the relationships (edges/arcs) that connect them either in space or through a more abstract dependency link [1]. Graph modeling represents the information either by a global vision or by a representation at lower scales (structural properties of networks) [14]. The objective is to use the structure and the semantics of the graph obtained to answer the problem of inverted domino effects and to produce indicators to characterize the inter-dependencies between networks in a given territory.

A model we found of particular interest was the multiplex graph of[11]. A multiplex graph is a graph composed of a set of vertices of the same type, linked by different types of relationships. A multiplex is therefore a multi-relational graph that is often represented by a multi-layer graph. Multi-layer graphs explicitly incorporate multiple channels of connectivity and constitute the natural environment to describe systems interconnected [2]. Layers can be interdependent and they contain information which would be lost if we only considered the corresponding aggregated network. It has also been shown recently that different types of dynamics that are run on top of multi-layer systems also provide new insights into the problems being modeled [4], [5]. So, a multi-layer systems consists of several distinct classical layers, each one encoding a specific type of information about the system. Many complex systems can be represented as networks consisting of distinct types of interactions, which can be categorized as links belonging to different layers [6]. The question is then how many layers are indeed necessary to accurately represent our structure of a multilayered complex system to model urban service systems.

Relation between graphs and layers

In this work we model urban services and associated technical networks with a multi-layer graph whose layers represent three levels of study: macro, meso, micro (relationships between the same components (micro level), relations between different infrastructures (meso level) and relations between different urban systems (macro level)). Assuming that urban services are defined by their infrastructure and components and are interconnected, urban services should not be defined as objects but rather through the networks that create them.

We model urban services (ex: electricity, water, railway, buses and metros, etc.) at one scale with vertices which can be detailed by connecting them to graphs of a different layer, representing a different scale. The aggregation approach by urban service and infrastructure leads us to consider not only urban technical networks (macro level) but also the infrastructures that structure each urban service (meso level) and their components (micro level). We thus build a particular type of multi-layer graph with the idea of a multiplex graph, that is to say a sequence of interconnected graphs.

To give an example (Figure 2, C and D), Paris RATP urban service (macro level) consists in a subset of infrastructures S such as metro stations, railway, etc. (meso level). The metro station concept belongs to the meso level while each individual metro station belongs to the micro level. The infrastructures of the micro level are spatialized. The metro stations taken one by one (e.g. Auber, Bercy, Créteil, etc. respectively $S1$, $S2$, $S3$) are geo-referenced and form a new layer(micro level). At each level, relationships and inter-dependencies exist between respectively urban services, infrastructures and spatialized components. Because of the peculiar interconnected structure, it is possible to move from one layer to another one.

3 Our model

The objective is to study the return to service strategies for different urban networks based on multi-layer graphs. Knowing that, we have chosen to focus on Paris' own urban technical networks: the RATP rail network, the ENEDIS electricity network and the road network as well as their infrastructures and components. In this section we present in details our model.

Methodology

We model the disruption of a network by the suppression of one or more arcs or vertices [12]. This modelling leads to the study of failure scenarios. In our funding project, flood are the failure causing event we are meant to investigate. To create these flood scenarios, two steps are necessary. First of all, we would like to simulate the impact of the crisis to obtain a graph representing the disturbed network. This first step is an application of deconstruction rules. Then, the objective is to reconstruct the graph in order to return to the state of the initial network. The idea in this second step is to propose schedules for the return to service of the installations. To this end, we would like to add semantic information on vertices and arcs of graphs, depending on the information and data available. This information will allow us to take into account the network disruption in the graph.

The multi-layer graph

The model is represented by a multi-layer graph $\mathbf{M}$. The structure is defined by : $\mathbf{M} = (G, C)$ where G is a directed graph such as:

$G = \{G^\alpha \mid \alpha \in [\![1..3]\!]\}$ with $G^\alpha = (V^\alpha, E^\alpha, \mu^\alpha, \varepsilon^\alpha)$

V^α is the set of vertices, E^α the set of arcs and

μ^α the set of weights/attributes on the arcs and

ε^α the set of weights/attributes on the vertices

C is the set of interconnections between vertices of different layers (G^α)

We write G^1 for the macro level, G^2 for the meso level and G^3 for the micro level. These G^α are called layers. We will call *semantic sub-graphs* sub-graphs of G^α connected through C to one vertex of $G^{\alpha-1}$ (Figure 1).
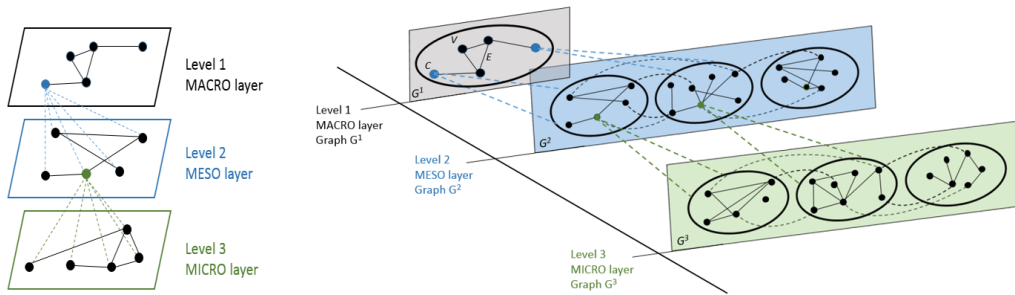
Figure 1 Illustration of a multi-layer graph at different levels representing three analysis layers: macro, meso, micro.

4   Semantic in the graph

Since networks consist of a large number of infrastructures and components, it is necessary to define the most important and primary elements (i.e.critical) which are essential for civil society. The aim is to determine whether the impacts and repercussions of an urban network, infrastructure or component are really disastrous or if they generate only low incidents during a flood. This is what we will call *criticality*. The concepts of vulnerability, resilience, and criticality are interrelated. Resilience is defined as "the capacity of a system to absorb disturbance and re-organize while undergoing change so as to still retain essentially the same function, structure, identity and feedbacks" [16]. We identify specific vulnerability to flooding of a particular critical facility by looking at factors such as its use, past flooding issues, location of critical systems like primary and back-up power to better understand its criticality. In addition, knowing the vulnerability and criticality of an element provides recommendations and/or resources to critical facility managers for short or long-term changes that could be made to reduce their facilities risk to flooding. If we take the example of the ENEDIS (electrical network) and its electrical installations (micro level), they have a variable vulnerability to flooding [8]. Lines, buried or overhead, are considered as not very vulnerable. On the other hand, some equipment clearly appears vulnerable, such as pylons in case of high flows [8]. The vulnerability of transformer stations (meso level) depends on water level (and turbidity too). The electrical network (macro level) is generally meshed to a high level of detail (almost everywhere up to distribution to individuals (micro level)). This mesh allows a station to temporarily or permanently take over from a failed station. However, in

the case of floods, this mesh provides little protection. Indeed, floods often affect very large areas, so all stations of the same mesh can be affected and cuts occur directly at customers' premises when they are flooded. If the power grid is very large aggressive towards other networks, it seems little dependent on other networks. Nevertheless, telecommunication networks and road networks are necessary for on-site repairs and communication in times of crisis. The criticality of the elements of the multi-layer network is modeled by vertices and arcs attributes. The idea is to evaluate the criticality of the actors of different networks according to their environment.

Criticality expression

At each level of detail, networks, infrastructure and components are vulnerable to flooding. They will have repercussions and impacts on their own networks. The criticality and vulnerability assessment process establishes priority between urban services, infrastructures and components.

Criticality on arcs is evaluated in order to know their vulnerability and to prioritize the actions to be taken (Figure 2, A and B) . This comparative assessment is based on certain criteria, in particular the dependence of the networks on each other, both in normal times and during floods. The notion of criticality is closely linked to the damage to the network, infrastructure and components (depending of the level of detail) after a flood [15]. Values of criticality may change during events depending on the water level. Over the three studied levels, a criticality scale is established and each edge $e \in E$ in the graph is valuated by this measure (*Criticalitylevel(e)*). The criticality levels are:

- Level 1: no special vigilance. The impact of the flood will be very low or even zero
- Level 2: pay attention. Impacts low but may become higher if the flood persists
- Level 3: very critical. Impacts heavy and will lead to a dysfunction in the resilience capacity of components, infrastructures and networks.
- Level 4: absolute criticality. Impacts exceptional and lead to a secession of activities.

For the macro and meso level, two other values (in percentage) value arcs. The first value (*networkDependencyNorm(e)*) corresponds to the level of dependence between the two vertices linked under normal circumstances, for users. The impact of flooding on vertex dependency defines the second value (*networkDependencyFlood(e)*). These values are intended to assess the fragility of the vertex and the importance of the vertex in the graph (Figure 2, A).

For example, RATP network (macro level) is very dependent on ENEDIS network. RATP network used by nearly 10 million passengers per day depends on ENEDIS' resources. RATP needs the electricital network to operate its transportation network. A simple network interruption on few lines can cause losses and repercussions for travellers to their workplace. On the other hand, in the event of a flood, ENEDIS' impact on RATP network will be low (provided that the network is not damaged, too) since in the event of flooding, RATP closes the exits to the transport routes with cofferdams. Above all, RATP wishes to resist damage caused by ice jam shocks and the pressures of other urban components. Thus, to give an example, the arc between RATP and ENEDIS network is valued as such: 4; 80; 30.

At the meso level, in the ENEDIS graph, the transformer used to transmit and distribute electricity (adapt the voltage) is highly dependent on the source stations. In the event of a flood, the flooding of the transformer (leakage problems) will lead to a deterioration of the source substations and will lead to a disruption or even interruption of the power supply. The transformer to the source stations therefore has a high criticality (for users), as well as a dependency and strong impacts during flood periods when normal weather conditions prevail. Thus, to give an example, the arc between transformer and source stations is valued as such: 3; 80; 80.

Criticality on vertices is taken into account according to their neighbourhood. Each vertex $v \in V$ is valuated with $d(v), d^-(v), d^+(v)$ (degree, input degree and output degree of a vertex v). Applied to this context, the notion of neighbourhood implies that a link with a poorly connected vertex is less critical than a link with a highly connected vertex (Figure 2, A, B).
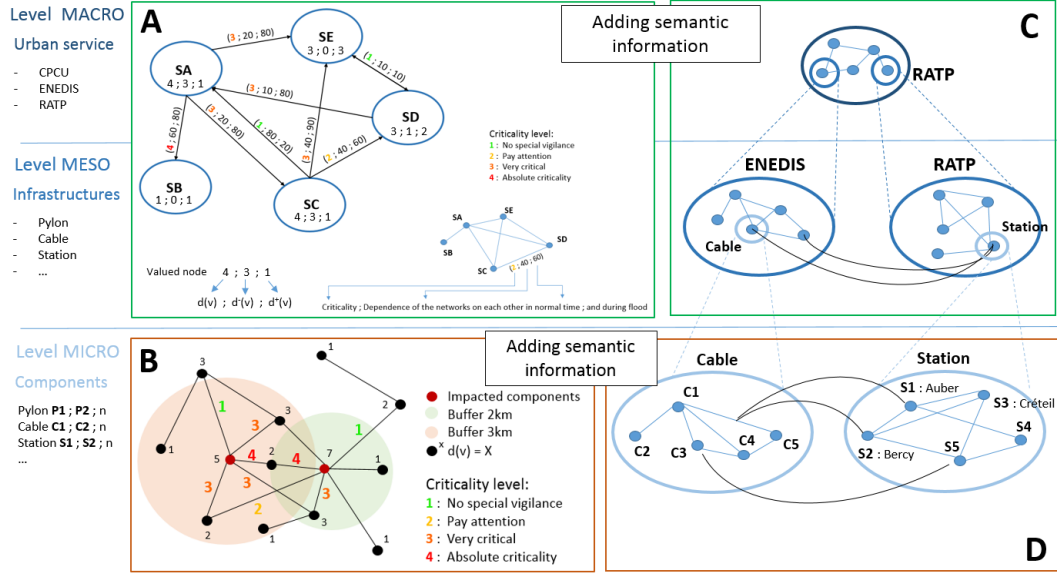
Figure 2 Formalization of inter-dependencies at three levels of study with the addition of semantics on our different graphs.

5   Environmental modelling

5.1   Formalisation

As illustrated in Figure 1, vertices from a G^α layer are connected to vertices of a $G^{\alpha+1}$ layer.

Let $hol^{\alpha+1}$ be:

$$hol^{\alpha+1} : \quad \begin{array}{ccc} \mathbf{V^{\alpha+1}} & \longrightarrow & \mathbf{V^\alpha} \\ v^{\alpha+1} & \longmapsto & v^\alpha \end{array}$$

For example if v^3 is the power station at the angle of street A and Avenue B, $v^2 = hol^3(v^3)$ is the infrastructure "Power station":"hol" is short for "holon" which describes the same notion in other formalisms.

Let RC^α be:

$$RC^\alpha : \quad \begin{array}{ccc} \mathbf{V^\alpha} & \longrightarrow & \mathscr{P}(\mathbf{V^{\alpha+1}}) \\ v^\alpha & \longmapsto & (hol^{\alpha+1})^{-1}(v^\alpha) \end{array}$$

"RC" is short for "returnComponent" which describes the same notion in other formalisms.

Finally we call $C^\alpha = \{(v_l^\alpha, RC(v_l^\alpha)), l \in [1.. \mid V^\alpha \mid]\}$ and $C = \bigcup_{\alpha=1}^{3} C^\alpha$ (section 3)

Macro level / Meso level

The graph $G^1 = (V^1, E^1, \mu^1, \varepsilon^1)$, models the macro level. It is defined by its set of vertices V^1 (urban service) and its set of arcs E^1 where

$V^1 = \{v_i^1 \mid i \in [1..nbServices]\}$ and
$E^1 = \{(v_k^1, v_j^1) \mid \exists$ an interdependence link between v_k^1 and v_j^1, $v_k^1, v_j^1 \in V^1, j \neq k\}$ with
$\mu^1 : V^1 \longrightarrow \mathbb{N} \times \mathbb{N} \times \mathbb{N}$
$\quad v \longmapsto (d(v), d^-(v), d^+(v))$
$\varepsilon^1 : E^1 \longrightarrow \mathbb{N} \times \mathbb{R} \times \mathbb{R}$
$\quad e \longmapsto (Criticalitylevel(e), networkDependencyNorm(e), networkDependencyFlood(e))$

The formalisation of the meso level is quite similar to the macro level except that it deals with infrastructures.

Micro level

The graph $G^3 = (V^3, E^3, \mu^3, \varepsilon^3)$, models the micro level. It is defined by its set of V^3 vertices (components) and its set of E^3 arcs where
$\quad V^3 = \{v_i^3 \mid i \in [| \ nbComponent \ |]\}$ and
$\quad E^3 = \{(v_k^3, v_j^3) \mid \exists$ a link of functional or spatial interdependence between v_k^3 and v_j^3, $v_k^3, v_j^3 \in V^3, j \neq k\}$

For the micro level, since the dependencies are not necessarily functional but geographical, the neighbourhood is based on the closest neighbours, by defining a buffer area. Each component have its own buffer with its own distance. The distance then becomes the radius of a circle since the vertex (representing our component) is a point, the induced surface of the circle, the buffer area (e.g Figure 2 B). According to this definition μ^3 and ε^3 formal definition is quite similar to macro and meso level without $networkDependencyNorm(e)$ and $networkDependencyFlood(e)$
$\quad \varepsilon^3 : E^3 \longrightarrow \mathbb{N}$
$\quad \quad e \longmapsto Criticalitylevel(e)$

5.2 Consideration of scheduling algorithms

This formalization gives the definition of the graph structure modeling the interdependency of technical networks. It models the relationships of the urban networks, at different levels of study and with their semantics. However, in order to identify strategies for rebuilding the network after a flood, different constraints must be taken into account, particularly in terms of resources, time and materials [3]. The objective is to prioritize these needs. The multi-layer graph models several networks. Our scheduling problem is therefore to give an order on operating tasks for the reconstruction of the activity of these urban networks while respecting the constraints. This structure, combined with scheduling algorithms, will make it possible to: identify and characterize "critical vertices" and their links, decide on the allocation of a network's need and arbitrate between the managers needs. After the flood, it can take days or even months for affected networks to return to normal operation. A poor consideration of this risk of impacts on the networks may lead to a significant additional delay before the territory is restored to normal.

6 Conclusion

This paper describes a methodology to model the vulnerability and resiliency of inter-dependant urban services networks (water, electricity, transportation etc.) and introduces a multi-layer approach to provide a sound support for resilience issues ans crisis management. We define a multi-layer graph to model each network, their infrastructural elements (duct, filter station etc.; pylon, transformer etc.; railway, station etc.) and their individual, spatialized components (the filter station at the angle of A street and B street; transformer number 1657

etc.). As it is, the graph model suggested is mainly dynamic how interactions and cascading affects can be modelled.

On this graph, short range reliance between two connected elements (two close pylons, a railway joining two stations) can be provided by the operators of the network. When rebuilding after a disaster, this operational knowledge is put to use to order the necessary operations. Nonetheless, long range dependency is much more difficult to assess for operators, especially if it requires knowledge external to the network they operate, and can therefore lead to clearly sub-optimal decisions. Topological operations on the proposed structure can compute this long range dependency, thus aiding rebuilding ordering decision making, and improving the resilience of the urban networks in the process.

References

1 Michael Batty. A new theory of space syntax, 2004.

2 Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesús Gómez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, and Massimiliano Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014.

3 Paolo Bocchini and Dan M Frangopol. Restoration of bridge networks after an earthquake: Multicriteria intervention optimization. *Earthquake Spectra*, 28(2):426–455, 2012.

4 Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025, 2010.

5 Emanuele Cozzo, Raquel A Banos, Sandro Meloni, and Yamir Moreno. Contact-based social contagion in multiplex networks. *Physical Review E*, 88(5):050801, 2013.

6 Manlio De Domenico, Vincenzo Nicosia, Alexandre Arenas, and Vito Latora. Structural reducibility of multilayer networks. *Nature communications*, 6:6864, 2015.

7 Ian Dobson, Jie Chen, JS Thorp, Benjamin A Carreras, and David E Newman. Examining criticality of blackouts in power system models with cascading events. In *Proceedings of the 35th annual Hawaii international conference on system sciences*, pages 10–pp. IEEE, 2002.

8 MEDD (Ministère de l'Environnement et du Développement Durable. Réduire la vulnérabilité des réseaux urbains aux inondations, 2005.

9 Yan Guo. Urban resilience in post-disaster reconstruction: Towards a resilient development in Sichuan, China. *International Journal of Disaster Risk Science*, 3(1):45–55, 2012.

10 Zhi-Hua Hu and Zhao-Han Sheng. Disaster spread simulation and rescue time optimization in a resource network. *Information Sciences*, 298:118–135, 2015.

11 R Kanawati. Co-authorship link prediction in multiplex bibliographical networks. In *Multiplex network workshop-European conference on complex systems (ECCS'13)*, 2013.

12 Serge Lhomme. Analyse théorique et quantitative de la remise en service de réseaux d'infrastructures: définition d'un cadre d'analyse fondé sur les propriétés structurelles de ces réseaux. In *Treizièmes rencontres Théo Quant*, 2017.

13 Timothy C Matisziw, Alan T Murray, and Tony H Grubesic. Strategic network restoration. *Networks and Spatial Economics*, 10(3):345–361, 2010.

14 Céline Rozenblat. Approches multiplexes des systèmes de villes dans les réseaux d'entreprises multinationales. *Revue d'Économie Régionale & Urbaine*, page 393–424, 2015.

15 Marie Toubin. *Améliorer la résilience urbaine par un diagnostic collaboratif, l'exemple des services urbains parisiens face à l'inondation.* PhD thesis, Université Paris-Diderot-Paris VII, 2014.

16 Brian Walker, Crawford S Holling, Stephen Carpenter, and Ann Kinzig. Resilience, adaptability and transformability in social–ecological systems. *Ecology and society*, 9(2), 2004.

Route Choice Through Regions by Pedestrian Agents

Gabriele Filomena[1]
Institute for Geoinformatics, University of Münster, Germany
gabriele.filomena@uni-muenster.de

Ed Manley
The Bartlett Centre for Advanced Spatial Analysis, University College London, London, UK
The Alan Turing Institute, London, UK
ed.manley@ucl.ac.uk

Judith A. Verstegen
Institute for Geoinformatics, University of Münster, Germany
j.a.verstegen@uni-muenster.de

─── **Abstract** ───

Simulation models for pedestrian movement are valuable tools to support decision-making processes in urban design. However, existing models of pedestrian behaviour are built on simplistic assumptions regarding people's representation of the urban space and spatial behaviour. In this work, a route-choice algorithm that takes into account regionalisation processes and the hierarchical organisation of geographical elements is adapted for pedestrian movement and incorporated into an agent-based model. The macro-level patterns emerging from two scenarios, one employing an angular-change minimisation algorithm and the other employing the regional algorithm here proposed, are compared for a case study in London, UK. Our routing algorithm led agents to recur to a higher number of street segments, i.e. routes were more diverse among agents. Though validation has not yet been performed, we deem the patterns resulting from the regional algorithm more plausible.

1 Introduction

The movement of people in cities, be them cyclists, pedestrians, drivers and transit users, has proved to be one of the most challenging subjects of study in urban dynamics research [21]. Urban travellers interact with the city environment and its manifold phenomena, shaping the city form as well as its economic and cultural structures. As such, gaining insights into people's movement and spatial behaviour may support cities in decision making as concerns transport infrastructure, wayfinding signage design, service allocation and urban configuration redevelopment.

In this context, geosimulation is considered a tool which "enhances our understanding of how cities function and evolve in space-time" [8, p. V]. In particular, Agent-Based Modelling (ABM) allows researchers and experts to understand how individuals' goals and choices mould flows at the macro level [7]. However, the ability of these models to capture such dynamics depends on the theoretical assumptions and design considerations of the modeller

---

[1] Corresponding author

about how an agent formulates routes across the urban environment [12]. In most of the existing representations, agents' route selection processes are modelled as functions derived from utility theory [1]. Herein, it is assumed that urban travellers make spatial choices and thereby generate routes by assigning costs to different alternatives. A utility measure is pursued and computed by the agent on the basis of time, distance or attractiveness [13, 17].

Simulation models for pedestrian movement in urban contexts are quite sporadic. Even harder is to find exhaustive attempts to implement cognitive representations of space in ABM. A set of works has been inspired by the Space Syntax approach and the idea that the configuration of the street network guides pedestrian movement [9]. Penn and Turner [16], integrating Space Syntax techniques and ABM, enriched agents with information regarding visibility at junctions. Jiang [10] devised an ABM for pedestrian simulation whose main postulate is that the interaction between agents and the street configuration alone may account for the self-organisation of pedestrian patterns. More recently, Omer and Kaplan [15] designed an ABM wherein agents choose destinations on the basis of a land-use attractiveness measure, and employ different kinds of path-selection criteria (Euclidean distance, number of turns and angular change minimisation).

These models mainly make use of street segments properties along the lines of utilitaristic approaches to spatial behaviour and, furthermore, do not contemplate agents endowed with symbolic representations of the urban space. Yet, other geographical elements are known to be important. Kevin Lynch [11] and successive research in cognitive geography widely suggest that individuals' representations of the city are built upon multiple categories of urban elements – nodes, paths, districts, landmarks and edges –, which are significant with respect to spatial behaviour, navigation and human-environment interaction. Moreover, several studies have gathered empirical evidence on the hierarchical organisation of these elements in human knowledge [14, 20], a type of structure which reflects the "degree of recognition and the idiosyncratic relevance of individual objects" [4, p. 257] in the urban environment. These findings may prompt a more realistic and complete representation of individuals' spatial behaviour in simulation models [6, 13].

The aim of this work is to advance an ABM for simulating pedestrian movement which embraces a cognitively-grounded, hierarchical routing framework. We include in the simulation a route-choice model built upon the framework presented in [13], adjusted for pedestrian movement. Therein, the author advances a bounded-decision making approach to route-choice behaviour in light of findings on the hierarchical organisation of spatial knowledge relative to urban elements, and regionalisation processes. In our ABM, we introduce a scenario in which agents are equipped with a simple cognitive, two-level hierarchical representation of the urban space, which comprises a coarse regional division of the city and fine-grained information about main street segments and junctions. Macro-level patterns emerging from the inclusion of such elements in the simulation are compared to the outcomes emerging from a scenario in which agents use a single-level cost-minimisation approach.

2 Methodology

In the ABM for pedestrian movement simulation here introduced, *agents* – representing walkers – complete trips through the *environment* – the street network of the case-study area – between pairs of origins and destinations (OD). Two different scenarios are designed: in the first case, agents use the common single-level utility approach, minimising angular change – *AC scenario* –, in the second, they employ the routing model presented below, here called regional routing algorithm – *RR scenario*.

The model proposed by Manley [13] embodies different planning levels in the route-choice process by representing an initial rough global plan, subsequently refined at higher granularity levels. This framework was designed and validated with taxi driver routing data and it is here adjusted and integrated into an ABM for pedestrian movement.

In summary, at first, nodes are extracted from a multilayer network and ranked by a centrality measure. Afterwards, functional regions are identified from the street network by means of a community detection technique, and finally employed within the route-choice model. Therefore, in the ABM, a spatial hierarchy is built at two levels: nodes are classified by salience and manipulated accordingly for the extraction of OD pairs; concurrently, a containment hierarchy is represented by a two-steps decision process, from the urban- to the street-level, when formulating a route.

Nodes and districts identification

Cognitive salient nodes are anchoring points, easy to remember and associated with the procedural component of the spatial knowledge. Centrality measures have proven to be able to differentiate between primary and secondary nodes [5]. In [6], *betweenness centrality* is employed to extract main nodes from the street network. However, we claim that the transit network should also be taken into account to capture meaningful urban nodes. Considering different urban layers, their interactions and their structure, allows to better understand how places are connected [19]. Therefore, the betweenness centrality of a node is here computed through a multilayer representation of the urban system [3] composed of two layers, the road network and the transit network (see figure 1 a).

Euclidean distance is used to weight links in the two networks as well as transfer edges (i.e. the distance between the street junction and the public transport station).
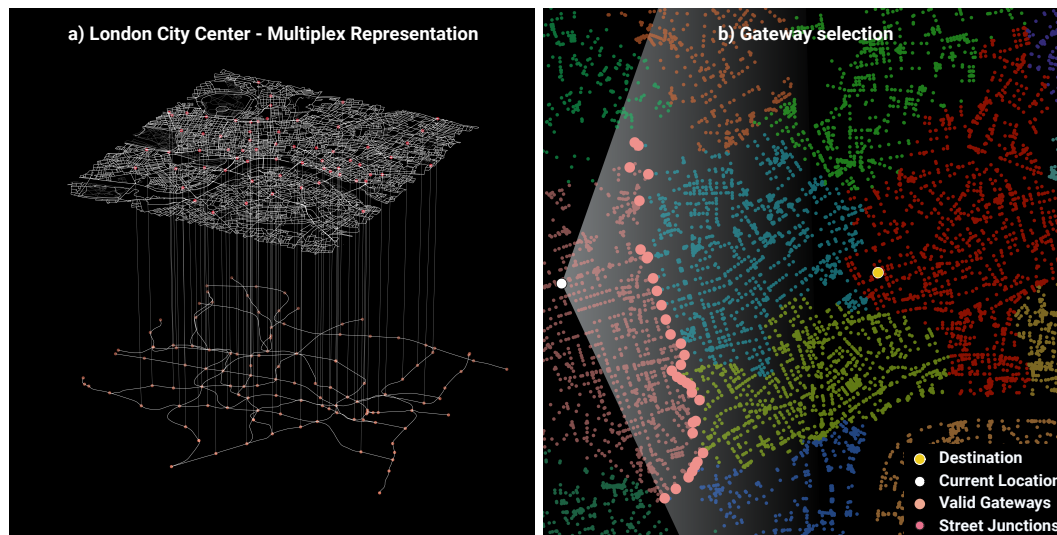
Figure 1 a) A multilayer-representation of the central area of London, UK: transit (below) and street (above) networks. b) Identification of possible gateways based on the location and the final destination of the agent.

The *modularity optimisation* algorithm [2] is employed to identify functional regions from the street layout. This algorithm is a community detection technique which optimises modularity, namely the robustness of a possible division in communities of a network. The

community membership of the street segments is derived from topological ties existing in a dual graph representation, namely a graph wherein nodes represent street segments, links represent connections amongst them. Afterwards, each street junction is assigned to its region.

Modelling route-choice behaviour

To begin with, when a trip is formulated, the origin and destination nodes are randomly chosen with a probability based on their betweenness centrality value, i.e. the betweenness centrality values are linearly re-scaled to probabilities, such that the node with the highest betweenness centrality has the highest probability to be selected as an origin or destination. Furthermore, the destination is picked drawing from nodes located outside the origin's region.

The route-choice approach adopted here [13] follows the hierarchical structure in which the urban environment is decomposed: the agents' decisions shift from the regional- to the street-level. In other words, it is assumed that a walker, before conceiving a detailed street-segment path, decides upon a sequence of regions to traverse to reach the destination. At this initial stage, the algorithm moves from one region to another until the destination region is found. The selection of each next region is performed making use of *gateways*, namely pairs of *exit* and *entry* nodes located at boundaries between regions. Such gateways are roughly evaluated every time a new region is entered on the basis of the following rules [13]:

- The Euclidean distance between the *destination* and the possible exit node must be shorter than the distance separating the *current location* from the destination node.
- The exit node should be in the direction of the destination node: the angle formed by the current location and the possible exit is supposed to be between the one formed by the current location and the destination $\pm\alpha$ degrees on each side. In this work, we subjectively set the α parameter to 70°, instead of 90° as in [13], to coerce the agent to exclude gateways with a high deviation from the destination, assuming that pedestrians are less inclined to take large detours compared to drivers (see figure 1 b).
- The *entry node* belonging to the next possible region should be in the direction of the destination as well.

The current location either corresponds to the origin of the route, or, across the computation, to an entry node. In a nutshell, such criteria constrain the gateway selection process to candidates that are towards the destination region, relative to the position of the agent. When multiple choices satisfy the minimum requirements, the gateway with the lowest deviation from the destination is selected. The search process moves to the next region until the destination region is reached.

At the street decision level, the agent formulates a more precise path, selecting nodes between each pair of gateways. Decisions are based on an intra-region cost-minimisation approach. Angular change minimisation is used as a criterion [18] for its ability to predict peoples' movement and account for cognitive heuristics. The series of regional-nodes are merged and the complete path is generated. Figure 2 presents a summary of the steps described above within the ABM environment.

The case study

London (UK) is chosen as a case study. The road network and the urban railway network (Underground, Overground and Docklands Light Railway lines) are used to generate the multilayer representation. In each ABM scenario, agents are set to perform 1000 trips across the city, between pairs of OD separated by a maximum distance of 4000 meters.
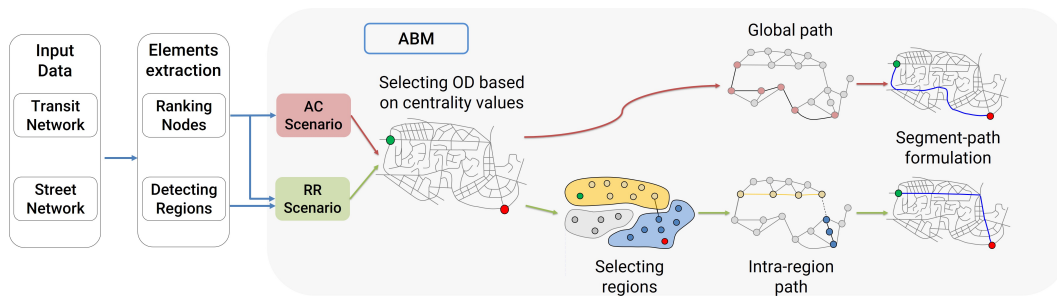
Figure 2 Summary of the methodology steps: nodes and paths between origin (red) and destination (green) are coloured in yellow, blue (RC scenario) or light red (AC scenario); graphics adapted from [13].

During the simulation, every single street segment records the number of times that it is traversed by an agent. In order to account for the randomness introduced by the selection of OD pairs, the scenarios are executed ten times; the mean of the flow of pedestrians across the different runs is calculated per segment and used thereby to compare the macro-level patterns emerging from the AC and RR scenario.

3 Results

The angular change shortest-path appears to bring about a low spatial variability of pedestrian segment usage across the case-study area (see figure 3). Most of the agents in this scenario made use of major roads to reach the city centre from the outer districts or vice versa. The A201 artery (including Farringdon Road and Blackfriars Bridge), in particular, was often traversed and emerged as the main link between the south and the north (some segments go to a maximum of 2400 crossings), from Elephant and Castle up to King's Cross. Likewise, the A40, along with the north bank of the Thames, was used to move from west to east. Many street segments were never crossed by the agents in this scenario (see figure 3 and 4).
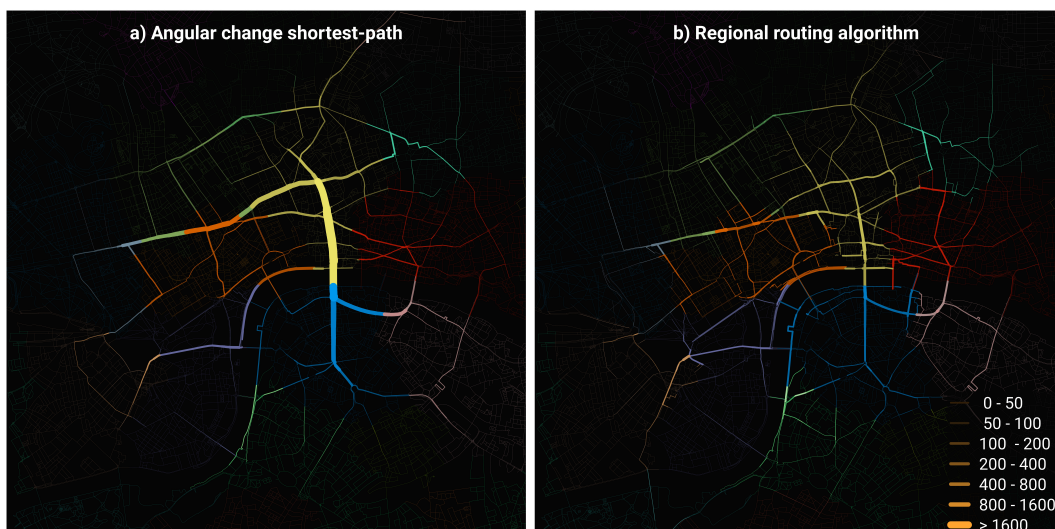
Figure 3 Street segments usage for 1000 trips in the ABM scenarios. Street segments are coloured by district membership; brightness and width indicate the number of agents' crossings.
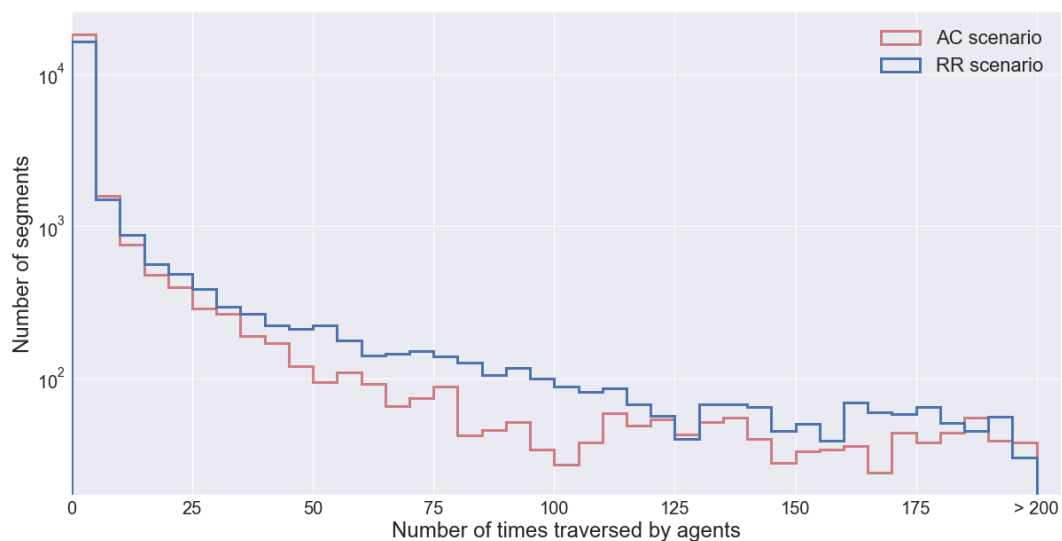
■ **Figure 4** Frequency distribution of pedestrian density values across street segment features in the two scenarios.

Even though the A201 played a big part in the RR scenario as well, the agents exploited a wider range of minor roads to reach their destinations, leading to a more diversified pattern. The central districts, coloured in orange and yellow, exhibit a higher number of street segments with relatively high agents densities: 545 street segments were crossed more than 200 times in the RR scenario, against 434 in the AC scenario. The district coloured in red, although displaying a quite defined pattern, was traversed slightly more regularly by agents in the RR scenario (168 and 145 segments respectively above 200 counts); as a link between the north and the south, street segments in this region were probably used as an alternative to the A201. Indeed, in the RR scenario, along this road, the highest number of crossing is between 800 and 1000, almost 60% less in comparison with the AC scenario. The South Bank (blue district) shows a higher spatial variability, in contrast to the other scenario. Visible paths along the southern riverfront even emerge towards the east, probably as a result of the recourse to the Millennium and the Southwark Bridges (coloured in red), nearly invisible in the routes of agents in the AC scenario.

Figure 4 summarises these observations: while the AC scenario displays a larger number of segments that were rarely or not even traversed, the RR scenario presents higher frequencies at almost each crossing category higher than 10. At the same time, the AC scenario also presents a higher number of segments crossed more than 200 times, further suggesting a more extreme distribution of the flows. Out of 1335946 kilometres of street network, considering an average distance per journey of 1648 (RR) and 1529 (AC) meters, 58095 km of street segments were featured by more than 200 crossings in RR, 49400 km in the AC scenario. 1047528 km were crossed at least once in the RR scenario, 977006 km in the AC scenario.

On the whole, the south and the central areas are the ones where most differences between the scenarios arise. Generally speaking, the outer regions of the case-study area are less traversed. This may be attributable to an edge-effect deriving from the centrality computation. The central-eastern part of the city seems to be the most preferred in both conditions, whereas the north-western street segments of the city centre do not exhibit relevant differences between the scenarios.

4   Discussion

When compared to a single-level cost-minimisation scenario, the results of the regional routing scenario seem more plausible both at the agent- and the macro-level. Regional routing led agents to take advantage of different streets and diversify routes, believably in relation to the gateways' positions. By travelling across alternative paths to major roads, regional routing agents spread out through the street network and determined more balanced flow patterns. Moreover, at the micro-level, the spatial constraints introduced by the morphological structure of the regions and their reciprocal connections reduced behavioural uniformity amongst agents.

In light of these preliminary results, the methodology here presented could be further developed, at different levels. The node hierarchy employed to manipulate the selection of OD pairs could be adapted to prevent agents to wander primarily in the central area, almost avoiding segments along the case-study boundaries. Concerning districts, the selection of gateways could be better tuned by taking into account the cognitive salience of nodes. Furthermore, individual differences between agents can be explicitly included in the simulation, assuming that urban explorers traverse specific junctions based on their knowledge of the environment. Finally, a validation of the ABM with observational-data could provide insights regarding the performance of the model and/or the routing algorithm. Such step will be carried out in the next phases of the model's development by comparing the distribution of pedestrian across the street networks, per each segment, with densities obtained by pedestrian GPS trajectory data.

References

1   Gustavo Kuhn Andriotti and Franziska Klügl. Agent-Based Simulation Versus Econometrics – from Macro- to Microscopic Approaches in Route Choice Simulation. In K. Fischer, I.J. Timm, E. André, and N. Zhong, editors, *Multiagent System Technologies. MATES 2006. Lecture Notes in Computer Science, vol 4196*, pages 61–72. Springer, Berlin, Heidelberg, 2006.

2   Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):1–12, October 2008. `doi:10.1088/1742-5468/2008/10/P10008`.

3   Stefano Boccaletti, Ginestra Bianconi, Regino Criado, C. I. del Genio, Jesus Gómez-Gardeñes, Miguel Romance, Irene Sendiña-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014. `doi:10.1016/j.physrep.2014.07.001`.

4   David Caduff and Sabine Timpf. On the assessment of landmark salience for human navigation. *Cognitive Processing*, 9(4):249–267, November 2008. `doi:10.1007/s10339-007-0199-2`.

5   Paolo Crucitti, Vito Latora, and Sergio Porta. Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(015113):1–9, 2006. `doi:10.1063/1.2150162`.

6   Gabriele Filomena, Judith A. Verstegen, and Ed Manley. A computational approach to 'The Image of the City'. *Cities*, 89:14–25, 2019. `doi:10.1016/J.CITIES.2019.01.006`.

7   Carlos Gershenson. Self-Organizing Urban Transportation Systems. In Juval Portugali, Han Meyer, Egbert Stolk, and Ekim Tan, editors, *Complexity Theories of Cities Have Come of Age*, pages 269–279. Springer, Heidelberg, 2012.

8   Marco Helbich, Jamal Jokar Arsanjani, and Michael Leitner, editors. *Computational Approaches for Urban Environments*. Springer, Cham, Heidelberg, 2015. `doi:10.1007/978-3-319-11469-9`.

9   Bill Hillier and Julienne Hanson. *The social logic of space*. Cambridge University Press, Cambridge, UK, 1984.

10   Bin Jiang and Tao Jia. Agent-based simulation of human movement shaped by the underlying street structure. *International Journal of Geographical Information Science*, 25(1):51–64, 2011. `doi:10.1080/13658811003712864`.

11   Kevin Lynch. *The image of the city.* MIT Press, Cambridge, MA, 1960.

12   Ed Manley, Tao Cheng, Alan Penn, and Andy Emmonds. A framework for simulating large-scale complex urban traffic dynamics through hybrid agent-based modelling. *Computers, Environment and Urban Systems*, 44:27–36, 2014. `doi:10.1016/j.compenvurbsys.2013.11.003`.

13   Ed Manley, Shepley W. Orr, and Tao Cheng. A heuristic model of bounded route choice in urban areas. *Transportation Research Part C: Emerging Technologies*, 56:195–209, 2015. `doi:10.1016/j.trc.2015.03.020`.

14   Timothy P. McNamara, James K. Hardy, and Stephen C. Hirtle. Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2):211–227, 1989. `doi:10.1037/0278-7393.15.2.211`.

15   Itzhak Omer and Nir Kaplan. Using space syntax and agent-based approaches for modeling pedestrian volume at the urban scale. *Computers, Environment and Urban Systems*, 64:57–67, July 2017. `doi:10.1016/J.COMPENVURBSYS.2017.01.007`.

16   Alan Penn and Alasdair Turner. Space Syntax Based Agent Simulation. In *1st International Conference on Pedestrian and Evacuation Dynamics*, Duisburg, 2001.

17   Greg Rybarczyk. Simulating bicycle wayfinding mechanisms in an urban environment. *Urban, Planning and Transport Research*, 2(1):89–104, 2014. `doi:10.1080/21650020.2014.906909`.

18   Edward K. Sadalla and Daniel R. Montello. Remembering changes in direction. *Environment and Behavior*, 21(3):346–363, 1989. `doi:10.1177/0013916589213006`.

19   Emanuele Strano, Saray Shai, Simon Dobson, and Marc Barthelemy. Multiplex networks in metropolitan areas: Generic features and local effects. *Journal of the Royal Society Interface*, 12(111):1–12, 2015. `doi:10.1098/rsif.2015.0651`.

20   Holly A. Taylor and Barbara Tversky. Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31(2):261–292, 1992. `doi:10.1016/0749-596X(92)90014-O`.

21   Paul M. Torrens. Computational Streetscapes. *Computation*, 4(37):1–38, 2016. `doi:10.3390/computation4030037`.

Modeling and Representing Real-World Spatio-Temporal Data in Databases

José Moreira[1]
Department of Electronics, Telecommunications and Informatics, University of Aveiro,Portugal
Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Portugal
jose.moreira@ua.pt

José Duarte
Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Portugal
hfduarte@ua.pt

Paulo Dias
Department of Electronics, Telecommunications and Informatics, University of Aveiro,Portugal
Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Portugal
paulo.dias@ua.pt

⎯⎯ Abstract ⎯⎯

Research in general-purpose spatio-temporal databases has focused mainly on the development of data models and query languages. However, since spatio-temporal data are captured as snapshots, an important research question is how to compute and represent the spatial evolution of the data between observations in databases. Current methods impose constraints to ensure data integrity, but, in some cases, these constraints do not allow the methods to obtain a natural representation of the evolution of spatio-temporal phenomena over time.

This paper discusses a different approach where morphing techniques are used to represent the evolution of spatio-temporal data in databases. First, the methods proposed in the spatio-temporal databases literature are presented and their main limitations are discussed with the help of illustrative examples. Then, the paper discusses the use of morphing techniques to handle spatio-temporal data, and the requirements and the challenges that must be investigated to allow the use of these techniques in databases. Finally, a set of examples is presented to compare the approaches investigated in this work. The need for benchmarking methodologies for spatio-temporal databases is also highlighted.

1 Introduction

Nowadays, there are many technologies and devices that capture large amounts of data about the position, shape and size of spatial phenomena over time. Although there are several commercial and open-access tools for storing and processing spatial data, few exist to deal

---

[1] Corresponding author

with the evolution of spatial data over time. There are two main approaches to handle spatio-temporal data called the discrete model and the continuous model. In the discrete model spatio-temporal data are represented as sets of cells (points) in time (1D) and space (2D). This approach is simple, but the resolution of the model depends on the size of the cells and the computational cost can be cumbersome when handling large amounts of data.

In the continuous model data are represented in vector mode using abstract data types usually referred to as moving objects, such as, moving points, moving lines, and moving polygons (also called moving regions). These are represented as ordered sequences of points, lines or polygons (observed values), and functions describing their spatial transformations (e.g., translation, rotation, scaling and deformation) during the time interval between two consecutive observations [4, 6]. Data models and query languages exist that largely meet the needs identified in the literature to model the evolution of spatial data over time, and there are also prototype systems such as Secondo [8] and Hermes [19] which have been developed manly for research and teaching. However, their use in real case studies is almost non-existent.

Reconstructing data between observations is not trivial. Current solutions focus on the creation of interpolations that are robust and valid, but do not create realistic approximations of the evolution of spatio-temporal phenomena between observations in some circumstances. In fact, the search for methods capable of obtaining interpolations that are robust and realistic at the same time, compatible with the requirements usually imposed to represent spatial and temporal data in databases, remains a research topic.

In this paper, our goal is to show that morphing techniques, used to create video animations among other applications, can be investigated to represent the evolution of spatio-temporal data in databases. We are particularly interested in modeling spatio-temporal data using continuous models of time and space. We show that morphing techniques can obtain interpolations that are closer to the actual evolution of real-world phenomena than the interpolations created using the methods proposed in the spatio-temporal databases literature, and we identify the constraints and the issues that must be studied to enable the use of morphing techniques to create spatio-temporal data. The aim is to find solutions to represent data with small errors of approximation, so that they can be used in scientific and engineering applications, and that can be used to develop methods and algorithms to perform spatio-temporal operations in databases.

The remainder of this paper is organized as follows. Section 2 presents an overview on modeling and querying spatio-temporal data in databases. Emphasis is given to the methods proposed in the literature to create spatio-temporal data from observations. Section 3 introduces the use of morphing techniques in the context of spatio-temporal databases. Two main issues are covered: polygon matching and the representation of the evolution of spatial data between known observations. Section 4 presents examples comparing the results obtained when using one of the main methods proposed in the spatio-temporal databases literature and a morphing technique. Section 5 concludes this work and presents guidelines for future research.

2   The region interpolation problem in spatio-temporal databases

2.1   Spatio-temporal data models and query languages

A moving object is a triple (T, S, f), where $T \subset \mathbb{R}$ is the time domain, $S \subset \mathbb{R}^2$ is the spatial domain and $f : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ is a function that gives the transformation of S during T, such that $(t, x, y) \in \mathbb{R} \times \mathbb{R}^2 \mid (\exists x')(\exists y')(\, t \in T \,\wedge\, (x', y') \in S \,\wedge\, (x, y) = f(t, x', y')\,)$ [10]. This

is an abstract definition that must be transposed into a discrete (finite) model suitable for implementation in a database system. Several data models and query languages have been proposed, but the most complete, and also the most successful, is, by far, the approach based on abstract data types [4, 6, 9]. This data model allows the representation of objects such as moving points, moving lines and moving regions that may have complex shapes, e.g., regions with holes. The authors also propose a comprehensive set of spatio-temporal operations, such as, projections (e.g., the shape of a moving region at a given time instant, the footprint of a moving region during a time interval or numerical measures, such as, the velocity or the size of a moving region over time), topological operations to evaluate the interaction of a moving object with other moving or static objects, distance operations and predicates. In addition, abstract data types can be smoothly integrated in the database management systems currently in use, which has also contributed to the success of this approach over other interesting proposals, namely, spatio-temporal constraints databases [3, 7].

One of the most interesting features of the data model proposed in [6] is the sliced representation (Figure 1).
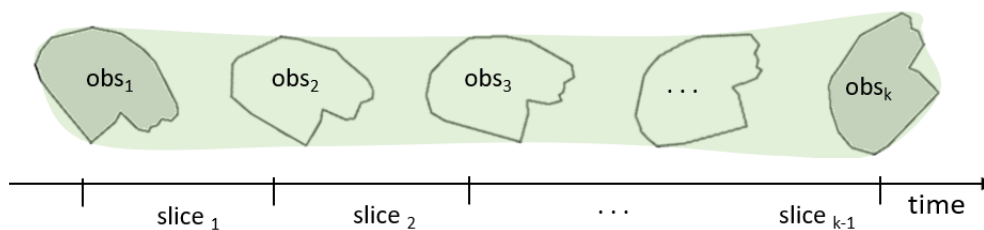
Figure 1 Sliced representation of a moving region.

The example shows the development of a moving region from left to right. Each polygon represents an observed value, i.e., the known shape of a moving region at a specific time instant. Each slice represents the development of the moving region between two consecutive observations. The definition of a slice includes, at least, the position and the shape of the moving object at a given time, and a function used to represent the spatial transformations (translation, rotation, skewing, etc.) of the object between consecutive observations. Moreover, constraints on the continuity of consecutive slices must be imposed.

A major challenge is to find a function that represents the spatio-temporal behavior of moving objects between observations as closely as possible. This problem is particularly complex when handling moving regions and moving lines, and it is often referred to as the Region Interpolation Problem. In this paper, this problem is discussed from two different points of view: special-purpose solutions proposed in the databases research community (Section 2.2) and general-purpose morphing techniques (Section 3).

2.2 Creating spatio-temporal data from observations

The region interpolation problem in spatio-temporal databases has been studied for the first time in [22]. The main contribution of this work is the so-called rotating-plane algorithm, that is used in subsequent works, which allows moving regions to be created from an ordered sequence of observations. Figure 2 illustrates the interpolation between a source (P) and a target (Q) region obtained using this algorithm.

The algorithm scans the line segments in P and Q one by one (e.g., in counter-clockwise order), starting with the ones having the smallest angle relatively to the x-axis in each shape ($\bar{p}_1$ and $\bar{q}_1$, respectively). The movement from P to Q is described by linear equations.
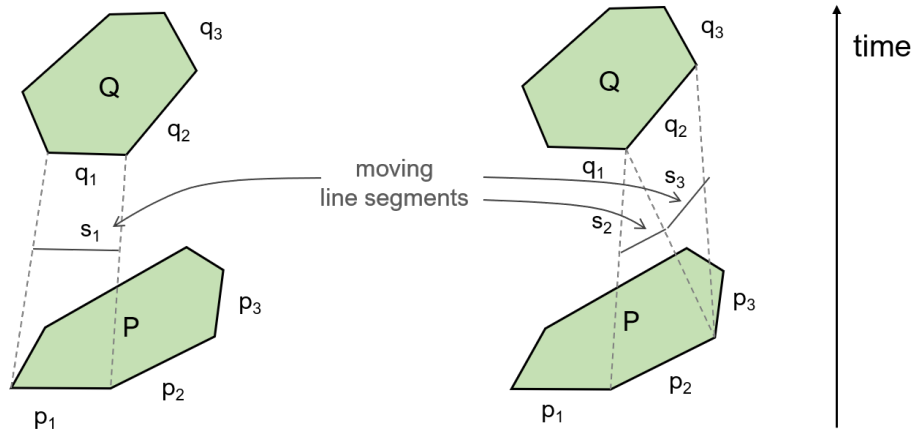
■ **Figure 2** Rotating-plane algorithm: when the line segments (e.g., $\bar{p}_i$ and $\bar{q}_j$) are parallel (left-hand side); when the angle of the line segments (e.g., $\bar{p}_{i+1}$ and $\bar{q}_{j+1}$) is different (right-hand side).

When the angles of the selected line segments in P and Q are equal (e.g., $\bar{p}_1$ and $\bar{q}_1$), a linear transformation of $\bar{p}_1$ into $\bar{q}_1$ is performed (see the moving line segment $\bar{s}_1$), and the algorithm goes to the next segment in both shapes ($\bar{p}_2$ and $\bar{q}_2$, respectively). When the angle of the selected segment in P is smaller than the angle of the selected segment in Q (e.g., $\bar{p}_2$ and $\bar{q}_2$), the first ($\bar{p}_2$) degenerates progressively (see the moving line segment $\bar{s}_2$) into the first point (in counter-clockwise order) of the other ($\bar{q}_2$) and the algorithm goes to the next segment ($\bar{p}_3$) of P. The procedure is equivalent when the angle of the selected segment in P (e.g., $\bar{p}_3$) is greater than the angle of the selected segment in Q (e.g., $\bar{q}_2$). In this case, a point in P becomes a line segment in Q (see the moving line segment $\bar{s}_3$). So, as depicted in the right-hand side of the figure, when the angle of the selected line segments in P and Q is different, the movement of a line segment from the source to the target is given by the movement of two (or more) line segments ($\bar{s}_2$ and $\bar{s}_3$).

This algorithm does not allow line segments to rotate, which is interesting to avoid invalid interpolations. An interpolation is invalid if, for example, there are line segment intersections during interpolation. This choice makes the implementation of spatio-temporal operations easier, because the movement of the vertices is described by linear equations, and it ensures that the 3D representation of the resulting moving region in (x, y, t)-space is a polyhedron. It also helps keeping compatibility with existing spatial DBMS, which, in most cases, are not able to handle curves. However, the approximation of the evolution of moving regions that rotate is poor, as illustrated in Figure 3.

This example shows a polyhedron representing a fixed moving region that rotates approximately 45 degrees. Although the shapes of the source and target are equal, we observe that the intermediate shape of the moving object estimated using the rotating-plane algorithm at the middle of the interpolation differs greatly from P and Q. This problem arises because, as exemplified in Figure 2, the rotation of a line segment is represented implicitly by the linear movement of two (or more) line segments.

As the rotating-plane algorithm is only able to handle convex shapes, [22] proposes to split non-convex shapes into convex features and to organize them in a convex hull tree. Each node (feature) is a hole in the convex hull of the shape in the parent node. Finding a correspondence between the features in the two convex hull trees is difficult. This issue is approached superficially in [11, 22].

time
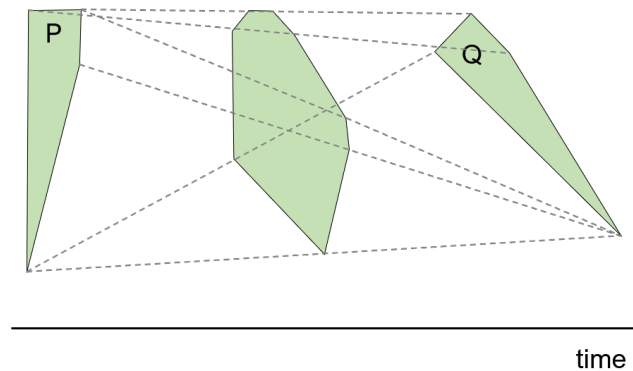
■ **Figure 3** Interpolation created using the rotating-plane algorithm.

In [16] a counter-example demonstrating that the rotation-plane algorithm is not robust is presented. The authors also argue that it is not always possible to create a single interpolation between a source and a target that is valid at all times. So, they propose an algorithm to split an interpolation into three parts at most to avoid line segment intersections during interpolation. Concavities are collapsed into or expanded from a single point (depending on whether the concavities are in the source or in the target, respectively), and intersecting concavities are detected and removed using a process called evaporation (concavities disappear), and condensation (concavities appear later in the interpolation). This can cause an anomalous deformation of the moving region during interpolation, but it has the advantage that the interpolations are always valid. This is demonstrated using an example involving highly complex (snail-shaped) shapes.

This work is extended in [15] to handle moving regions with holes and with a variable number of components (multi-regions). This includes dealing with transformations, such as, splitting and merging regions during interpolation. Almost at the same time, [11] revisited the work in [22] to make it robust by using a strategy similar to the one used in [16].

All the methods mentioned above are based on the rotating-plane algorithm. Yet, this algorithm has well-known issues, particularly, when representing moving regions that rotate and when representing concavities. For that reason [12] presents a different approach to handle moving regions with fixed shape. It presents a data model that can handle curves, and algorithms to compute spatio-temporal operations, namely, the spatial footprint of a moving region, the intersection of a moving region with a static object and the intersection of a moving region with a moving point. Moving segments are allowed to rotate, but the shape of the region is fixed, i.e., the only spatial transformations allowed are translation and rotation.

3 Using morphing techniques in spatio-temporal databases

3.1 Creating interpolations using morphing techniques

The transformation of a source into a target is a problem that has been studied since the beginning of the 1990s in areas such as video animation, gaming and medical imaging. Many techniques were proposed for the interpolation of 2D images, free-form curves, planar shapes (e.g., polygons and polylines) and volumetric representations (3D objects) [20]. The morphing of free-form-curves and planar shapes shares some similarities with the region interpolation problem in spatio-temporal databases. The main objective is to obtain an interpolation that is smooth and realistic, providing visually appealing animations.

Some morphing techniques use iterative methods, which allow for sophisticated interpolations, but the computational costs are high and therefore they would not be suitable for processing queries on large datasets. Yet, there are approaches to estimate the shape of a moving region between observations using formulas. This is the case of the approach proposed in [1], which takes as input two meshes created using a compatible triangulation algorithm, e.g., [14, 21]. Thus, the source and the target meshes have the same number of triangles and there is a one-to-one correspondence between them. The interpolation is given by the affine transformation of each triangle of P into the corresponding triangle in Q and is obtained using Single Value Decomposition. Since the transformation of each triangle is independent of the transformation of its neighbors, it is necessary to calculate a unique position for the shared vertices. Simple solutions, such as, computing the midpoint or using a least squares formulation, can be used for that purpose. A reformulation of this problem using normal equations is presented in [2]. Figure 4 displays an interpolation created using this approach.
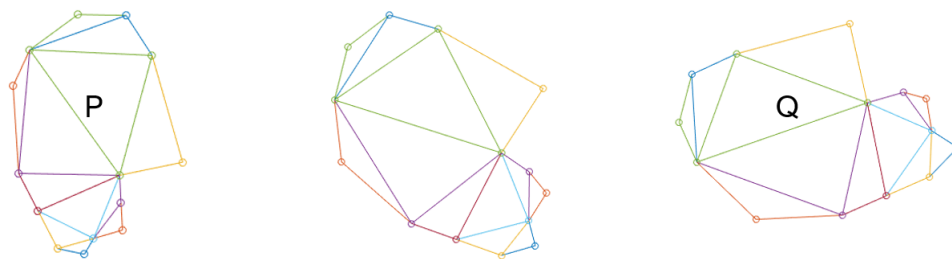
Figure 4 Interpolation created using the method proposed in [1, 2].

The results presented in [1, 2] show that it is possible to obtain realistic interpolations even when the shapes are very different and complex (Figure 5).
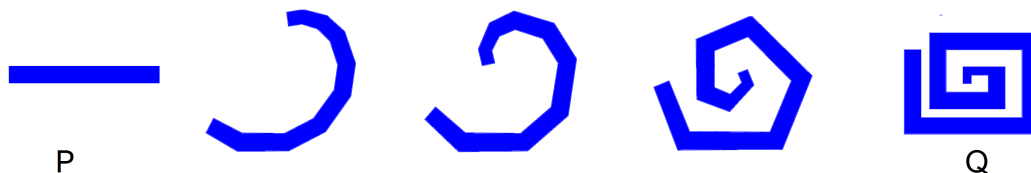
Figure 5 Another interpolation created using the method proposed in [1, 2].

The representation of a complex shape as a mesh of triangles allows using divide-and-conquer strategies to implement complex operations. For instance, two moving regions intersect if any triangle of P intersects at least one triangle of Q during the interpolation. In addition, there are already several algorithms to deal with triangles in computational geometry that may be useful to implement spatio-temporal operations. Finally, since the interpolation is given by a transformation matrix, the shape of a moving region can be estimated at any given time using a single formula.

3.2   Polygon matching

When transforming a shape into another shape some notion of correspondence or matching between the two shapes must be defined. This leads to another problem known as the vertex correspondence problem or polygon matching. Polygon matching consists of finding a correspondence between the elements (e.g., vertices or edges) of two shapes. This paper outlines two representative approaches to solve this problem.

The first uses the concept of feature points [13, 17], which are a subset of vertices that best represent the shape of the objects (e.g., the numbered vertices in Figure 6). A feature point is described using measures, such as the angle and the distance to its neighbors, which ideally should be invariant under translation, rotation and scaling. The mapping between the features points in a source and a target is done using similarity functions. If the number of vertices between two features points in the source is different from the number of vertices between the two corresponding feature points in the target, new vertices are added. This creates a one-to-one correspondence between all vertices in the source and the target, as required by many interpolation algorithms.
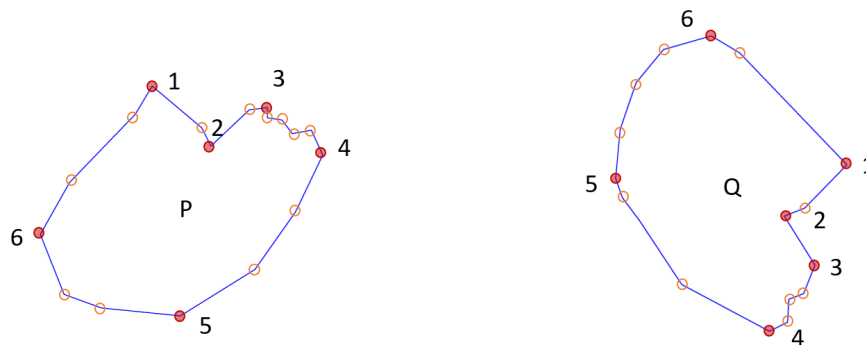
Figure 6 Polygon matching using feature points.

The second approach uses turning functions to find a correspondence between the vertices of two shapes. The method consists of representing the length of the edges (d) and the turning angles α of a source and a target in a two-dimensional chart (Figure 7).
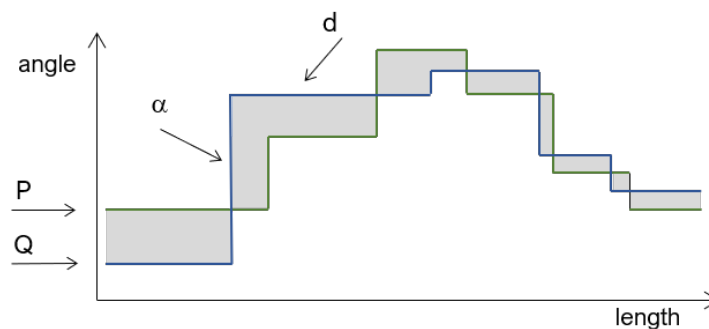
Figure 7 Polygon matching using turning functions.

The sum of the lengths of the segments in each polygon must be equal to 1. The algorithm consists of shifting edges from left to right or vice-versa, to find the mapping of the vertices in P and Q that minimizes the sum of the colored areas (gray rectangles) in the figure.

4   Representation of spatio-temporal data using morphing techniques

4.1   Examples

This section compares the methods discussed in section 2 with the morphing technique presented in Section 3, focusing on the representation of real-world phenomena in spatio-temporal databases. Since the representation of spatial transformations such as translation

and scaling is easy, the emphasis is given to moving regions rotation and morphing. The data used in the examples were extracted from satellite images monitoring the movement of two large blocks of an iceberg in the Antarctic region. The images were segmented to extract the shape of the icebergs at different dates, and the correspondence between the shapes of consecutive pairs of observations of an iceberg was obtained using the method proposed in [13].

The example in Figure 8 shows the evolution of the shape of an iceberg between two consecutive observations (P and Q). The three snapshots in the middle were estimated using the method proposed in [11]. The predominant spatial transformation is rotation and the shape of the iceberg has a large concavity.
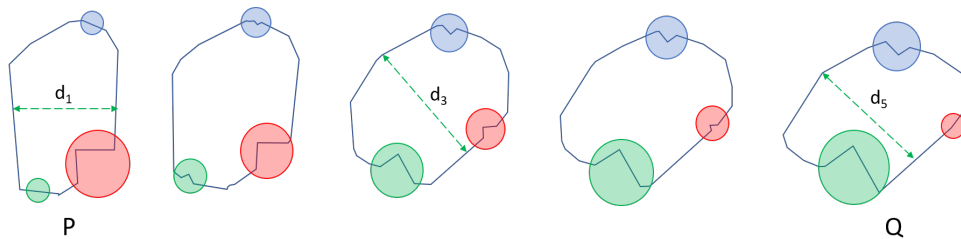
Figure 8 Interpolation created using the method proposed in [11].

This example highlights the main issues with the methods presented in section 2 when considering deformable moving regions. First, it is observable that the shape of the moving region tends to inflate (expand) until it reaches the middle of the interpolation and has an opposite behavior during the second half of the interpolation (the length of d_3 is greater than the lengths of d_1 and d_5). This is also observable in Figure 3. This means that numerical measures, e.g., the area, tend to increase and decrease during the interpolation, which usually is not the expected behavior in real-world phenomena. The anomalous deformation is due to the constraints imposed by the rotating-plane algorithm, which does not allow line segments to rotate (the rotation of a line segment is simulated by the movement of two or more line segments that move linearly). This algorithm is also used in [15] and [11, 22], and so, the results are similar. It is also important to note that the polygon matching step is not performed in the rotating-plane algorithm: the first pair of line segments to be processed is chosen using heuristics, e.g., the smallest angle relatively to the x-axis, which may not be a good choice in many cases.

This figure also highlights issues on handling concavities. The light blue circles show a normal case where the concavity at the top appears progressively during the interpolation. However, the concavity on the bottom right does not go along with the rotation of the object from P to Q. Instead, it is artificially divided in two: the concavity in P, which disappears progressively (marked by the light red circles), and the concavity in Q, which appears progressively (marked by the light green circles) during the interpolation[2]. This is caused by the method proposed in [22] to map the concavities in the convex-hull trees of the source and target. The mapping is based on heuristics, such as, the distance between centroids or the percentage of overlapping between features (in this context, a feature is a convex-polygon representing a hole), but, as shown in [18], this method is not safe, particularly with noisy data.

---

[2] Best illustrated in the video animation accessible from `http://most.web.ua.pt/resources/cosit2019/concavities.mp4`

When using the method proposed in [16], concavities tend to vanish because each concavity is mapped to a point in the convex-hull of the other shape, and they reappear during the second half of the interpolation. This is a good strategy to guarantee that interpolations are robust, but may cause important deformations as shown in Figure 9.
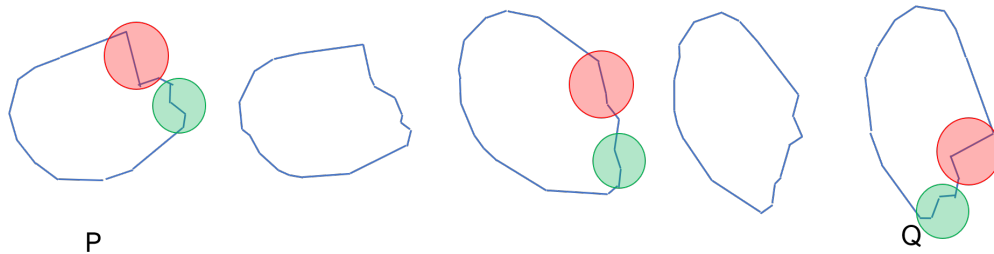
Figure 9 Interpolation created using the method proposed in [16].

Figure 10 displays an interpolation created using compatible triangulation [14] and the morphing technique proposed in [1, 2].
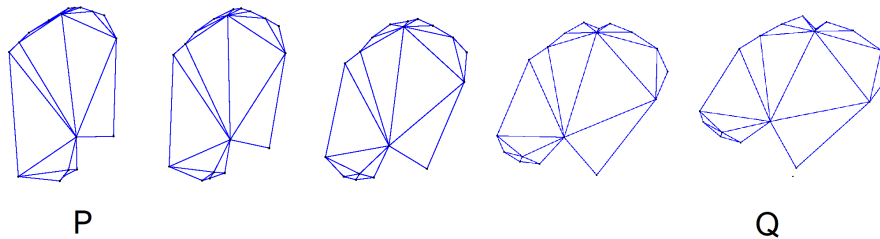
Figure 10 Interpolation created using the morphing algorithm proposed in [1, 2].

Each triangle in P is transformed into the corresponding triangle in Q. Because there is a prior correspondence between the vertices of P and Q, obtained using for example, a method based on feature points, the mapping between the features in P and Q, including concavities, is better than in the previous examples. In addition, the method tries to preserve the rigidity of the object. As a consequence, the moving region does not tend to inflate or deflate artificially during the interpolation.

Experiments were carried out to quantify the error associated with the interpolation method proposed in [1, 2], however, it was difficult to quantify the weight of the noise in the results. In some cases, errors caused by noise prevail over those of the interpolation method, and so, it was difficult to make convincing conclusions [5].

The algorithms used to create the examples based on morphing techniques [1, 2] and on the method proposed in [16, 15] were implemented by the authors of this paper, following the specifications presented in the original papers. Since we could not identify clearly the strategy used in [16, 15] to choose the point in the convex-hull of a moving object to which a concavity should converge to, we evaluated several alternatives. The results were slightly different for each choice, but, as they do not influence the discussion and the conclusions of this work, we present only the results of one of the alternatives. The examples concerning the methods proposed in [11, 12] were created using the tools in the virtual machine provided by the authors at the Secondo website[3].

---

[3] `http://dna.fernuni-hagen.de/secondo/`

4.2   Research issues

As we have seen in Section 2, the solutions proposed in the literature impose constraints on the movement and on the morphing of the objects to make the implementation of spatio-temporal operations easier. Removing these constraints, e.g., allowing moving segments to rotate and the shape of the region to change, simultaneously, is challenging. [12] proposes a method that allows moving segments to rotate, but handles only regions with fixed shape. The authors observe that, under these circumstances, the movement of the vertices of a moving region can be represented by trochoids and the curves traced by the segments can be represented by ravdoids. The parametric equations of these curves are then used to implement spatio-temporal operations. Figure 11 displays the curves (trochoids) defined by two vertices on the boundary of a moving object (the region in gray) whose center of rotation moves horizontally.
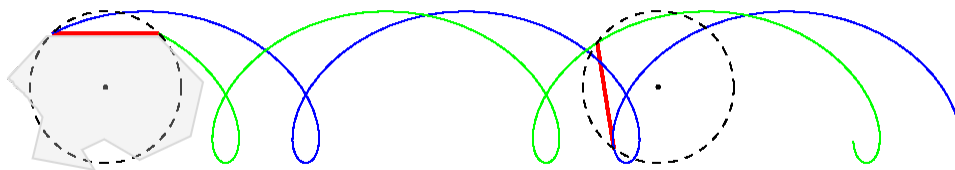
■ **Figure 11** Path of two vertices in the boundary of a rigid moving object.

Extending this idea to handle regions with non-fixed shape is challenging because the parametric equations that represent the paths traced by the vertices and the segments of the region are no longer known. The new resulting curves are exemplified in Figure 12.
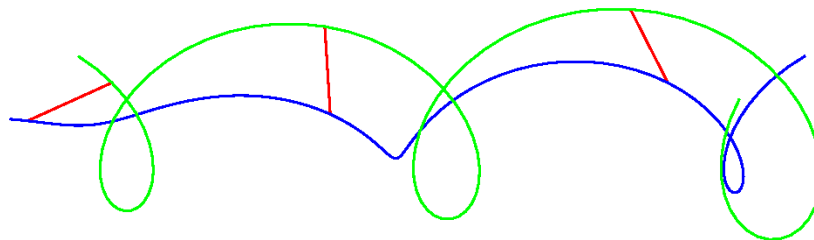
■ **Figure 12** Path of two vertices on the boundary of a deformable moving object.

Moreover, the path traced by the vertices and the segments of the moving regions with fixed shape is independent of the interpolation method used as long as it preserves the rigidity of the object during interpolation. This is no longer true when the moving region is deformable, because the trace of the vertices and the segments generated using two interpolation methods will probably differ. Whether or not these curves are related or there exists a family of curves that can be used in this case is a subject for further investigation. There are at least two paths for investigation here. Finding a general solution that is independent of the interpolation method used or finding a solution for a specific method. In both cases, after the parametric equations have been found, the implementation of spatio-temporal operations should follow similar strategies.

Ideally, we should find the parametric equations of these curves and use analytical methods to implement spatio-temporal operations. This means that it would be possible to obtain exact and computationally efficient solutions. An alternative is to use morphing techniques

to obtain the best possible representation of the evolution of the spatio-temporal phenomena and then use approximation functions, e.g., splines, to store that information. In this way, it would be possible to have simpler equations and the algorithms developed to implement spatio-temporal operations would be independent of the interpolation method used. However, the effect of the approximation on the characteristics of the interpolation, e.g., robustness and computational costs, must be minimized.

Another solution is to use numerical (iterative) methods to compute spatio-temporal operations. There are already solutions proposed in the literature, but this topic is not discussed in this paper.

4.3 Discussion

The examples above show that morphing techniques can give important insights and well-established solutions to the modeling and representation of spatio-temporal data in databases. However, the use of morphing techniques in databases raises new challenges that need investigation.

First, unlike the methods presented in Section 2, where rotation is represented implicitly (i.e., it is simulated by splitting line segments into parts that move linearly), morphing techniques allow rotation to be represented explicitly. This allows to obtain closer representations of the real evolution of the phenomena, but it also increases the complexity of the algorithms, because it becomes necessary to deal with curves. This problem is partially investigated in [12], but the proposed methods only apply to moving regions with fixed shape and the algorithms presented implement only a subset of the operations that should be provided by a spatio-temporal database. For instance, finding whether two rigid moving regions intersect is an open issue. On the other hand, using triangulation-based methods that allow decomposing complex shapes into triangles can make the development of algorithms easier.

Second, the focus of research in the spatio-temporal databases literature to solve the region interpolation problem is on creating efficient and robust interpolations, to ensure that the shape of the moving regions is always valid during the interpolation. Conservative approaches are followed, but the constraints imposed to the methods may cause unnatural deformations of the objects during the interpolation. This is an important issue when dealing with real-world data. On the other hand, morphing techniques are less conservative with respect to robustness and few studies exist on how to deal with complex transformations such as merging and splitting moving regions.

It is important to note that robustness is commonly evaluated using high complex examples, but no formal proof exist that the methods are robust. The proof that a method is not robust is usually given by counter-example. So, we also need to develop methods to detect invalid interpolations to ensure that all data in the database are valid.

Third, creating spatio-temporal data, such as moving regions (e.g., icebergs) and moving lines (e.g., the front-line of a forest fire), from raw data (e.g., an ordered sequence of images, videos or time-lapse videos) is difficult and time-consuming. Previous work is evaluated using synthetic data and so, it is difficult to conclude whether an interpolation represents well a given phenomenon, because one does not know how the shape should be between the source and the target observations. The alternative is to use a ground truth. For instance, given an ordered sequence of observations it is possible to create a sample (e.g., the observations in odd positions) and compare the results of the interpolation created using the sample with the 'observations in the even positions [18]. However, the creation of spatio-temporal data goes through several stages, for example, segmentation to extract the shape of the object from a sequence of images, simplification to reduce the number of vertices, polygon

matching to establish a correspondence between the source and target shapes, and finally, the interpolation of the shapes. Thus, the errors measured when running the interpolation algorithms may have been partially caused at the previous stages. Consequently, a ground truth and a benchmarking methodology are needed to measure and compare the accuracy of the interpolation methods.

5   Conclusion

This article deals with the representation of moving objects in spatio-temporal databases using continuous models in time and space, i.e., spatial data are represented in vector format, as well as their evolution over time. Despite advances in this field, there are open issues that need to be investigated so that spatio-temporal databases can be effectively used in real-world problems. A notable example is the region interpolation problem. [12] has recently proposed a new approach to interpolate moving regions of fixed shape, which is a step forward relatively to previous work. Nevertheless, the representation of general-purpose moving regions remains an open issue.

In this paper, we argue that there are topics studied in the field of morphing techniques that may help to solve important issues in spatio-temporal databases research. However, even though these techniques are widely used in visualization, their use in databases is unusual or nonexistent. Several challenges must be investigated to allow the use of morphing techniques in the context of databases:

(a) Robustness – additional constraints must be set to enforce data integrity in the database (there are constraints on the representation of spatial and spatio-temporal data that, in general, are not considered in animation and visualization);

(b) Operability – the use of new interpolation methods, data structures (e.g., triangle meshes) and models, requires developing new methods and computational geometry algorithms to implement spatio-temporal operations, such as, projection, distance and topological relationships;

(c) Optimization – the development of new data structures and algorithms may require the development of new optimization techniques to provide fast response times when querying large datasets;

(d) Context-awareness – there are static and dynamic factors that may affect the evolution of the spatio-temporal phenomena between observations (e.g., wind direction and speed, and the slope of the terrain affect the movement of the front-line of forest fires). Currently, the methods proposed in the literature to solve the region interpolation problem and the morphing techniques ignore these contextual factors.

The strategies presented in this article are being investigated in a research project financed by national funds, that started in 2018. The aim is to develop methods and tools to enable quantitative analysis of spatio-temporal data, guaranteeing levels of objectivity, precision and reproducibility compatible with the completion of scientific and engineering work.

Two case studies will be considered. The propagation of controlled forest fires and the morphological changes of living cells. The objective of the first is to compute the emissions of gases to the atmosphere using the representation of the propagation of the fire-line provided by a database. The data will be captured using drones and a meteorological station that will measure the gases emissions to the atmosphere. The objective of the second case study is to represent the continuous evolution of living cells in a database. The data will be captured using electronic microscopes and recorded as time-lapse videos. These case studies will provide real data and requirements that involve the modeling of several spatio-temporal

features. This is an important input to create a ground truth or a benchmark to evaluate the methods proposed in the spatio-temporal databases research community, thus addressing an important limitation found in previous work, where only synthetic data are used.

References

1 Marc Alexa, Daniel Cohen-Or, and David Levin. As-rigid-as-possible shape interpolation. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques - SIGGRAPH '00*, pages 157–164, 2000.

2 William Baxter, P Barla, and K Anjyo. Rigid shape interpolation using normal equations. In *NPAR '08 Proceedings of the 6th international symposium on Non-photorealistic animation and rendering*, pages 59–64, 2008.

3 Jan Chomicki, Yuguo Liu, and Peter Z. Revesz. Animating Spatiotemporal Constraint Databases. In *Proceedings of the International Workshop on Spatio-Temporal Database Management*, STDBM '99, pages 224–241, London, UK, 1999. Springer-Verlag.

4 J Cotelo Lema, Luca Forlizzi, Ralf Hartmut Guting, Enrico Nardelli, and Markus Schneider. Algorithms for Moving Objects Databases. *The Computer Journal*, 46(6):680–712, 2003.

5 José Duarte, Paulo Dias, and José Moreira. An Evaluation of Smoothing and Remeshing Techniques to Represent the Evolution of Real-World Phenomena. In *Advances in Visual Computing - 13th International Symposium, ISVC 2018, Las Vegas*, pages 57–67, 2018. `doi: 10.1007/978-3-030-03801-4_6`.

6 Luca Forlizzi, Ralf Haxtmut Giiting, Enrico Nardelli, and Markus Schneider. A Data Model and Data Structures for Moving Objects Databases. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pages 319–330, 2000.

7 Stephane Grumbach, Manolis Koubarakis, Philippe Rigaux, Michel Scholl, and Spiros Skiadopoulos. *Chapter 5: Spatio-temporal Models and Languages: An Approach Based on Constraints*, pages 177–201. Springer, September 2003.

8 Ralf Hartmut Güting, Thomas Behr, and Christian Düntgen. SECONDO : A Platform for Moving Objects Database Research and for Publishing and Integrating Research Implementations. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 33(2):56–63, 2010.

9 Ralf Hartmut Güting, Michael H Böhlen, Martin Erwig, Christian S Jensen, Nikos A Lorentzos, Markus Schneider, and Michalis Vazirgiannis. A Foundation for Representing and Querying Moving Objects. *ACM Trans. Database Systems*, 25(1):1–42, 2000.

10 S. Haesevoets and B. Kuijpers. Closure properties of classes of spatio-temporal objects under Boolean set operations. In *Proceedings Seventh International Workshop on Temporal Representation and Reasoning. TIME 2000*, pages 79–86, July 2000.

11 Florian Heinz and Ralf Hartmut Güting. Robust high-quality interpolation of regions to moving regions. *GeoInformatica*, 20(3):385–413, 2016.

12 Florian Heinz and Ralf Hartmut Güting. A data model for moving regions of fixed shape in databases. *International Journal of Geographical Information Science*, 32(9):1737–1769, 2018.

13 Ligang Liu, Guopu Wang, Bo Zhang, Baining Guo, and Heung-Yeung Shum. Perceptually based approach for planar shape morphing. In *Proceedings of the 12th Pacific Conference on Computer Graphics and Applications, 2004. PG 2004*, pages 111–120, October 2004.

14 Zhiguang Liu, Howard Leung, Liuyang Zhou, and Hubert P. H. Shum. High Quality Compatible Triangulations for 2D Shape Morphing. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, VRST '15, pages 85–94, New York, NY, USA, 2015. ACM. `doi:10.1145/2821592.2821594`.

15 Mark Mckenney and Roger Frye. Generating Moving Regions from Snapshots of Complex Regions. *ACM Trans. Spatial Algorithms Systems*, 1(1):1–30, 2015.

16   Mark Mckenney and James Webb. Extracting Moving Regions from Spatial Data. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 438–441, San Jose, 2010.

17   Q Meng, Z Fu, Y Huang, and M Shen. A FAST MATCHING APPROACH OF POLYGON FEATURES. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-2:45–49, July 2012.

18   J. Moreira, P. Dias, and P. Amaral. Representation of continuously changing data over time and space: Modeling the shape of spatiotemporal phenomena. In *Proceedings of the 2016 IEEE 12th International Conference on e-Science, e-Science 2016*, 2017.

19   Nikos Pelekis, Yannis Theodoridis, Spyros Vosinakis, and Themis Panayiotopoulos. Hermes - A Framework for Location-Based Data Managment. In *EDBT'06 Proceedings of the 10th international conference on Advances in Database Technology*, pages 1130–1134, Munich, 2006.

20   Majken K. Rasmussen, Esben W. Pedersen, Marianne G. Petersen, and Kasper Hornbæk. Shape-changing Interfaces: A Review of the Design Space and Open Research Questions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 735–744, New York, NY, USA, 2012. ACM.

21   Vitaly Surazhsky and Craig Gotsman. High quality compatible triangulations. *Engineering with Computers*, 20:147–156, 2002.

22   Erlend Tøssebro and Ralf Güting. Creating Representations for Continuously Moving Regions from Observations. In *Proceedings of the 7th International Symposium on Advances in Spatial and Temporal Databases*, pages 321–344. Springer-Verlag Berlin Heidelberg, 2001.

Towards a Qualitative Reasoning on Shape Change and Object Division

Momo Tosue
School of Science and Technology, Kwansei Gakuin University, 2-1, Gakuen, Sanda, 669-1337, Japan
eyi98098@kwansei.ac.jp

Kazuko Takahashi
School of Science and Technology, Kwansei Gakuin University, 2-1, Gakuen, Sanda, 669-1337, Japan
ktaka@kwansei.ac.jp

──── **Abstract** ────

We propose a qualitative representation for handling shape change and object division. We model the shape of a smooth curve in a two-dimensional plane together with its temporal change, using curvature extrema. The representation is based on Process-Grammar, which gives a causal account for each shape change. We introduce several rewriting rules to handle object division, that consist of making a tangent point, reconstruction, and separation. On the treatment of the division process, the expression can clarify the relative locations of multiple objects. We show formalization and application to represent a sequence of shape changes frequently observed in an organogenesis process.

1 Introduction

There are many examples of shape changes in dynamic systems. Usually simulation is applied using quantitative data to show the process of what happened or what will happen. An alternative way to represent shape change is to use algebraic formulas, for example, differential equations. However, it is difficult to imagine the shapes solely with differential equations and impossible to perform logical reasoning directly via algebraic formulation. We sometimes would like to establish the reasons that something happened to facilitate future predictions. For example, maybe one is interested in why a certain shape has been made or what will happen if a pair of objects become attached. These types of problems can be addressed using logical reasoning based on qualitative data to provide a symbolic description.

In biology and life sciences, division of an object and shape change in a single object are frequently observed. When two smooth curves (probably portions of the same closed circuit) contact each other at a point, we call the contact point *a tangent point*. In the process of division, the shape of an object gradually changes in such a way that a concave part is generated, a tangent point of the border is made, and then separation occurs at that point. Therefore, if we want to analyze such dynamic systems, we must first establish a better understanding of the underlying change mechanism.

There has been almost no symbolic treatment of shape change in the dynamic systems found in the life sciences. Tosue et al. proposed a symbolic expression to represent a

shape that enables reasoning about its temporal change for an organogenesis process. They approximated the shape using straight edges ignoring lengths and regarded the object as a polygon; transformation rules for the expression were described [12]. More specifically, they represented the border of a polygon using a sequence of rotation angles made by the subsequent edge. They defined shape changes as a set of rewriting rules on this expression, and presented an algorithm for drawing a figure corresponding to the expression [13].

However, it is difficult to perceive transformations using a coarse approximation such as straight edges, because the borders of objects are usually curved. If we wish to apply a more intuitive model using this method, we must use a more refined approximation, which complicates the rewriting rules and introduces high computational complexity.

In this study, we adopt a method that allows the curve to be represented qualitatively without using straight edges. The method is based on Process Grammar proposed by Leyton [7, 8], in which curvatures and extrema are used to represent the shape. Here, a curvature extremum is a part of the curve, where the curvature is at a maximum or minimum when tracing the boundary in a designated direction. Leyton considered that an extremum of a closed curve was formed gradually from a simple convex shape, and he aimed to infer the history of the construction of the shape. For example, the outline of an object in Figure 1(a) is changed to that in (b) by adding the force in the direction shown by the arrow; then to (c) if the force continues; and to (d) if the force diverges into two directions. Leyton formalized this transition as Process Grammar, which is a rewriting rule for symbols.

(a)   (b)   (c)   (d)
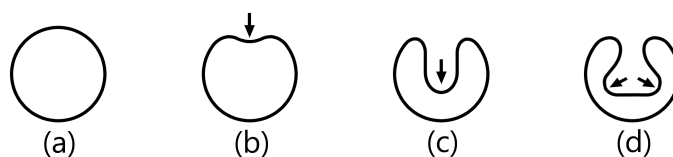
Figure 1 History of changing a shape.

Leyton's Process Grammar treats only a smooth curve that does not cross itself and has no cusps. Moreover, the division of an object was outside of his focus.

Here, we extend Leyton's representation to handle the division of an object. To this end, we define an expression of a shape that can discriminate (1) shapes, i.e., whether the curve has concavity and/or a tangent point, and (2) relative locations of objects, more specifically, whether an object is in the inner or outer part of another object. The second point is an essential factor in the treatment of multiple objects in two-dimensional planes (e.g., [15]).

The division process proceeds as follows. First, a border of a single curve extends to make a tangent point on itself. The connection is then reconstructed so that two closed curves are connected at the tangent point. Finally, the two closed curves are separated. A tangent point is made by connecting two points of a single curve. Therefore, the rewriting rules with respect to the division are defined over the entire expression, while the original Process Grammar is defined as rewriting a symbol locally.

The crucial point in the process of division is reconstruction. For example, in an organogenesis process, the borders of each object consist of a sequence of cells, and a certain force on the cells causes changes in the reconstruction. Here, we introduce a reconstruction rule to reflect such a phenomenon.

The remainder of this paper is organized as follows. In Section 2, we describe Leyton's Process Grammar. In Section 3, we introduce the description language for representing a shape. In Section 4, we define transformation rules for a shape change, and in Section 5, we

apply it to the transformations of objects in the organogenesis process. In Section 6, we discuss the extension of the proposed method and also compare our method to related works. Finally, in Section 7, we present our conclusions and future work.

2   Process-Grammar

Process-Grammar is a means of recovering the process history of a smooth shape from its curvature extrema, and expressing that evolution in terms of transitions at these extrema [7]. Here, a smooth curve never intersects itself and has no tangent point nor cusp. The target is the boundary of an object between the solid and the empty. A smooth line is represented by a sequence of curvature extrema, traveling along the curve so that the solid lies on the left side of the curve. Leyton showed that in a two-dimensional plane the evolution of any smooth shape of a smooth curve can be expressed in terms of six process transitions; he named this a "Process-Grammar." In Process-Grammar, a process is understood as creating the curvature extrema. It shows how the shapes form over time, and a direction of change of a curve is shown by an arrow to the curve in the figure. Here, we refer to the cause for the shape change as a "force."

There are four types of extrema curvatures: two maximum extrema M^+ and M^- and two minimum extrema m^+ and m^-. Each one shows how the shapes form over time: M^+ indicates a protrusion that is sharpening outwards, m^- indicates an indentation that is sharpening inwards, m^+ indicates a squashing that is flattening inwards, and M^- indicates an internal resistance that is flattening outwards. The polarity represents the convexity: "+" indicates a convex shape, while "−" is a concavity.

extremum type	explanation	force type	force direction	convexity
M^+	protrusion	sharpening	outwards	convex
m^-	indentation	sharpening	inwards	concave
m^+	squashing	flattening	inwards	convex
M^-	internal resistance	flattening	outwards	concave

A smooth curve in a two-dimensional plane is expressed as a sequence of these symbols. Figure 2 shows an example.
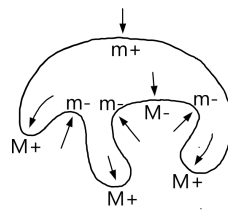
Figure 2 A figure and the corresponding expression.

A Process-Grammar is the transition rule over these sequences to represent changes in the shapes. There are two kinds of rules: continuation (the names of the rules begin with "C") and bifurcation (the names of the rules begin with "B") of the force at each extremum. Below we show the rules associated with the description of changes in the shape[1].

---

[1] In [7], the symbol "0" was used to represent an inflection point whose curvature is zero; we do not use it, because it can be deduced.

[Rules]   continuation and bifurcation

Cm^+:	$m^+ \to m^-$	(squashing continues until it indents)
CM^-:	$M^- \to M^+$	(resistance continues until it protrudes)
BM^+:	$M^+ \to M^+m^+M^+$	(shield formation)
Bm^-:	$m^- \to m^-M^-m^-$	(bay formation)
Bm^+:	$m^+ \to m^+M^+m^+$	(breaking through of a protrusion)
BM^-:	$M^- \to M^-m^-M^-$	(breaking through of an indentation)

For example, if protrusion (M^+) continues in the same direction, the shape of the extremum will become steeper, but its shape type does not change; if the force branches forward, then the extremum will move both to the left and right sides, and the original position will be flattened, which is formalized as BM^+ rule. The correspondence between these rules and shape changes are shown in Figure 3. In each figure, the arrow towards a curve indicates a force; the bold black arrow indicates the added force and the white arrow is a newly emerged force.
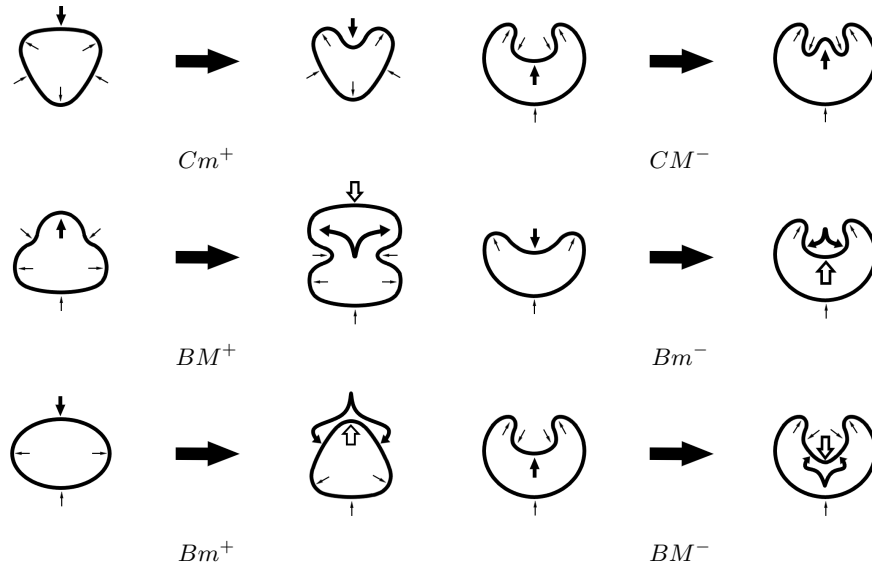
Figure 3 Process-Grammar defined by Leyton.

3   Description Language

We extend the Process Grammar formalism to describe the process of division. Our target figure is a set of smooth closed curves without an intersection. To simplify the problem, we first restrict the case in which there are at most two closed curves with at most one tangent point.

We introduce a description language $\mathcal{L}$ based on the Process Grammar. The language consists of two types of symbols: one with a dot and one without a dot

$\mathcal{L} = \{M^+, m^-, m^+, M^-, \dot{M}^+, \dot{m}^-, \dot{m}^+, \dot{M}^-\}$.

$\dot{M}^+, \dot{m}^-, \dot{m}^+$ and $\dot{M}^-$ denote that there exists a tangent point on the extrema M^+, m^-, m^+ and M^-, respectively. We call the symbols $\dot{M}^+, \dot{m}^-, \dot{m}^+$, and $\dot{M}^-$, *dotted elements*. We also call $M^+, M^-, \dot{M}^+$ and $\dot{M}^-$, *M-elements*, and $m^+, m^-, \dot{m}^+$ and $\dot{m}^-$, *m-elements*, respectively.

An expression is a finite sequence of elements in $\mathcal{L}$. For example, the expression for the figure in Figure 4 is $M^+\dot{m}^-M^+m^+M^+\dot{m}^-M^+m^+$
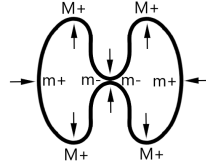
■ **Figure 4** A single curve with a tangent point.

An expression for a single closed curve is cyclic, that is, expressions $e_ie_{i+1}\ldots e_ne_1\ldots e_{i-1}$, for all i $(1 \leq i \leq n)$ show the same shape. For example, an expression for a simple oval in Figure 5(a) is represented either as $M^+m^+M^+m^+$ or $m^+M^+m^+M^+$. If we use more elements to represent a closed curve, then we can express the shape in a more refined manner (Figure 5(a)(b)).

(a) (b)
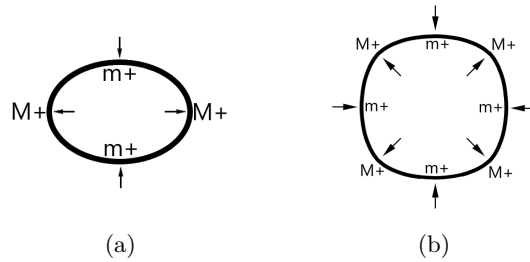
■ **Figure 5** Simple ovals.

Let $\mathcal{E}$ be a set of expressions
$\mathcal{E} = \{e_1e_2\ldots e_n \mid e_i \in \mathcal{L} \ (1 \leq i \leq n)\}$.
We define an inverse function on $\mathcal{E}$ as follows:

$inv(M^+) = m^-, \quad inv(\dot{M}^+) = \dot{m}^-,$
$inv(m^+) = M^-, \quad inv(\dot{m}^+) = \dot{M}^-,$
$inv(M^-) = m^+, \quad inv(\dot{M}^-) = \dot{m}^+,$
$inv(m^-) = M^+, \quad inv(\dot{m}^-) = \dot{M}^+,$
$inv(e_1e_2\cdots e_n) = inv(e_n)\cdots inv(e_2)inv(e_1)$ for $e_1e_2\cdots e_n \in \mathcal{E}$ $(1 \leq i \leq n)$.

Let $\dot{s}$ be an expression that includes exactly one dotted element. Then, s is the expression obtained by replacing the dotted element in $\dot{s}$ by the corresponding non-dotted element. That is, for $\dot{s} = e_1\cdots e_n$, there $\exists 1i$ $(1 \leq i \leq n)$ such that $e_i = \dot{e}$ where $e = M^+, m^-, m^+, M^-$, s denotes $e_1\cdots e_{i-1}ee_{i+1}\cdots e_n$.

An expression for a smooth closed curve satisfies the following conditions (C1) and (C2).

(C1) For $e_1\ldots e_n \in \mathcal{E}$, n is more than three.
(C2) For $e_1\ldots e_n \in \mathcal{E}$, if e_i is an M/m-element, then e_{i+1} is an m/M-element for all i $(1 \leq i \leq n, e_{n+1} = e_1)$.

The first condition requires at least four extrema to form a closed curve in a two-dimensional plane, according to the four-vertex theorem in differential geometry (e.g., [4]). The second condition requires that both M-element and m-element appear in turn, which is critical for smooth curve formation. Specifically, it indicates that there are no cusps between tangent points and guarantees the balance of inward and outward forces.

If there are two closed curves, then we can combine the expressions for each curve. The *combined expression* is either in the form σ, $\sigma \,||\, \tau$ or $\sigma[\tau]$, where σ and τ are expressions that satisfy the above conditions (C1) and (C2). In case it is in the form of σ, then it includes either a no-dotted element or two dotted elements that are not next to each other. In case it is in the form of $\sigma \,||\, \tau$ or $\sigma[\tau]$, both σ and τ have exactly one dotted element. $\sigma \,||\, \tau$ shows that τ is located in the external part of σ, and $\sigma[\tau]$ shows that the closed curve τ is located in the inner part of σ. In the latter case, σ has a hole τ in its inner part. $\sigma \,||\, \tau$ and $\tau \,||\, \sigma$ show the same figure.

This representation can be used to discriminate between the location of closed curves and also the existence of tangent points. We show several simple combined expressions for the shapes shown in Figure 6.

(a) $M^+m^+M^+m^+ \,||\, M^+m^+M^+m^+$
(b) $M^+m^+M^+m^+[M^-m^-M^-m^-]$
(c) $M^+m^+\dot{M}^+m^+ \,||\, M^+m^+\dot{M}^+m^+$
(d) $M^+m^+M^+\dot{m}^+[M^-\dot{m}^-M^-m^-]$

(a)   (b)   (c)   (d)
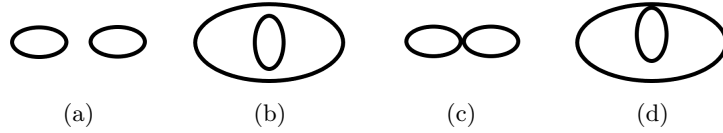
Figure 6 Combined expressions.

There are two closed curves. In cases (a) and (b), they are disconnected; in case (c), they are externally connected; and in case (d), they are internally connected. The tangent point is represented by the dotted expressions. Moreover, in cases (b) and (d), one is inside of the other.

4 Transition System

In addition to continuation and bifurcation rules, we introduce several rewriting rules over the description language to formalize object division: making a tangent point, reconstructing closed curves, and separation.

4.1 Making a tangent point

A tangent point is made by connecting a pair of extrema, which have grown by receiving a force. For example, an extremum m^- grows to reach another extremum m^+, then a tangent point is made both at m^- and m^+. The type of connection is either internal or external depending on the direction of the added force. Only four pairs have the possibility to make a tangent point.

1. internal connection
   A pair of extrema has received forces inward and at least one of them is concave. The pair satisfying this condition is either a pair of m^- and m^-, or a pair of m^- and m^+.
2. external connection
   A pair of extrema has received forces outward and at least one of them is convex. The pair satisfying this condition is either a pair of M^+ and M^+, or a pair of M^+ and M^-.

The transition rule for each pair is as follows.

[Rules]   making a tangent point

$$TH : sm^-tm^- \to s\dot{m}^-t\dot{m}^-$$
$$TU : sm^+tm^- \to s\dot{m}^+t\dot{m}^-$$
$$TO : sM^+tM^+ \to s\dot{M}^+t\dot{M}^+$$
$$TP : sM^+tM^- \to s\dot{M}^+t\dot{M}^-$$

where $s, t \in \mathcal{E}$ that satisfy (C2), and t contains at least one M^+ and m^- in the rules TH and TO, respectively.

4.2   Reconstruction

Reconstruction is a crucial part of the division process.

When we deal with an alveolus whose boundary is a sequence of cells, a pair of the sequences reconnect with each other in the reconstruction process. Actually, this occurs within a thick boundary at the tangent point. Here, we make a model in which the structure of the boundary is reconstructed.

In reconstruction, the sequences of the extrema located around a tangent point are decomposed and connected differently with new pairs of extrema.

On tracing a boundary which has a tangent point, we find two smooth curves encountered at the tangent point. Considering the directions of these curves on passing the tangent point, there are only two possibilities shown in Figure 7, since there is a constraint that a boundary never crosses. The reconstruction is the process of changing from (a) to (b) or from (b) to (a) in Figure 7. Therefore, we get four types of reconstructions, each of which corresponds to a type of tangent point; type P is divided into P_l and P_r, which are symmetric.

(a)   (b)

Figure 7 Directions of curves on tracing a boundary.

[Rules]   reconstruction

$$RH : s\dot{m}^-t\dot{m}^- \to s\dot{m}^+ \,||\, t\dot{m}^+ \qquad \text{(Type } H)$$
$$RU : s\dot{m}^+t\dot{m}^- \to sm^+\dot{M}^+m^+ \,||\, tm^+\dot{M}^+m^+ \qquad \text{(Type } U)$$
$$RO : s\dot{M}^+t\dot{M}^+ \to s\dot{M}^-[t\dot{M}^-] \qquad \text{(Type } O)$$
$$RP_l : s\dot{M}^-t\dot{M}^+ \to sM^-\dot{m}^-M^-[tM^-\dot{m}^-M^-] \qquad \text{(Type } P_l)$$
$$RP_r : s\dot{M}^+t\dot{M}^- \to sM^-\dot{m}^-M^-[tM^-\dot{m}^-M^-] \qquad \text{(Type } P_r)$$

where $s, t \in \mathcal{E}$ that satisfy (C2), $|s|, |t| \geq 3$, and s is the expression for the outer curve in the rules RO, RP_l, and RP_r.

The symbols "H," "U," "O," and "P" used in the names of the rules are based on the entire shape of an object when a tangent point is made.

The constraint on the length of the expressions s and t is applied to obtain a combined expression that satisfies the conditions (C1) and (C2) after reconstruction. If this constraint is not satisfied, then the transformation process should transit to an intermediate state by applying bifurcation rules (BM^+, Bm^-, Bm^+ and BM^-) before applying the reconstruction rule. The other constraint is for distinguishing the locations of curves.
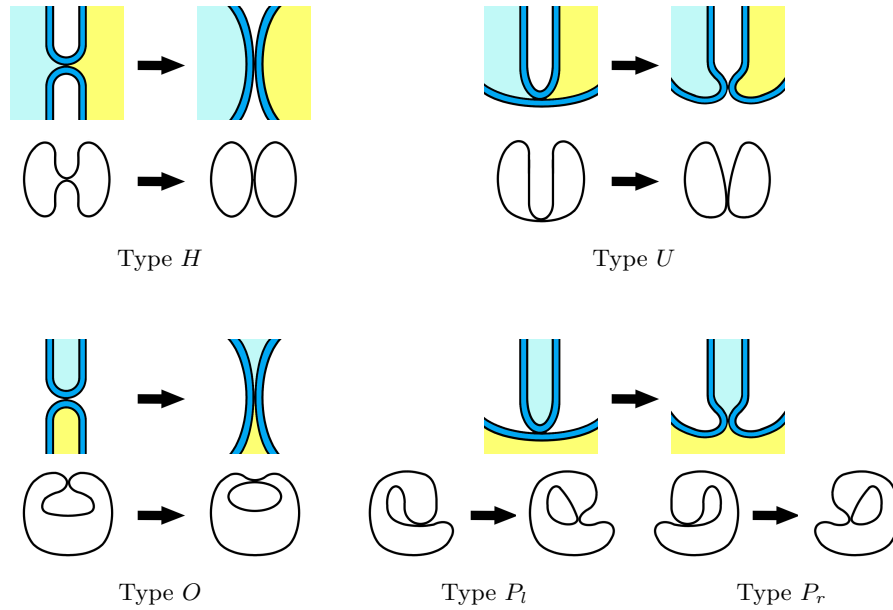
Figure 8 Reconstruction for each type.

Figure 8 illustrates the reconstruction for each type. The details of the neighbor to the tangent point are shown at the top, and an example of the shape of an object is shown at the bottom for each type. The shaded area indicates a solid part, that is, the inside of the object.

The rules of types H and U are the changes in which a part(s) of the border is extended inward to connect itself from the inside; as a result, two externally connected, closed curves are obtained. The rules of type O and P are the changes in which a part(s) of the border is extended to connect itself from the outside; as a result, a hole is made inside.

For example, in type H, two extrema m^- approach to make a tangent point (left side of the figure); then, the directions of the forces are changed to m^+ (right side).

4.3 Separation

The rules for separation are simple. We separate the two closed curves by removing the tangent point.

[Rules]   separation

$SE : \dot{s_1} \, || \, \dot{s_2} \to s_1 \, || \, s_2$
$SI : \dot{s_1}[\dot{s_2}] \to s_1[s_2]$

where $s_1, s_2 \in \mathcal{E}$ that satisfy (C1) and (C2).

5 Application of rules for division processes

We show an application of four types of rules for the division process that frequently appear in an organogenesis process [14].

We start with a convex shape, whose expression is $M^+m^+M^+m^+M^+m^+M^+m^+$. We set this shape as $S_0{}^2$.

---

[2] We follow the Leyton viewpoint that a pure circle cannot be represented as a process. We take a simple

5.1 Type H

In this case, starting from S_0, two protrusions are made, come near, make an internal tangent point, and then the object is separated into two pieces (Figure 9).

$S_0 : M^+m^+M^+m^+M^+m^+M^+m^+$

$\quad \Downarrow \quad (Cm^+)$

$S_1 : M^+m^-M^+m^+M^+m^-M^+m^+$

$\quad \Downarrow \quad (TH)$

$S_2 : M^+\dot{m^-}M^+m^+M^+\dot{m^-}M^+m^+$

$\quad \Downarrow \quad (RH)$

$S_3 : M^+\dot{m^+}M^+m^+ \,||\, M^+m^+M^+\dot{m^+}$

$\quad \Downarrow \quad (SE)$

$S_4 : M^+m^+M^+m^+ \,||\, M^+m^+M^+m^+$
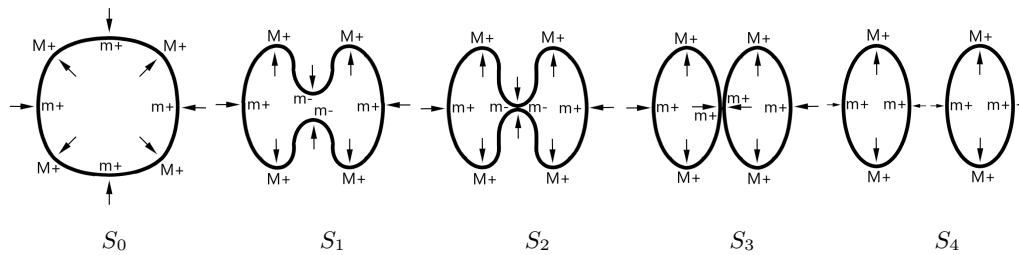
Figure 9 Division process for Type H.

5.2 Type U

In this case, starting from S_0, one indentation arises and reaches another part of the border, makes an internal tangent point, and then the object is separated into two pieces (Figure 10).

$S_0 : M^+m^+M^+m^+M^+m^+M^+m^+$

$\quad \Downarrow \quad (Cm^+)$

$S_5 : M^+m^+M^+m^+M^+m^-M^+m^+$

$\quad \Downarrow \quad (TU)$

$S_6 : M^+\dot{m^+}M^+m^+M^+\dot{m^-}M^+m^+$

$\quad \Downarrow \quad (RU)$

$S_7 : M^+m^+\dot{M^+}m^+M^+m^+ \,||\, \dot{M^+}m^+M^+m^+M^+m^+$

$\quad \Downarrow \quad (SE)$

$S_8 : M^+m^+M^+m^+M^+m^+ \,||\, M^+m^+M^+m^+M^+m^+$

5.3 Type O

In this case, the same change as that of type U occurs before S_5. After S_5, the protrusion branches and extends to connect together, and makes an external tangent point, and then the curve is separated into two closed curves, one of which is enclosed by the other (Figure 11).

$S_0 : M^+m^+M^+m^+M^+m^+M^+m^+$

$\quad \Downarrow \quad (Cm^+)$

$S_5 : M^+m^+M^+m^+M^+m^-M^+m^+$

oval as an initial state and apply Bm^+ to get S_0.
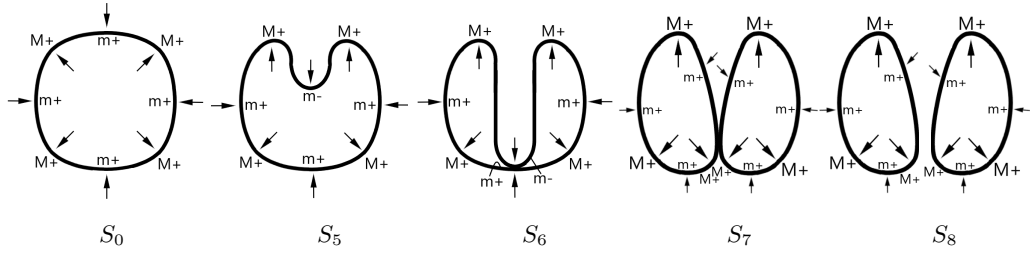
Figure 10 Division process for Type U.

$$\Downarrow \quad (Bm^-)$$
$$S_9 : M^+m^+M^+m^+M^+m^-M^-m^-M^+m^+$$
$$\Downarrow \quad (TO)$$
$$S_{10} : M^+m^+M^+m^+\dot{M}^+m^-M^-m^-\dot{M}^+m^+$$
$$\Downarrow \quad (RO)$$
$$S_{11} : M^+m^+M^+m^+M^+m^+\dot{M}^-m^+M^+m^+[M^-m^-\dot{M}^-m^-]$$
$$\Downarrow \quad (SI)$$
$$S_{12} : M^+m^+M^+m^+M^+m^+M^-m^+M^+m^+[M^-m^-M^-m^-]$$
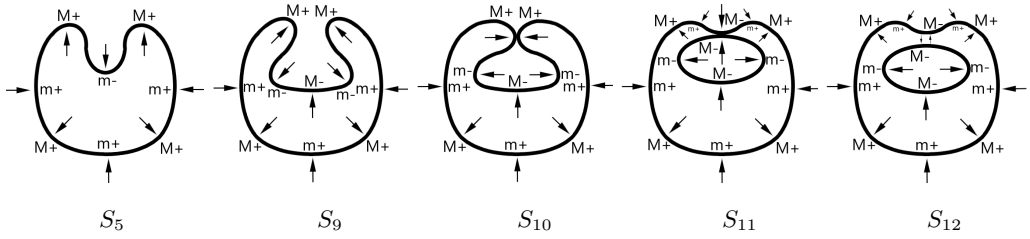
Figure 11 Division process for Type O.

5.4   Type P_l

In this case, the same change as that of type O occurs before S_9. After S_9, only one protrusion extends and bends to reach another part of the border; this makes an external tangent point, and then the curve is separated into two closed curves, one of which is enclosed of by the other (Figure 12). The process is similarly described for type P_r.

$$S_0 : M^+m^+M^+m^+M^+m^+M^+m^+$$
$$\Downarrow \quad (Cm^+)$$
$$S_5 : M^+m^+M^+m^+M^+m^-M^+m^+$$
$$\Downarrow \quad (Bm^-)$$
$$S_9 : M^+m^+M^+m^+M^+m^-M^-m^-M^+m^+$$
$$\Downarrow \quad (BM^+)$$
$$S_{13} : M^+m^+M^+m^+M^-m^-M^-m^-M^+m^+M^+m^+M^+m^+$$
$$\Downarrow \quad (TP)$$
$$S_{14} : M^+m^+M^+m^+\dot{M}^-m^-M^-m^-\dot{M}^+m^+M^+m^+M^+m^+$$
$$\Downarrow \quad (RP_l)$$
$$S_{15} : m^+M^+m^+M^+m^+M^+m^+M^+m^+M^-\dot{m}^-M^-[m^-M^-m^-M^-\dot{m}^-M^-]$$
$$\Downarrow \quad (SI)$$
$$S_{16} : m^+M^+m^+M^+m^+M^+m^+M^+m^+M^-m^-M^-[m^-M^-m^-M^-m^-M^-]$$

$$S_9 \qquad S_{13} \qquad S_{14} \qquad S_{15} \qquad S_{16}$$
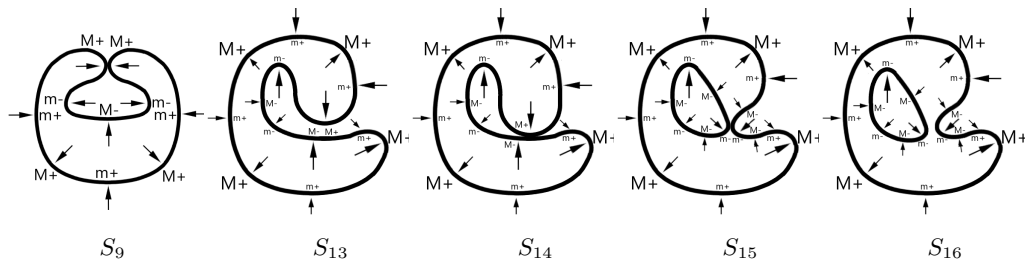
Figure 12 Division process for Type P_l.

Note that after separation in cases of type O and type P_l, one closed curve becomes a hole, that is, the inner side of the hole is an outside of the original solid object. Therefore, in case of type O, the expression of the inner closed curve is $M^-m^-M^-m^-$, which is equivalent to $inv(m^+M^+m^+M^+)$.

6 Discussion

6.1 Generalization

In the previous sections, we restricted the target figure to one with at most two closed curves and at most one tangent point. We can drop these two restrictions easily.

We can represent multiple closed curves in any location by defining a combined expression recursively: let $\mathcal{E}_C$ be a set of combined expressions; the combined expression is defined either in the form of σ, $\sigma \,||\, \tau$ or $\sigma[\tau]$ where $\sigma, \tau \in \mathcal{E}_C$ that satisfy the conditions (C1) and (C2), respectively. We can represent multiple tangent points by adding elements to a description language: "numbered-dotted element" is used instead of dotted element to discriminate each tangent point because tangent points do not affect one another. For example, Figure 13 shows one possible division process in the case of two tangent points, and the following is the corresponding reconstruction rule, in which two tangent points are represented using a single dot and a double dot, respectively.

[Rules]   reconstruction for double tangent points

$RU_2 : s\dot{m}^+t\ddot{m}^+[u\ddot{m}^-v\dot{m}^-] \to sm^+\dot{M}^+m^+um^+\ddot{M}^+m^+ \,||\, tm^+\ddot{M}^+m^+vm^+\dot{M}^+m^+$
where $s, t, u, v \in \mathcal{E}_c$ that satisfy (C2), $|s|, |t|, |u|, |v| \geq 3$.

In this figure, the inner circuit is expanded to reach the border of the outer circuit at two distinct points, and two tangent points are generated (Figure 13(b)). Next, the reconstruction occurs at these tangent points, respectively and as a result, two new curves are generated that are externally connected (Figure 13(c)). The object is then separated into two pieces (Figure 13(d)).

6.2 Extension

So far, we have discussed shape change starting from a simple convex form in the direction in which concave parts are created. Moreover, we have not considered shape change after separation. Then, the following question arises: if an object has a concave part after separation, how does this affect the shape change? The shape may change similarly with the process before the separation; however, it may change to recover the convex form.
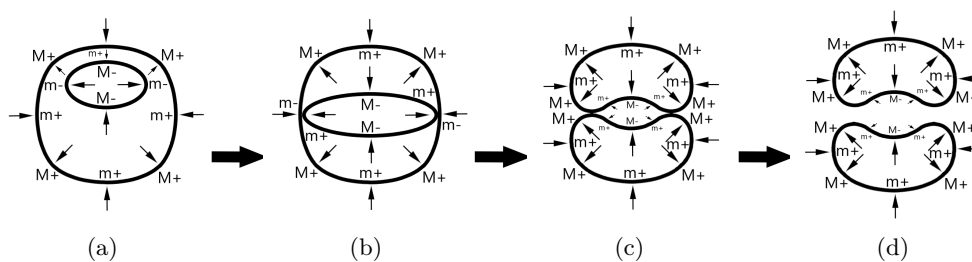
Figure 13 Division process in case of double tangent points.

Figure 14 Stabilization rules.

To address this issue, first, we introduce the concept of a stable state. When an expression consists of only M^+ and m^+, we call it *stable*. It can be considered that a stable curve changes by receiving some force, and an unstable curve likes to change to become stable. To treat this possibility, we need to allow an application of the rules introduced so far in the opposite direction. As such, the following rules are required.

[Rules]   stabilization

$$Stm^+ : m^+M^-m^+ \to m^+$$
$$StM^- : M^-m^+M^- \to M^-$$

The first rule shows that if the concave part of the curve is pressed continuously from the inside then this part vanishes. The second rule shows that if the concave part of the curve is pressed continuously from the outside, then the protrusion vanishes (Figure 14). In each figure, the bold black arrow indicates the changing force, and the white arrow is the force that vanishes.

Using all of these rules together, we generate typical shape changes that appear during the organogenesis process (Figure 15).

6.3   Related works

Generally, it is difficult to represent the shape of an object qualitatively compared to other spatial features such as the relative positional relationships and relative directions. The most popular approach is to divide the boundary of an object into segments and represent its shape as a sequence of attributes such as length, direction, curvature, and so on that are attached to each segment. In these methods, the more attributes each segment has, the more accurately a figure can be drawn. This also requires more reasoning rules to interpret more complicated data.

Museros et al. introduced a qualitative shape descriptor (QSD) of each boundary using length, angle, curvature and so on [2, 5]. They also extended this scheme to a juxtaposition of objects in point-point, point-line, or line-point connecting types. They defined this
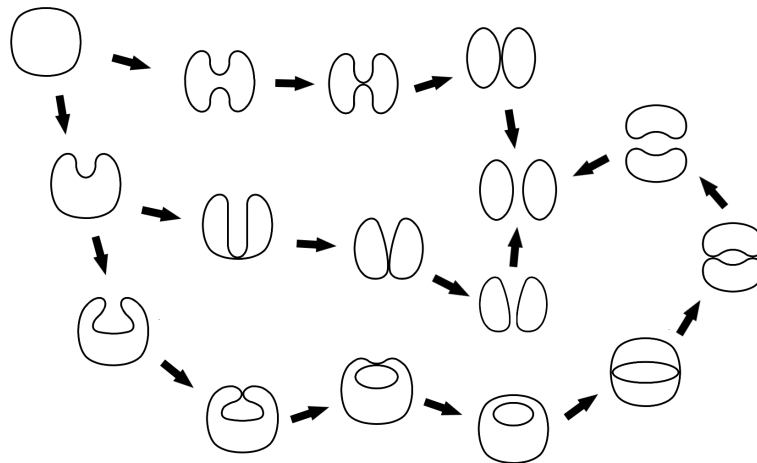
Figure 15 Application for explaining an organogenesis process.

juxtaposition as a shape composition that derives a new shape by this operation [1, 9] and treats rigid objects with non-deformable boundaries. Their focus was to provide a qualitative description of an object and formalize their composition, whereas we describe the change in the shape of an object with deformable boundaries using rewriting rules.

Galton et al. proposed a grammar scheme to describe changes in shapes, including a cusp [6]. Unlike QSD, they addressed deformable boundaries. They did not use extrema but rather local shape patterns to represent a closed curve; additionally, they created a number of transition rules by enumerating possible local changes. However, they did not describe the reason for the change. In contrast, here, we consider the forces involved in deforming the boundaries. Moreover, they did not address tangent points nor the division of an object, whereas we address these aspects.

Cohn used a mereotopological approach to formalize the shape of an object. He proposed a qualitative representation of a concave region using predicates [3]. Various shapes can be distinguished by representing relative position, size, and the direction of concave parts in a refined manner. He also discussed the continuous shape change. Because the number of possible shape descriptions is generally unbounded, he showed an example of a possible continuous transformation under some restricted forms. In contrast to an approach using rewriting rules, it is difficult to define the continuous transformation in the logical framework, and no formal explanation was given regarding this transformation.

In some research activities, shape is represented as a sequence of symbols, and its change is formalized as a set of rewriting rules. Shape grammar is a set of rules applied to an initial shape to generate designs [11]. It is mainly applied to show the structure of architectures. As the rules are defined to transform the initial shapes, the user may decide which rule can be used to achieve the desired outcome. Leyton's Process Grammar uses a set of rewriting rules. It can be considered as an abstract rewriting system [16, 10]. The main reason for our choice of Process Grammar is that it is suitable for resolving dynamic changes in a curve, as the history of shape change can be explained in terms of forces applied to the curve.

The biggest difference between our work and previous works is that our method can address the division of an object. We have defined a language and transition rules to handle the reconstruction of closed curves and the locations of multiple closed curves, which are the main issues involved in the treatment of a division frequently observed in an organogenesis process.

7   Conclusion

We have proposed a system to handle qualitative shape change, using the curvature and extrema of the curve. The proposed system enables the representation of a transformation qualitatively, including the division of an object, and gives a causal account for each transformation.

Our method has the following main features:

- direct representation of a smooth curve, as opposed to using an approximation such as a polygon,
- the ability to accommodate a tangent point and a division process, and
- the ability to describe the relative positional relationships of multiple closed curves.

Our approach can be applied to shape changes in various fields such as an alveolar division in a life science, analysis of a tumor in immunology, change in terrain shape in geomorphology, and so on.

As a future work, we would like to prove the completeness of this transition system, that is, the set of expressions that cover all possible transformations. It may be suitable to make several distinct models, including a conceptual, a theoretical, and a realistic model. We are currently looking into the rules necessary for describing these possible transitions more precisely, depending on the model.

References

1 Lledó Museros Cabedo, Luis González Abril, Francisco Velasco Morente, and Zoe Falomir. A Pragmatic Qualitative Approach for Juxtaposing Shapes. *Journal of Universal Computer Science*, 16(11):1410–1424, 2010.

2 Lledó Museros Cabero and M. Teresa Escrig. A Qualitative Theory for Shape Representation and Matching for Design. *Proceedings of the Sixteenth Eureopean Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004*, pages 858–862, 2004.

3 Anthony Cohn. A hierarchical representation of qualitative shape based on connection and convexity. *Spatial Information Theory. A Theoretical Basis for GIS, COSIT'95*, pages 311–326, 1995.

4 Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Inc., 1976.

5 Zoe Falomir, Luis González Abril, Lledó Museros Cabedo, and Juan Antonio Ortega. Measures of similarity between objects based on qualitative shape descriptions. *Spatial Cognition & Computation*, 13(3):181–218, 2013.

6 Antony Galton and Richard Meathrel. Qualitative outline theory. *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 1061–1066, 1999.

7 Michael Leyton. A process-grammar for shape. *Artificial Intelligence*, 34:213–247, 1988.

8 Michael Leyton. Inferring causal-history from shape. *Cognitive Science*, 13:357–387, 1989.

9 Albert Pich and Zoe Falomir. Logical Composition of Qualitative Shapes Applied to Solve Spatial Reasoning Tests. *Cognitive Systems Research*, 52:82–102, 2018.

10 David Sabel and Hans Zantema. Transforming Cycle Rewriting into String Rewriting. *Leibniz International Proceedings in Informatics*, 36:285–300, 2015.

11 George Stiny. Introduction to shape and shape grammars. *Environment and Planning B: Planning and Design*, 7(3):343–351, 1980.

12 Momo Tosue, Sosuke Moriguchi, and Kazuko Takahashi. Qualitative Shape Representation and Reasoning Based on Concavity and Tangent Point. In *Proceedings of the 31st International Workshop on Qualitative Reasoning*, 2018.

13 Momo Tosue, Sosuke Moriguchi, and Kazuko Takahashi. Operations for Shape Transformations Based on Angles. In *Proceedings of the 11th International Conference on Agents and Artificial Intelligence*, to appear.

14 Lewis Wolpert and Cheryll Tickle. *Principles of Development*. OUP Oxford, 2011.

15 Michael Worboys. Using maptrees to characterize topological change. In *Spatial Information Theory. 11th International Conference, COSIT 2013*, pages 74–90, 2013.

16 Hans Zantema, Barbara Koenig, and H. J. Sarnder Brugging. Termination of cycle rewriting. *Proceedings of Joint International Conference, RTA-TLCA 2014*, pages 476–490, 2014.

Reproducible Research in Geoinformatics: Concepts, Challenges and Benefits

Christian Kray
Institute for Geoinformatics (ifgi), University of Münster, Germany
c.kray@uni-muenster.de

Edzer Pebesma
Institute for Geoinformatics (ifgi), University of Münster, Germany
edzer.pebesma@uni-muenster.de

Markus Konkol
Institute for Geoinformatics (ifgi), University of Münster, Germany
m_konk01@uni-muenster.de

Daniel Nüst
Institute for Geoinformatics (ifgi), University of Münster, Germany
daniel.nuest@uni-muenster.de

Abstract

Geoinformatics deals with spatial and temporal information and its analysis. Research in this field often follows established practices of first developing computational solutions for specific spatiotemporal problems and then publishing the results and insights in a (static) paper, e.g. as a PDF. Not every detail can be included in such a paper, and particularly, the complete set of computational steps are frequently left out. While this approach conveys key knowledge to other researchers it makes it difficult to effectively re-use and reproduce the reported results. In this vision paper, we propose an alternative approach to carry out and report research in Geoinformatics. It is based on (computational) reproducibility, promises to make re-use and reproduction more effective, and creates new opportunities for further research. We report on experiences with executable research compendia (ERCs) as alternatives to classic publications in Geoinformatics, and we discuss how ERCs combined with a supporting research infrastructure can transform how we do research in Geoinformatics. We point out which challenges this idea entails and what new research opportunities emerge, in particular for the COSIT community.

1 Introduction

Spatial and temporal information and their analysis play a central role in many scientific disciplines such as Geography or Economics and are essential in understanding and solving many pressing societal issues. Regardless of which disciplinary perspective is taken, the way in which the corresponding research is carried out and reported is very similar. On an abstract level, a researcher will work on a specific issue or problem that has a spatiotemporal component, for example, computing a route that meets certain criteria, or checking whether

a number of spatiotemporal constraints can be fulfilled given a set of moving objects. She might develop a representation to computationally describe the problem and/or an algorithm/approach that can be applied to the problem in order to answer the research question. These might be entirely new or derived from existing ones. Or, she might carry out some studies to explore spatiotemporal aspects in the real world (e.g. how humans perform when given specific navigational instructions) and then derive some potentially generalisable insights (e.g. performance varies according to the length of instructions). She will then report her findings in a publication, i.e. a static document consisting of text, tables, and figures that concisely describe what she found out, how she went about her investigation, and (ideally) how others can independently repeat the steps she took to confirm her findings. Once peer-reviewed and published, others will use her results for further research.

This overall approach is very similar to many other scientific disciplines, and has existed in this form for a long time. While generally well established and effective, there are a number of key shortcomings inherent to this process. Firstly, it is usually not possible to actually or easily reproduce results as the required data, analysis procedures (e.g. source code), and the necessary configuration to re-run the analysis (e.g. compiler version) on the data are not available. Even when they are, the effort of reproducing the results can be prohibitive, e.g. due to having to restore the exact configuration of the computational environment that was used by the author of the paper. This configuration may come with little documentation, e.g. since it evolved over a long time period in a trial-and-error fashion. If this information is available at all, it can still involve issues such as having to retrieve and install outdated libraries, which also hinders re-use [15]. Secondly, the paper is usually provided as unstructured text for human readers that makes a deeper (computational) analysis difficult. For example, finding all papers that investigate the same spatial region or use the same spatiotemporal calculus is not trivial. In addition, this lack of semantic structure prevents presenting results in different ways that might be better suited for different (non-academic) audiences or purposes. Thirdly, with a static document it is very difficult for the reader or reviewer to explore the results more deeply. For example, a reader cannot easily check whether the reported results are robust (e.g. when parameters are slightly modified or a particular axiom is weakened), or whether they are consistent with previously reported results that might have used a different dataset. These issues apply for authors as well as readers and reviewers.

In this paper, we present several concepts and ideas for overcoming these issues for research which deals with spatiotemporal information and its computational analysis. After first reviewing the status-quo of how research in Geoinformatics is done currently, we describe the concept of executable research compendia (ERCs), which addresses some of the issues outlined above. Based on these concepts, we then propose our vision for how computational research on spatiotemporal aspects can be carried out in the future and propose an open research infrastructure for Geoinformatics (OpenRIG). This approach comes with a number of key challenges and opportunities, which we discuss subsequently. The paper concludes by outlining limitations and summarising our key contributions.

2 Background

Spatial and temporal information is essential when dealing with a variety of problems. In the past decades, several scientific sub-disciplines have emerged which tackle these kinds of problems across traditional disciplinary boundaries and often act as an integrator, e.g. between computer scientists and geologists. *GI Science* can be defined in different ways (cf. [8] for a short review of definitions). The research conducted under this label stretches from

semantic interoperability and ontologies, theories of spatial-temporal information systems, and spatial data modelling and services, to user-generated data [26]. Donoho critically discusses *Data Science* as a scientific field in its own right, which concerns not only the extraction of information from data, but "each and every step that the professional must take, from getting acquainted with the data all the way to delivering results based upon it, and extending even to that professional's continual review of the evidence about best practices of the whole field itself" [7, p. 794]. Spatiotemporal data is one important type of data covered by Data Science, with *Spatial Data Science* being an emerging subfield focusing on this. We define *Geoinformatics* as the discipline that generally deals with spatial and temporal information computationally (e.g. geostatistics, geosimulation, geovisualisation, interaction with spatiotemporal information). The definitions listed above overlap and describe similar disciplines that each take a different perspective. Whereas GI Science emphasises the geographic aspect and the use of geographic information systems (GIS), Data Science focuses on data and its properties, while Geoinformatics more generally considers spatial and temporal information and its analysis and use.

Besides investigating spatiotemporal aspects, the scientific disciplines introduced above share another commonality: the way in which a substantial part of research is carried out and reported today. After inception of an idea, scientists formulate a hypothesis or research questions, review existing work on the topic, gather data and information, and analyse them. At some point, researchers author a scholarly publication that describes their methods and results and then submit it to an academic outlet, where it undergoes peer-review. If accepted, a document is created by the publisher and made available to readers of the outlet. This process today mostly takes place in the digital realm – frequently including the research work itself [12]. Specific instances of such a publication process may vary, especially with recent initiatives for more openness and quality such as public reviews [27], preregistration (e.g. against HARKing) [4], Open Access licensing [24], preprint servers [2], independent non-profit journal publishers, open data publishing [11], FAIR data [31], or open code [9].

Although these practices have demonstrated their potential, the habits of researchers are slow to change and require action by all involved stakeholders [19]. That is why a paper published today is still very much a snapshot in time describing the results of a researcher's work. As Buckheit and Donoho put it: "An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures" [3, p. 59]. The parts that make up the publication are difficult to re-use and applying the "good enough" practices [32] is still perceived as additional work, not as a way to improve the review process and make an article's argument stronger. At the same time, the citation as a mechanism to give credit is disputed, with software and data citations [13, 22] as well as altmetrics [23] emerging as complementary means to recognise and refer to activities that demand a lot of time and effort from researchers.

In summary, we can thus conclude that though there are different disciplines dealing with spatiotemporal information, scientists in these fields do research in a similar way and face similar challenges. This way is common in other scholarly disciplines as well but suffers from a number of key issues such as low reproducibility, difficult re-use, and no deeper inspection. The following sections will propose an alternative approach that tackles these problems in Geoinformatics. We focus on Geoinformatics here due to its computational emphasis while dealing with spatiotemporal information but the ideas reported below are applicable to other fields as well. It is important to point out that our focus is on computational research here. Other types of research carried out in the fields discussed above (e.g. qualitative or explorative research) are outside of the scope of this paper.

3   Executable Research Compendia

In order to overcome the issues we highlighted above, we have developed an approach that combines essential elements related to a specific research activity in Geoinformatics into a coherent compendium. In the following we first describe the basic concept and its realisation before discussing how it fits into the research workflow. We also briefly contrast this approach with current practice and highlight implications of using the new approach.
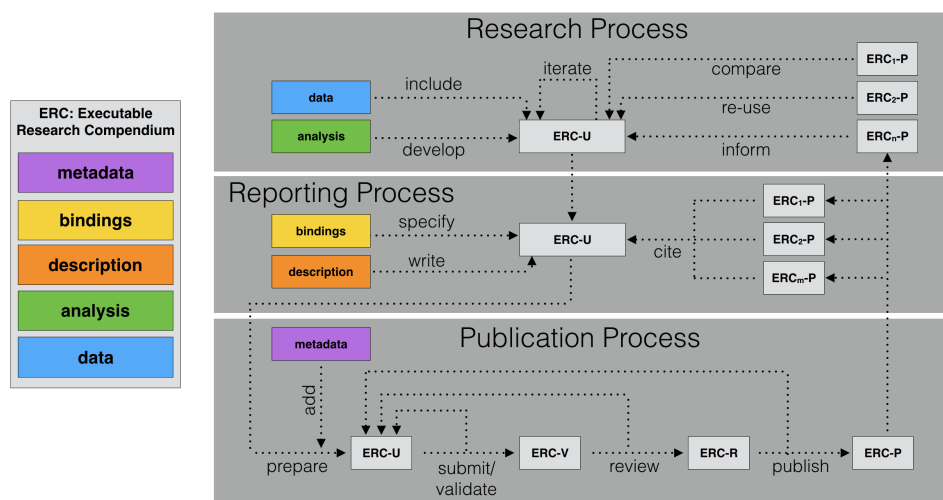
Figure 1 Executable Research Compendium (ERC) with its five key components (left) and how ERCs can be integrated into the research, reporting and publication processes (right). ERC-U stands for an unvalidated ERC, ERC-V for a validated one, ERC-R for a reviewed ERC, and ERC-P for a published one. Processes are sequentialised to make the figure more readable.

3.1   Concept and realisation

An *Exectuable Research Compendium (ERC)* includes all research components that are needed to reproduce the computational results in a paper [20]. It is meant to replace the "classic" static paper, e.g. a PDF or HTML file made up of textual, tabular, and graphical elements that describe the research work, its outcomes, context, and meaning. The five ERC components are depicted in Figure 1 (left) and consist of the following five items (from bottom to top):

- the *data* that was the input to a process creating the results that are being reported; this could consist, e.g., of satellite imagery, a formal specification of spatio-temporal configurations or a transcript of navigational instructions produced by study participants.

- the *analysis* or computational steps that were applied to the data in order to generate the results; examples for this component may include source code implementing a geostatistical method, a set of rules formally specifying a spatial reasoning process, or a code snippet that generates word clouds of textual navigational instructions.

- the *description* of the reported research; this part corresponds largely to a "classic" paper: it contains text providing motivation and background of the research as well as a description of the process, its outcome and the implications/meaning of the latter

- the *bindings* are an optional[1] component describing the links between the three components listed above on a fine-grained level. They can specify which part of the analysis produced which result reported in the description using which part of data. For example, a binding can encode that a specific figure in the description was produced using a particular function included in the analysis component that was applied to a specific dataset included in the data component. Bindings can also specify user interface (UI) controls that enable readers to deeply interact with results, for example by moving a slider next to a figure to change a parameter in the computations that produced this figure, which in turns results in a re-run of the computations and an updated figure.
- the *metadata* component contains meta information about the entire research component. For example, it may include author names, keywords or version information about the interpreters and libraries that were used to produce the results reported in the ERC.

By combining these five elements in a coherent compendium, all relevant information is readily available to carry out a number of essential research activities that currently are difficult to perform with "classic" publications. For example, an ERC makes it easy to re-run the computational steps behind the reported results both for human readers and systems supporting ERCs as a digital object. Section 3.3 includes a more in-depth comparison between ERCs and traditional publications.

The feasibility of the proposed concept is demonstrated by a prototypical platform that was implemented using containerisation to encapsulate the five components of an ERC [21]. The container recipe, in case of Docker[2]: the *Dockerfile*, also specifies the computational environment. The current prototype supports research that uses the R language[3] [28] to carry out spatiotemporal analysis on arbitrary data. The prototypical platform can be extended to include other analysis methods (e.g. the Python language[4] or a specific theorem prover). The prototype also provides methods for creating, running and comparing ERCs as well as further functionalities such as support for interactive figures and "one-click reproduce".

3.2 Integrating ERCs into the research workflow

While ERCs on their own offer some useful properties that can benefit researchers individually, it makes sense to consider their use in the larger context of a typical research workflow. Generally speaking, researchers carry out multiple, potentially overlapping activities when investigating a particular topic. In addition to performing the research itself, for example developing, testing and evaluating a model to simulate navigation, they also work on reporting and publishing the research so that other researchers can use (i.e. extend or build upon) it in their work. ERCs can be used at each step in this overall workflow as detailed in the following and depicted in Figure 1 (right).

Starting with the *research process* itself, a researcher can work on an unvalidated ERC (ERC-U) while deciding which data to include and while developing the analysis procedure. In doing so, she can use previously published ERCs (ERC-P) in several ways. She can compare her data, analysis or results to those reported in an ERC-P. She can also directly re-use parts of an ERC-P, for example, its data to test her analysis. Finally, she can use an ERC-P much like a "classic" publication to inform her work on a more general level, e.g. to

---

[1] Bindings are optional as they require additional effort from authors and basic reproducibility can already be achieved without them.
[2] `https://www.docker.com`, accessed on August 21, 2019.
[3] `https://www.r-project.org`, accessed on August 21, 2019.
[4] `https://www.python.org`, accessed on August 21, 2019.

modify the assumptions or thresholds she is using in her analysis. Once the research has progressed sufficiently, the researcher will want to *report* on it so that others can benefit from her insights. For this purpose, she will write a description much like a 'classic' publication, consisting of text, tables and figures, and add it to the ERC-U that contains the data and analysis she developed previously. During this process, she will contextualise her work by referring to related work, citing other ERCs or parts of them. In addition, she can create a bindings component for her ERC-U to make explicit how data, analysis and description are connected at a fine-grained level. For example, she might specify for a figure which function in her analysis component was applied to what part of the data to generate that figure.

Typically, the researcher will then want to publish the compendium in an academic outlet (e.g. a journal or a conference) to ensure interested readers learn about the new insights she gained through her research and to receive feedback. For this purpose, she first adds relevant metadata to ensure that others can easily use the ERC she has produced and that the ERC conforms to the requirements of the outlet. (Semi-)automatic validation mechanisms can help with this preparation and ensure that the analysis component in the ERC produces the results in the description component. Once the validation is successful, the validated ERC (ERC-V) can be submitted to peer-review. Reviewers are now able to investigate a much larger part of the scholarship underlying the "classic" paper, which is usually not available at all or not integrated into the review process. For example, if reviewers possess the skills to review the code, they can do so, but even if they do not, they can at least confirm that the computations produce the results reported in the paper and check how the results vary when parameters are changed. Depending on the outlet the review process can take different forms and may result in several iterations during which the author needs to revise her ERC. If accepted, the ERC can then be published, resulting in a published ERC (ERC-P). The final publishing process includes multiple steps, such as updating the meta-information to specify which issue/year the ERC was published in. The ERC-P then becomes available for other scientists, who can use it for their own research.

3.3   Benefits and challenges

ERCs come with a number of benefits that enrich a reader's workflow while studying a paper [14]. First, the source code is directly reproducible/executable in a predefined computational environment. In contrast, today's papers are, if at all, supplemented by a folder including files that contain data and code or an (incomplete) reference to the software that was used (e.g. without version, and not persistently stored). This leaves readers with the (daunting) task of figuring out how to run the included analysis. Reviewers face the same challenge with traditional papers while ERCs enable them to easily validate the analysis described in a paper by rerunning it. Second, ERCs can also enhance how researchers work with scientific publications. When searching for relevant articles, they can do so in a more fine-grained way, e.g. by also considering spatial and temporal information derived from the data or by specifying spatiotemporal constraints. While reading an article, they can in parallel investigate how the authors produced a specific figure. In case the analysis code exposes control parameters, it is possible to provide readers with UI controls that enable them to interactively manipulate the initial configuration within a range predefined by the author to see how the results change. Finally, researchers can substitute the original input dataset by data that resulted from their own experiments or another paper, or replace the included code with an alternative one. To fully realise those benefits, an easy-to-use UI is highly desirable not only for inspecting results or interacting with them but also for comparing the computational outputs to quickly identify differences. Otherwise, it might be difficult to spot the differences of an original figure and the one produced by changing a parameter.

The use of ERCs also introduces a number of challenges that need consideration. One such issue is that the concept of an ERC can only fully work if the used research software and data are open source and available with a suitable license (e.g. analysis code written in R and the data could be provided under Creative Commons Zero). Otherwise, it is not possible to include all necessary software to reproduce the results of the paper, to create bindings or to fully re-use the analysis/data. However, many computational analyses are realised using restrictively licensed and proprietary software or data formats, such as Matlab [18] or ArcGIS [10]. Though specific licenses or managed execution by the companies producing the proprietary software might facilitate their use in ERCs, the inherent lack of source code counteracts the principle of complete transparency of the research. A further challenge linked to ERCs is that researchers may be unable (e.g. due to privacy concerns) or reluctant to open their research to the degree required by ERCs despite the known benefits of open data [25] or efficiency, continuity, and reputation [17]. Researchers might fear that others "steal' their data/ideas, that the additional information exposes them to heightened scrutiny, or that the code they wrote for the analysis is not worth publishing or a publication might be harmful [1]. Finally, if an ERC reports on research that involves time-consuming computations or very large data sets, then local execution or transmission of the ERC become unfeasible.

In order to establish ERCs as a desirable alternative to current practice and to realise the benefits outlined above, it is thus necessary to not only overcome some technical and legal challenges but also to change the mindset and behaviour of researchers working in Geoinformatics. Clearly communicating the potential benefits and adding further ones (such as interactive figures, one-click-reproduce or easy re-use) can help to address the latter issue. While we cannot deny that creating ERCs requires additional effort compared to submitting a PDF file, we believe that making research results easier to find, easier to understand, and reusable does not only lead to a higher impact but also strongly benefits authors and the discipline as a whole. The legal aspects are more difficult to tackle. Requiring researchers to only use open source software with permissive licenses could negatively affect their work. It also seems unlikely that vendors of proprietary software would easily agree to their software becoming part of an ERC as this would enable third parties to run their software for free. Many of the (technical) challenges outlined above (big data/long computation) can be addressed by the support infrastructure for ERCs we envision in the following section.

4   An Open Research Infrastructure for Geoinformatics (OpenRIG)

ERCs encapsulate individual research contributions and by combining all key elements linked to those contributions in a coherent way, they offer a number of benefits as outlined in section 3.3. In order to realise their full potential, there needs to be a supporting infrastructure that provides various functions, in particular those that affect multiple ERCs. Since previous work has focused on the realisation of ERCs individually [14, 15, 21], we will outline our vision for such an infrastructure in the following sections. We first give a rationale for why it is needed and then discuss components and services it should provide. In addition, we point out key challenges and opportunities that result from working towards and with such an infrastructure. Figure 2 provides an overview of what we envision for this infrastructure.

4.1   Rationale

There are several reasons why a research infrastructure for ERCs is beneficial. From a technical perspective, local execution (i.e. on the researcher's computer) is not always feasible: the computer might not be powerful enough to execute the ERC, the computation might take
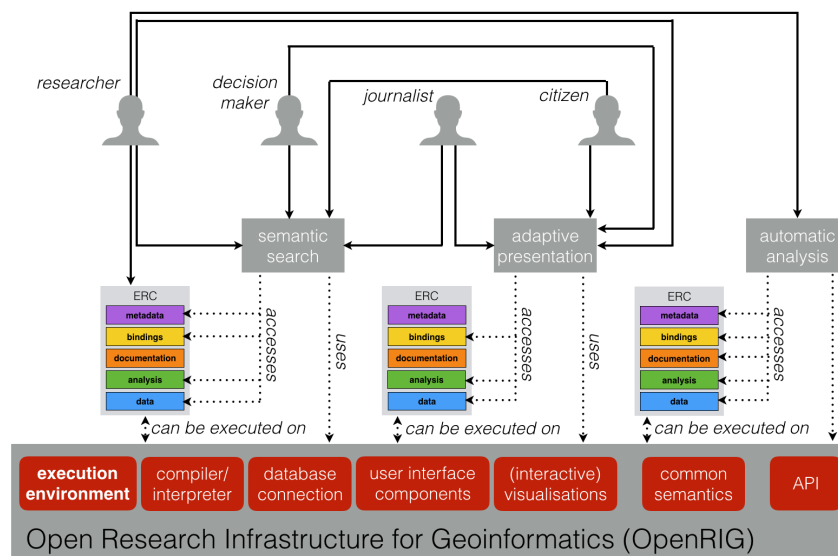
Figure 2 Open Research Infrastructure for Geoinformatics (OpenRIG): key components (red), essential functionalities enabled by it (grey boxes) and different stakeholders wanting to access them.

too long, and/or the amount of data might be to large to transmit/store. In addition, some desirable functions such as comparing multiple ERCs / their components or a *semantic search* require access to *many* ERCs at the same time, which is not feasible inside an individual ERC. For the semantic search in particular, it is necessary to analyse specific components of several ERCs and to apply constraints/filters across them.

Another opportunity that arises from the structure inherent to ERCs and the inclusion of bindings in particular is the ability to adapt the way in which results are presented. This is beneficial, for example, to harmonise diagrams across multiple ERCs so that it is easier to relate the reported results to one another. *Adaptive presentations* also allow for generating figures and diagrams that meet the needs and preferences of different stakeholders. For example, researchers might want a higher level of detail than journalists and colour-blind readers might prefer contrast-rich colour palettes.

Furthermore, ERCs also pave the way for *automatic analysis* processes. Unlike traditional papers, ERCs are highly structured, semantically annotated, and they include data and code. Automatic processes can use this structured information to perform tasks such as automatic validation of individual ERCs (are the results reported in the description component produced by the included analysis applied to the data) or cross-validation (are the results reported in an ERC consistent with those reported in ERCs with the same/similar data set and/or the same/similar analysis method). More sophisticated automated analyses are also possible, e.g. by using spatiotemporal reasoning to combine results from multiple papers.

Additionally, there are a number of practical reasons why a research infrastructure for ERCs is beneficial. It would allow for the seamless realisation of the publication process depicted in Figure 1. In addition, it could also provide a means for archival of ERCs so that long-term storage (and execution) can be guaranteed. Finally, a common infrastructure could also contribute towards ensuring that all researchers have access to the same resources thereby levelling the playing field.

4.2   Components

An infrastructure supporting functionalities as discussed in the previous sections fundamentally needs to be open, i.e. to facilitate the inclusion of runtime environments and to enable bindings. The term "open " in this context specifically refers to "open source " and a license that is permissive. We envision an *Open Research Infrastructure for Geoinformatics (OpenRIG)* that contains the following core components/elements.

The most obvious component of the OpenRIG is an *execution environment* that can execute ERCs "in the cloud ". This is particularly important in cases where the analysis included in the ERC is computationally demanding, e.g. would take hours, days or longer on a standard PC. Such an execution environment would also benefit the realisation of automatic analysis processes for similar reasons. Automatic analysis processes could potentially be realised as ERCs as well: the analysis itself would be contained in the analysis component and the documentation could describe the purpose of the automatic analysis process. However, this might require means to persist the state of the analysis between subsequent executions of the automatic analysis. Otherwise, the periodic re-running of such a process might entail unnecessarily re-analysing all ERCs that were processed already in previous executions. Automatic execution of ERCs might also come into play in case of updates of a software library that was used in the analysis code. The resulting changes could affect the final output making it necessary to re-execute the analysis and to see if the results are still the same. A further possibility is related to metadata which is usually entered manually: an automatic extraction might relieve authors from this task, which is disliked by many researchers. An automatic analysis could also be part of an "ongoing " meta-analysis determining, for example, how many papers address research about a specific topic. Once the computational analysis is implemented, the input data could be updated regularly.

Independently of how such automatic analysis processes are realised, they have a substantial potential to improve the way in which research is done in Geoinformatics. For example, they could automatically compare ERCs that are newly published to similar existing ERCs, e.g. to check for anomalies or to infer new knowledge resulting from a newly added ERC. A new method to predict navigation errors might thus be applied to the data in previous papers that investigated the same issue. This, in turn, might help to better determine the overall performance of the new method compared to previously published ones.

Another useful element of a research infrastructure for Geoinformatics is a persistent and standardised *connection to relevant databases*. Particularly when dealing with very large datasets such as satellite data, it is not feasible to include a full copy of the data in each ERC that used it in its analysis. Instead, using a persistent, versioned interface to access this data from an ERC does not only reduce the size of an ERC but also has benefits in terms of efficiency. Given such a database connection, the infrastructure can make sure that the analysis is executable "near " the actual data, e.g. on the same server, to avoid unnecessary delays resulting from transmitting large amounts of data.

Two further beneficial and related elements of an execution environment are a set of *user interface (UI) components* and *(interactive) visualisations*. The former could provide a library of controls that enable authors to add interactivity to their ERCs, e.g. to provide readers with means to explore more deeply how results change when certain assumptions change. The latter targets visualisations particularly as they play a key role in the understanding of scientific publications [6]. In addition to standard diagrams such as bar charts or box plots, this includes in particular spatiotemporal visualisations such as maps or space-time cubes. When ERCs use these sets of standardised UI components and visualisations, this opens up the possibility to change how specific results are presented. For example, the projection,

colour palette and colour break values of a map in one ERC could be adapted to be the same as the one used in another ERC to facilitate comparison. Different stakeholders might also prefer certain visualisations (e.g. varying degree of detail, colour scheme), which constitutes another type of adaptive presentations that could be realised in this way.

In addition, we envision the OpenRIG to contain a *common semantic framework* that formally describes core spatiotemporal concepts, the components of ERCs as well as their relationships and interactions. This will not only enable reasoning about individual ERCs but also strengthen the capabilities of semantic search functions and automatic analysis processes as they can operate on a more abstract level. Semantic information about spatiotemporal aspects of ERCs can be included in their metadata component. Furthermore, semantically describing different visualisations and UI components opens the door for reasoning about how information is presented in line with the idea of interface plasticity [5]. This can also help addressing one of the key challenges in geovisualisation: adapting geovisualisations [16].

The final element that a research infrastructure for Geoinformatics should provide is a standardised *Application Programming Interface (API)*. This more technical aspect is important for providing easy access to all the functions provided by the OpenRIG. For example, in order to implement semantic search, adaptive presentations or automatic analysis, it is necessary to interact with various elements of the OpenRIG. Providing a versioned and persistent API is not only important to ensure automatic analysis processes remain operational but also for the long-term archival (and execution) of ERCs.

The open research infrastructure for Geoinformatics (OpenRIG) envisioned in this section thus facilitates various desirable functions. In addition, it provides a number of interesting opportunities for future research and poses several challenges that require further investigation. The following section provides an overview over these aspects.

4.3   Opportunities and Challenges

Among the opportunities offered by the envisioned research infrastructure, being able to *perform reasoning* on top of it is very promising. Obviously, a well-designed semantic framework is required to realise this, which constitutes another opportunity for interesting research. The envisioned reasoning includes the automatic analysis processes mentioned above but can be extended substantially – given that ERCs are highly structured, semantically annotated entities with a strong spatiotemporal component. An example for such work could be research into which example cases the region connection calculus (RCC) [29] was applied to in different publications. The collected datasets describing the example cases could then be used as a corpus to compare RCC to a different approach in terms of whether the latter can solve all example cases as well, and whether the results are identical for both approaches. In addition to reasoning about the content of an ERC, there is an opportunity to further investigate how key results are presented. As mentioned in section 4.2, bindings semantically link data, analysis and description. Formally describing descriptions of spatiotemporal results – for example, by establishing equivalence between different visualisation types – would enable the seamless adaptation of how outcomes are presented to a human reader.

A further opportunity is the assessment of *meaningfulness* of a specific type of analysis approach to check whether the computations making up the analysis of a specific ERC are meaningful (and not just mathematically possible). Spatiotemporal properties can play a key role in determining whether a specific computation is meaningful. In the context of spatial prediction and aggregation an initial approach of this type has been proposed [30]. With the OpenRIG and a broad availability of ERCs, this idea could be extended to other (qualitative) calculi such as RCC and use the metadata of an ERC and the common semantics of the

OpenRIG as a basis. Eventually, this could even become a safeguard that is incorporated into the validation process of ERCs, e.g. upon submission to an outlet.

The solutions presented here also come with several challenges. Some papers published in Geoinformatics do not rely on computational analysis but on qualitative data and analysis, e.g. interviews. While such publications thus are not executable, they still include data. One option to integrate them with the proposed approach could be to extend the bindings concept to guide readers through the data analysis, for example, by connecting statements in the results section of the paper with (anonymised) quotes and the aggregated higher-level themes. Still, for some types of scientific output, e.g. purely conceptual papers, although relevant in general, the proposed approach is potentially less beneficial. The higher-level challenge is of course to create opportunities and benefits so good that it incentivises researchers in Geoinformatics to actively adopt open science principles. Over the last decade we have learned that only talking about open science has not led to a substantial change in research practice, despite broad agreement about research ethics. Changing practice will need a concerted effort from scholars, reviewers, publishers, and libraries as well as science-funding bodies. Creating rewards for open science activities that go beyond publishing text papers is an important component of this. All of these parties must collaborate to operate the OpenRI(G), or instances of it. It would make sense to extend and connect existing building blocks to tackle the financial and organisational issues of operating such an infrastructure. These building blocks include, for example, review and publishing services by libraries or scholarly societies, computing resources of research institutes, data hosting by domain observatories, and funding schemes for sustainable software development or open access journals. We also envision that an OpenRIG would be useful to scholars from many disciplines. It would thus make sense to extend an existing research infrastructure such as Zenodo[5] with spatiotemporal search- and link-capabilities. That would provide this functionality to scholars in any discipline where spatially and temporally referenced data are used, including hydrologists, ecologists, meteorologists, geographers, archaeologists, and so on.

5   Concluding Remarks

In this paper, we propose the development and the adoption of an *open research infrastructure for Geoinformatics* to help us, scientists, to transform the centuries old process of *only* sharing textual and pictorial descriptions of our research findings into one where also the data and the data analysis *procedures* are shared, comprehensibly and reproducibly. This will not only have the advantage of increased transparency and enable more trust in science in general, but also create new options for interacting with the data and procedures, searching and finding particular datasets or applications of methods, and ways to link together research components. This also directly applies for some of the research that is reported at COSIT, e.g. new calculi or algorithms to tackle specific spatiotemporal problems and apply them to example scenarios to demonstrate their usefulness. In addition, the vision outlined in this paper provides new opportunities for fundamental research in spatial information theory, e.g. in terms of designing a semantic framework capturing spatiotemporal aspects of ERCs so that deep reasoning about multiple ERCs is enabled. Previous work shows that the technical realisation of such an infrastructure is by all means possible [20, 21, 14, 15]. In order to make this rather disruptive proposal for a transformation to open science a reality, the key challenge is a social one: although scientists agree that openness is essential, most of them

---

[5] `https://zenodo.org`, accessed on August 21, 2019.

hesitate to bear the (initial) costs themselves. As a starting point, we suggest that scientists who review manuscripts reporting on computational research start to decline doing this in cases where data and reproducible procedures are not made available to the reviewers.

──── **References** ────

1 Nick Barnes. Publish your computer code: it is good enough. *Nature News*, 467(7317):753–753, October 2010. `doi:10.1038/467753a`.

2 Philip E. Bourne, Jessica K. Polka, Ronald D. Vale, and Robert Kiley. Ten simple rules to consider regarding preprint submission. *PLOS Computational Biology*, 13(5):e1005473, May 2017. `doi:10.1371/journal.pcbi.1005473`.

3 Jonathan B. Buckheit and David L. Donoho. WaveLab and Reproducible Research. In Anestis Antoniadis and Georges Oppenheim, editors, *Wavelets and Statistics*, number 103 in Lecture Notes in Statistics, pages 55–81. Springer New York, 1995. `doi:10.1007/978-1-4612-2544-7_5`.

4 Andy Cockburn, Carl Gutwin, and Alan Dix. HARK No More: On the Preregistration of CHI Experiments. In *Proceedings of CHI 2018*, pages 141:1–141:12, New York, NY, USA, 2018. ACM. `doi:10.1145/3173574.3173715`.

5 Joëlle Coutaz. User interface plasticity: Model driven engineering to the limit! In *Proceedings of ACM EICS 2018*, pages 1–8. ACM, 2010.

6 David DiBiase, Alan M MacEachren, John B Krygier, and Catherine Reeves. Animation and the role of map design in scientific visualization. *Cartography and geographic information systems*, 19(4):201–214, 1992. `arXiv:https://doi.org/10.1559/152304092783721295`.

7 David Donoho. 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, 26(4):745–766, October 2017. `doi:10.1080/10618600.2017.1384734`.

8 Matt Duckham. GI Expertise. *Transactions in GIS*, 19(4):499–515, 2015. `doi:10.1111/tgis.12166`.

9 Steve M Easterbrook. Open code for open science? *Nature Geoscience*, 7(11):779, 2014. `doi:10.1038/ngeo2283`.

10 Esri. ArcGIS Pro product page. `https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview`. Accessed April 15, 2019.

11 Virginia Gewin. Data sharing: An open mind on open data. *Nature*, 529(7584):117–119, January 2016. `doi:10.1038/nj7584-117a`.

12 Simon Hettrick. It's impossible to conduct research without software, say 7 out of 10 UK researchers. *Software and research*, 5:1536, 2014.

13 Daniel S. Katz and Neil P. Chue Hong. Software Citation in Theory and Practice. In James H. Davenport, Manuel Kauers, George Labahn, and Josef Urban, editors, *Mathematical Software – ICMS 2018*, Lecture Notes in Computer Science, pages 289–296. Springer International Publishing, 2018. `doi:10.1007/978-3-319-96418-8_34`.

14 Markus Konkol and Christian Kray. In-depth examination of spatio-temporal figures in open reproducible research. *Cartography and Geographic Information Science*, 2018. `doi:10.1080/15230406.2018.1512421`.

15 Markus Konkol, Christian Kray, and Max Pfeiffer. Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. *International Journal of Geographical Information Science*, pages 1–22, 2018. `doi:10.1080/13658816.2018.1508687`.

16 Alan M MacEachren and Menno-Jan Kraak. Research challenges in geovisualization. *Cartography and geographic information science*, 28(1):3–12, 2001.

17 Florian Markowetz. Five selfish reasons to work reproducibly. *Genome Biology*, 16:274, 2015. `doi:10.1186/s13059-015-0850-7`.

18 MathWorks. Matlab product page. `https://www.mathworks.com/products/matlab.html`. Accessed April 15, 2019.

19 Daniel Nüst, Carlos Granell, Barbara Hofer, Markus Konkol, Frank O Ostermann, Rusne Sileryte, and Valentina Cerutti. Reproducible research and GIScience: an evaluation using AGILE conference papers. *PeerJ Preprints*, 6:e26561v1, 2018.

20 Daniel Nüst, Markus Konkol, Edzer Pebesma, Christian Kray, Marc Schutzeichel, Holger Przibytzin, and Jörg Lorenz. Opening the publication process with executable research compendia. *D-Lib Magazine*, 23(1/2), 2017.

21 Daniel Nüst. Reproducibility Service for Executable Research Compendia: Technical Specifications and Reference Implementation, December 2018. `doi:10.5281/zenodo.2203844`.

22 Hyoungjoo Park and Dietmar Wolfram. An examination of research data sharing and reuse: implications for data citation practice. *Scientometrics*, 111(1):443–461, April 2017. `doi:10.1007/s11192-017-2240-2`.

23 Heather Piwowar. Altmetrics: Value all research products. *Nature*, 493:159, January 2013. `doi:10.1038/493159a`.

24 Heather Piwowar, Jason Priem, Vincent Larivière, Joan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6(e4375), February 2018. `doi:10.7717/peerj.4375`.

25 Heather A. Piwowar and Todd J. Vision. Data reuse and the open data citation advantage. *PeerJ*, 1:e175, October 2013. `doi:10.7717/peerj.175`.

26 Hardy Pundt and Fred Toppen. 20 Years of AGILE. In Arnold Bregt, Tapani Sarjakoski, Ron van Lammeren, and Frans Rip, editors, *Societal Geo-innovation*, pages 351–367. Springer, Cham, 2017. `doi:10.1007/978-3-319-56759-4_20`.

27 Ulrich Pöschl. Interactive open access publishing and public peer review: The effectiveness of transparency and self-regulation in scientific quality assurance. *IFLA Journal*, 36(1):40–46, March 2010. `doi:10.1177/0340035209359573`.

28 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL: `https://www.R-project.org/`.

29 David A Randell, Zhan Cui, and Anthony G Cohn. A spatial logic based on regions and connection. *KR*, 92:165–176, 1992.

30 Christoph Stasch, Simon Scheider, Edzer Pebesma, and Werner Kuhn. Meaningful spatial prediction and aggregation. *Environmental Modelling & Software*, 51:149–165, 2014.

31 Mark D. Wilkinson *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018, March 2016. `doi:10.1038/sdata.2016.18`.

32 Greg Wilson, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Nederbragt, and Tracy K. Teal. Good enough practices in scientific computing. *PLOS Computational Biology*, 13(6):e1005510, June 2017. `doi:10.1371/journal.pcbi.1005510`.

Cross-Corpora Analysis of Spatial Language: The Case of Fictive Motion

Ekaterina Egorova (ORCID)
Massey Geoinformatics Collaboratory, Massey University, Auckland, New Zealand
e.egorova@massey.ac.nz

Niloofar Aflaki (ORCID)
Massey Geoinformatics Collaboratory and School of Natural and Computational Sciences,
Massey University, Auckland, New Zealand
n.aflaki@massey.ac.nz

Cristiane K. Marchis Fagundes (ORCID)
Department of Geomatics, Federal University of Paraná, Curitiba, Paraná, Brazil
cristiane.fagundes@ufpr.br

Kristin Stock (ORCID)
Massey Geoinformatics Collaboratory, Massey University, Auckland, New Zealand
k.stock@massey.ac.nz

—— **Abstract** ——

The way people describe where things are is one of the central questions of spatial information theory and has been the subject of considerable research. We investigate one particular type of location description, fictive motion (as in, *The range runs along the coast*). The use of this structure is known to highlight particular properties of the described entity, as well as to convey its configuration in physical space in an effective way. We annotated 496 fictive motion structures in seven corpora that represent different types of spatial discourse – news, travel blogs, texts describing outdoor pursuits and local history, as well as image and location descriptions. We analysed the results not only by examining the distribution of fictive motion structures across corpora, but also by exploring and comparing the semantic categories of verbs used in fictive motion. Our findings, first, add to our knowledge of location description strategies that go beyond prototypical locative phrases. They further reveal how the use of fictive motion varies across types of spatial discourse and reflects the nature of the described environment. Methodologically, we highlight the benefits of a cross-corpora analysis in the study of spatial language use across a variety of contexts.

1   Motivation and Background

Any language is a network of options and, whether consciously or not, its speakers regularly make linguistic choices in everyday communication situations [23]. Spatial language in particular offers multiple options for the subtle windowing of attention of a listener to particular aspects of described scenes [21]. Borrowing an example by Matlock, imagine the scene behind *The table goes from the kitchen wall to the sliding glass door* [9]. In hearing this sentence, our minds automatically draw a table that is long and narrow – we do not visualise a round kitchen table or a small square coffee table. This interpretation is evoked by the use of fictive motion (henceforth FM), a linguistic structure that includes motion information in the description of the location of a static entity [20].

Understanding patterns of FM use is important for spatial information theory for several reasons.

First, multiple studies in cognitive linguistics and psychology have shown how the use of FM in language reflects the focus of attention of the speaker, while also affecting the mental representation of a scene as constructed by the listener [13, 9, 11, 12]. From the speaker's side, the use of FM signals the conceptual primacy of a spatial entity and its configuration in space, since FM offers an efficient way for conveying information about the physical layout of a scene [13, 9]. Further, FM can highlight and even construe spatial properties of the described entity, as in the example with the table above, where the table "becomes lengthened through dynamic construal" [12, p. 548]. Other properties reported in previous studies on FM include vertical orientation (as in, *route plunges*) and complex shape or structure (as in, *glacier spills*, *mountains roll*) [4]. For the listener, FM and the semantics of the verb induce a mental simulation of motion, which results in a particular mental representation of a spatial scene [11]. A number of experiments with participants have demonstrated that people process semantically equivalent expressions with FM and without FM differently [10, 8, 13]. For example, they would draw a longer tattoo if the latter was described with *The tattoo runs along his spine*, as opposed to *The tattoo is next to the spine* [11]. Studying patterns of FM use can thus provide insights into particular aspects of the construction and communication of mental representations of space.

Second, FM is described as "pervasive across languages" and is known to often occur "when people are describing physical space" [8, p. 1390]. Given this pervasiveness, the importance of studying patterns of FM use has been acknowledged in the line of work that explores spatial concepts found in various types of spatial discourse [15]. Developing a spatial language annotation scheme, Pustejovsky and Yocum introduce the MOTION SENSE attribute to distinguish between different interpretations of motion events, one of them being FICTIVE [16]. In a corpus-based study, Egorova et al. further report on types of scenes and spatial concepts encoded by FM: actual motion of the observer (as in, *The second icefield led much more quickly than anticipated*), general encyclopedic knowledge (as in, *The range runs east west across the central part of the Tibet plateau*), vistas (as in, *Far off, a great red buttress rose steeply*)[4]. Building upon [4], another study develops a rule-based approach for the automated extraction and classification of FM from text, demonstrating the non-trivial nature of such tasks but pointing out that exploring patterns of use of figurative language such as FM is necessary if we want to develop algorithms that exhibit spatial awareness [3].

Fictive motion has been thoroughly studied in cognitive linguistics, but research is largely based on introspection (e.g. [21]) or general corpora such as BNC or novels (e.g. [19]). To the best of our knowledge, no studies have actually examined the frequencies and patterns of its use across different types of spatial discourse. Conducting a cross-corpus analysis of FM

can contribute to our knowledge of spatial description strategies used in different contexts. It can also make a step towards a more nuanced automated annotation of spatial information in text, which is crucial given that the correct interpretation of motion has a "lasting effect on the interpretation of a text with respect to spatial information" [15, p. 992]. To address this gap and explore the opportunities, the following research questions have been formulated for this study:

- RQ1: How frequent is FM in various types of spatial corpora?
- RQ2: How do motion verbs in FM and their semantic classes differ across corpora?
- RQ3: What does this tell us about spatial discourse production in different contexts?

2   Data and Methods

2.1   Corpora

Four corpora were used in our analysis, and one of the four – **Nottingham Corpus of Geospatial Language (NCGL)** [18] – was divided into four subcorpora, allowing analysis across different domains.

The **NCGL: News** sub-corpus includes 1592 geospatial sentences from 14 news web sites from the USA, Australia, New Zealand and South Africa.

The **NCGL: Travel and Tourism** sub-corpus includes 3380 sentences from 9 web sites including travel blogs (e.g. Seat61), tourism agency sites (e.g. Tourism NZ) and tourism publishers (e.g. Lonely Planet).

The **NCGL: Outdoor pursuits** sub-corpus contains 1822 sentences from 6 sites, mainly focused on walking (e.g. Arizona Trails, BBC Walks).

The **NCGL: Local History** sub-corpus contains 2104 sentences and focuses on local history, harvested from 11 sites, mostly from the UK (with one site from Australia).

The **Geograph** corpus includes descriptions from Geograph, an online project that collects geographically representative images (and their descriptions) for every square kilometre of the British Isles.[1] Descriptions used in this study refer to images of six neighbouring squares within the urban area of London and are represented by 3153 sentences.

The **Where am I?** corpus was created using human subjects experiments [17], in which respondents were shown an array of photos of a particular location, and were asked to imagine that they had witnessed an accident and to describe the location to emergency services. 178 native English speakers responded to the experiment, which resulted in a corpus of 737 sentences.

The **National Soils Database**[2] (**NSD**) is a collection of descriptions of locations of soil specimens gathered by Manaaki Whenua - Landcare Research in New Zealand. Our corpus consists of a subset numbering 1389 sentences.

2.2   Fictive motion annotation and analysis

FM annotation involved two steps: automated annotation of FM candidates in the corpora and manual annotation of FM among the candidates. In the first step we automatically identified sentences containing motion verbs, based on part-of-speech tags and lemmas of motion verbs as compiled from two sources [4, 6]. In the second step the candidates were split into

---

[1] `https://www.geograph.org.uk`
[2] `https://soils.landcareresearch.co.nz/soil-data/national-soils-data-repository-and-the-national-soils-database/`

■ **Table 1** Characteristics of corpora and FM in them.

Corpus	N of sentences	N of FM candidates	N of FM	FM per candidate
LOCAL HISTORY	2104	239	114	0.48
OUTDOOR PURSUITS	1822	478	192	0.4
NEWS	1592	291	27	0.09
TRAVEL AND TOURISM	3380	575	87	0.15
NSD	1389	107	5	0.05
GEOGRAPH	3153	312	27	0.09
WHERE AM I?	737	249	44	0.18
Total	14177	2251	496	n/a

three sub-corpora and FM structures were annotated in each sub-corpus by one of the first three authors of the paper. A motion event was considered fictive if the noun linguistically represented as a moving entity was a static entity. To measure the inter-coder agreement, we randomly sampled 5% of candidate sentences from each corpus, and the resulting corpus of 112 candidate sentences was independently annotated by the three mentioned annotators. Further, we performed a cross-corpora comparison of FM using a combination of quantitative and qualitative methods. First, we calculated the ratio between the number of FM structures and the number of FM candidates as a proxy for evaluating the pervasiveness of FM in each corpus. Second, we classified the verbs according to their semantics and compared the distribution of classes across corpora in order to examine how the semantics of verbs reflects the nature of spatial discourse. The motion verbs' classification scheme was borrowed from [22] where path verbs include "Source-originating" (e.g. *leave*), "Goal-oriented" (e.g. *reach*), "Vertical" (e.g. *ascend*), "Trajectory" (e.g. *cross*) and "Change in direction verbs" (e.g. *turn*), while manner verbs include "Complex shape trajectory" (e.g. *wind*) and "Trajectory of unspecified shape" (e.g. *roll*) verbs.

3 Results and Interpretation

In total, 496 FM structures were identified in all corpora. The average pairwise Cohen's Kappa, a standard measure of agreement [2], is 0.78, which is a good positive indication of the reliability of the annotation.

3.1 Fictive motion frequency

The highest proportion of FM in relation to the number of FM candidates[3] is found in **Local history** – 0.48 (see Table 1). This might reflect the nature of the corpus – focussing on the history of rural England, it is rich in descriptions of vistas of local (both built and natural)

---

[3] This metrics essentially represents the ratio between the number of all motion verbs in the corpus and the number of verbs used in FM. We chose to report this ratio (instead of verbs used in FM per number of sentences) for two main reasons. First, it avoids the problem arising from the presence of multiple FM in one sentence. Second, this ratio is more revealing about the use of motion verbs, and is more relevant for the line of work that captures spatial information in text and distinguishes between various interpretations of motion events [16].

landscapes. Additionally, it has a largely poetic flavour and creatively deploys FM to convey nuanced properties of the environment (as in, *To the left is the mansion, skirted by the gloomy cedars, and beyond, the lake expanding into a noble sheet of water is embosomed in magnificent woods*). Another corpus with a high ratio of FM is **Outdoor pursuits** (0.4). From the perspective of spatial discourse, texts in this corpus are mostly descriptions of trails or narratives about completed walks, whereby FM appears to offer an effective way of communicating spatial information about trail-like features (as in, *The path goes between trees by the side of the lake*). The ratio of FM in **Travel** was found to be quite low (0.15) which can be explained by the fact that the corpus mostly describes travelling by trains and buses. Thus, while spatial descriptions including motion events abound, they mostly include means of travel (e.g. *ferry*, *train*) and factive motion (as in, *The Livorno train heads down the Rhine Valley in the early evening, past castles, Rhine river barges and vineyards*).

The smallest ratios are found in NSD (0.05), News (0.09), Geograph (0.09) and Where am I? (0.18). In the case of **NSD** (0.05), this might be explained by the nature of geographic information that it contains – namely, precise quantitative metrically-grounded descriptions of small-scale locations. In a rather similar way, FM is rarely used in **Where am I?** (0.18), which might also be explained by the necessity to describe one's own position as precisely as possible. This results in a high frequency of prototypical locative phrases and descriptions of landmarks (as in, *There's a pedestrian crossing and a disabled parking spot in front of the school building. The building is brick and concrete fame with blue walls*). Furthermore, the scenario given to respondents in the Where am I? survey was rather utilitarian and urgent in nature, allowing little room for consideration of different modes of expression. The sample from **Geograph** (0.09) also represents descriptions of urban vistas (captured in images), but since there is no task of describing the location, descriptions of space are rather scarce – instead, the focus is often on people and events (as in, *A guard stands to attention as the people walk by*). Finally, in the case of **News** (0.09) the low frequency of FM reflects the focus on events and their locations, mostly represented by the first- and second-order political entities such as countries and regions (as in, *Net traffic will travel to the satellite through Hughes' Earth station near Los Angeles*).

3.2   Verbs in fictive motion

57 different motion verbs occurred in 496 FM structures in the corpora. The most frequent verbs were *run* (86 inst.), *lead* (73 inst.), *pass* (50 inst.), *go* (47 inst.), *cross* (34 inst.), *turn* (31 inst.), *take* (26 inst.), *follow* (25 inst.), *climb* (17 inst.), *wind* (16 inst.). Among the verb classes, the most prominent class is "Trajectory of unspecified shape", followed by "Vertical" and "Trajectory" classes. A cross-corpora comparison of the distribution of classes in the three largest sub-corpora of fictive motion (**Outdoor pursuits**, **Local history**, and **Travel**) invites for several observations.

"Complex shape trajectory", while almost absent in **Local history** (2.63%), has slightly higher proportions in the two other corpora (7.35% for **Outdoor pursuits**, 8% for **Travel**) where it is represented by verbs such as *wind*, *meander*, *snake*, *wrap*, *twist*. This class of verbs is mostly used to describe the shape of water bodies (as in, *The Nile snakes through upper Egypt*) or trails (as in, *The track meanders through gullies*).

"Vertical" verbs are similarly frequent in **Local history** (14%) and **Outdoor pursuits** (15.1%). However, in **Outdoor pursuits** the two dominant verbs are *climb* and *drop*, usually collocated with a trail-like entity. In contrast, in **Local history** this class is overly represented by *rise* that is frequently used to describe human-built parts of the landscape (as in, *The spire of Edwinstowe Church rises gracefully from among the old oaks*).

"Trajectory of unspecified shape" is the most frequent class in both **Local history** (35%) and **Travel** (42.5%). Verbs in this group mostly encode the spatial extension of an entity (as in, *A small wood stretched from Jenny Burton's Hill to near her cottage*). In **Outdoor pursuits**, in contrast, "Trajectory of unspecified shape" represents 18% only, while the most frequent class is "Trajectory" (verbs such as *cross*, *follow*, *traverse*). This reflects the focus on the path of the motion event in the context of outdoor activities, where locomotion is an important part of navigation, as in *The footpath initially follows the right hand field boundary.*

4   Conclusion and Outlook

A cross-corpora analysis of fictive motion has provided us with several insights that have important implications and invite for further investigations.

First, our findings suggest that the relative frequency of FM in a particular type of spatial discourse depends on aspects such as the scale of described scenes (we found more FM in the descriptions of vistas of landscapes in **Local history** and spatial layouts of trails in **Outdoor pursuits**), required precision of spatial information (we saw less instances of FM in **NSD**, where preference is given to metric locative phrases), as well as the main theme of spatial descriptions (we saw less instances of FM in **News**, where spatial information mostly relates to the location of events, and not spatial entities). Second, the semantic classes of verbs found in FM further reflect the peculiarities of each corpus, both from the perspective of the described environment and from the perspective of spatial information in focus. Example of the former is the low frequency of "Trajectory of unspecified shape" verbs in **Local history**, which might be a result of the absence of features such as large winding rivers in the described area. Example of the latter is the high frequency of "Trajectory" verbs in **Outdoor pursuits**, which might reflect the focus on the path in the context of walking and hiking.

These findings have practical implications for several lines of work within the spatial information theory. For the line of work developing spatial annotation schemes and capturing spatial information in text [15, 14], this study highlights the fact that taking FM into account is especially legitimate when working with corpora similar in their characteristics to **Local history** and **Outdoor pursuits**. It also provides insights which are of key importance for the development of parsers that are capable of distinguishing between factive and fictive motion in text [3]. The findings are further relevant for the development of spatial language generation systems [5, 7, 24], given that the use of FM has the capability of inducing a more effective processing of spatial information through the simulation of motion [8]. Finally, for the line of research looking into spatial language use across various contexts, the study brings an important message of the utility of a cross-corpus analysis.

In future work, we plan to enhance our understanding of FM use through more controlled, hypothesis-driven studies. In the next step, we aim at performing a more systematic analysis of verbs' classes and types of spatial entities occurring in FM, as well as at exploring potential conventionalised FM structures. More broadly, further work is also required for a better understanding of how we could model the representations of spatial scenes encoded by FM, and how situatedness and context impact FM use and interpretation [1].

References

1 John A Bateman, John Hois, Robert Ross, and Thora Tenbrink. A linguistic ontology of space for natural language processing. *Artificial Intelligence*, 174(14):1027–1071, 2010.

2 Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960.

3 Ekaterina Egorova, Ludovic Moncla, Mauro Gaio, Christophe Claramunt, and Ross S Purves. Fictive motion extraction and classification. *International Journal of Geographical Information Science*, 32(11):2247–2271, 2018.

4 Ekaterina Egorova, Thora Tenbrink, and Ross S Purves. Fictive motion in the context of mountaineering. *Spatial Cognition & Computation*, 18(4):259–284, 2018.

5 Mehdi Ghanimifard and Simon Dobnik. Knowing when to look for what and where: Evaluating generation of spatial descriptions with adaptive attention. In Laura Leal-Taixé and Stefan Roth, editors, *Computer Vision – ECCV 2018 Workshops. ECCV 2018. Lecture Notes in Computer Science, vol 11132*, pages 153–161. Springer, Cham, 2018.

6 Beth Levin. *English verb classes and alternations: A preliminary investigation*. University of Chicago press, Chicago, London, 1993.

7 Tomoyuki Maekawa and Wataru Takano. Spatial representation of context-dependent sentences and its application to sentence generation. *Advanced Robotics*, 31(15):780–790, 2017.

8 Teenie Matlock. Fictive motion as cognitive simulation. *Memory & Cognition*, 32(8):1389–1400, 2004.

9 Teenie Matlock. The conceptual motivation of fictive motion. In Günter Radden and Klaus-Uwe Panther, editors, *Studies in linguistic motivation*, pages 221–248. Mouton de Gruyter, Berlin, New York, 2004.

10 Teenie Matlock. Depicting fictive motion in drawings. In June Luchjenbroers, editor, *Cognitive Linguistics Investigations: Across languages, fields and philosophical boundaries*, volume 15, pages 67–85. John Benjamins Publishing, Amsterdam, 2006.

11 Teenie Matlock. Abstract motion is no longer abstract. *Language and Cognition*, 2(2):243–260, 2010.

12 Teenie Matlock and Till Bergmann. Fictive Motion. In E Dąbrowska and D. Divjak, editors, *Handbook of Cognitive Linguistics*, pages 546–562. Berlin, Mouton de Gruyter, 2014.

13 Teenie Matlock and Daniel C Richardson. Do eye movements go with fictive motion? *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26), 2004.

14 Ludovic Moncla, Mauro Gaio, Javier Nogueras-Iso, and Sébastien Mustière. Reconstruction of itineraries from annotated text with an informed spanning tree algorithm. *International Journal of Geographical Information Science*, 30(6):1137–1160, 2016.

15 James Pustejovsky. ISO-Space: Annotating static and dynamic spatial information. In Nancy Ide and James Pustejovsky, editors, *Handbook of Linguistic Annotation*, pages 989–1024. Springer, Dordrecht, 2017.

16 James Pustejovsky and Zachary Yocum. Capturing motion in ISO-Spacebank. In *Proceedings of the 9th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation*, pages 25–34, 2013.

17 Kristin Stock and Luciene Delazari. Where am I? The challenges of interpretation of natural language descriptions of geographical location. In *SAI Computing Conference (SAI), 2016*, pages 1335–1338. IEEE, 2016.

18 Kristin Stock, Robert C Pasley, Zoe Gardner, Paul Brindley, Jeremy Morley, and Claudia Cialone. Creating a corpus of geospatial natural language. In Thora Tenbrink, John Stell, Antony Galton, and Zena Wood, editors, *Spatial Information Theory. COSIT 2013. Lecture Notes in Computer Science, vol 8116*, pages 279–298. Springer, Cham, 2013.

19 Dejan Stosic and Laure Sarda. The many ways to be located in French and Serbian: the role of fictive motion in the expression of static location. In Brala M. Vukovic and Gruic L. Grmusa, editors, *Space and Time in Language and Literature*, pages 39–60. Cambridge Scholars Publishing, 2009.

20   Leonard Talmy. Fictive motion in language and "ception". In Paul Bloom, Merrill F Garrett, Lynn Nadel, and Mary A Peterson, editors, *Language and Space*, pages 211–277. MIT Press, Cambridge, MA, 1996.

21   Leonard Talmy. *Toward a cognitive semantics.* MIT Press, Cambridge, MA, 2000.

22   Piia Taremaa. Fictive and actual motion in Estonian: Encoding space. *SKY journal of linguistics*, 26:151–183, 2013.

23   Thora Tenbrink. Cognitive discourse analysis: Accessing cognitive representations and processes through language data. *Language and Cognition*, 7(1):98–137, 2015.

24   Jette Viethen and Robert Dale. The use of spatial relations in referring expression generation. In *INLG'08 Proceedings of the Fifth International Natural Language Generation Conference, Salt Fork, Ohio, June 12 - 14, 2008*, pages 59–67. Association for Computational Linguistics, Stroudsburg, PA, USA, 2008.

Talk of the Town: Discovering Open Public Data via Voice Assistants

Sara Lafia 🆔
Department of Geography, University of California, Santa Barbara, USA
slafia@ucsb.edu

Jingyi Xiao
Department of Geography, University of California, Santa Barbara, USA
jingyi_xiao@ucsb.edu

Thomas Hervey 🆔
Department of Geography, University of California, Santa Barbara, USA
thomasahervey@ucsb.edu

Werner Kuhn 🆔
Department of Geography, University of California, Santa Barbara, USA
werner@ucsb.edu

──── **Abstract** ────

Access to public data in the United States and elsewhere has steadily increased as governments have launched geospatially-enabled web portals like Socrata, CKAN, and Esri Hub. However, data discovery in these portals remains a challenge for the average user. Differences between users' colloquial search terms and authoritative metadata impede data discovery. For example, a motivated user with expertise can leverage valuable public data about transportation, real estate values, and crime, yet it remains difficult for the average user to discover and leverage data. To close this gap, community dashboards that use public data are being developed to track initiatives for public consumption; however, dashboards still require users to discover and interpret data. Alternatively, local governments are now developing data discovery systems that use voice assistants like Amazon Alexa and Google Home as conversational interfaces to public data portals. We explore these emerging technologies, examining the application areas they are designed to address and the degree to which they currently leverage existing open public geospatial data. In the context of ongoing technological advances, we envision using core concepts of spatial information to organize the geospatial themes of data exposed through voice assistant applications. This will allow us to curate them for improved discovery, ultimately supporting more meaningful user questions and their translation into spatial computations.

1 Open Data: Of, By, and For the People?

Open data, also called public sector information, aspire to increase the transparency of government activities and their accountability to the public [10]. In the United States, mandates for open data are often satisfied in part by the adoption of platforms, like CKAN and ArcGIS Hub, which mediate public access to government data catalogs. The platforms often include both geospatial (e.g. parcel maps) and non-geospatial data (e.g. tax tables).

However, the production, maintenance, and dissemination of such authoritative open public data is costly for providers and the effort does not guarantee increased public engagement [6]. Furthermore, work remains to be done to help governments keep track of the direct and indirect benefits of their open data policies, measured in part by tracking data reuse [2].

Impediments to the uptake of open public data include challenges with data discovery and usability. *Discovery* is understood broadly as a mode of exploratory search that involves browsing for task appropriate data, while *usability* describes the fitness of data for a defined task[1]. A major impediment to data discovery in human-system communication is reflected by the "vocabulary problem" [5], in which users rarely agree on what to call the things that they want to find. This makes effective keyword-based search and discovery difficult to accomplish in public open data portals. To address this problem, some open data portals like Esri's ArcGIS Hub have partnered with community organizations to develop ontologies that map the terms used to describe community level initiatives from authoritative vocabularies (e.g. the USGS Thesaurus) to users' colloquial terminology [8].

While this strategy addresses data discovery, it does not address underlying issues with data usability. Even if users are better able to identify task-appropriate data, they generally do not know how to assess the fitness of data for a given task and are still expected to manipulate and analyze data to gain insights. Given these constraints, services like data dashboards are being developed, allowing users to track vital community issues (e.g. pedestrian fatalities[2]) without requiring them to manipulate, clean, or visualize data. Even more empowering are alternative modalities, such as those offered by voice assistants, which are growing in popularity[3] and have implications for open public data discovery and use. Governments have suggested that voice assistants might offer new interfaces for connecting community members to public services and information exposed through open public data portals.

In this paper, we explore the current capabilities of various voice assistants under development by local governments across the United States. We focus on the application areas that these systems are designed to address and examine how (if at all) they leverage geospatial data. Next, we discuss the challenges that voice assistants face when answering geospatial questions. Finally, we envision using core concepts of spatial information [7] to organize the geospatial themes of data that users want to discover, with the goal of supporting a broader range of user questions and spatial computations on them. We focus primarily on improvements to discovery for existing systems that also carry benefits for data usability.

2 State of the Art for Government Voice Assistants

Voice assistants are now widely available on commercial smart speakers, such as Google Assistant and Amazon Alexa. A recent survey[4] has projected that half of all U.S. homes will own smart speakers by the end of 2019. The same survey also reported that the most common interactions with voice services include asking questions, performing online searches, performing basic research like confirming information, and asking for directions.

While today's voice assistants are used to control home automation systems and perform other basic daily tasks, interest has shifted to more intelligent interactions such as enabling natural conversations and answering questions. Users are able to ask questions about real-

---

[1] https://www.force11.org/group/fairgroup/fairprinciples
[2] http://visionzero.lacity.org/map/
[3] https://www.citylab.com/solutions/2018/10/amazon-alexa-smart-speakers-city/573412/
[4] https://www.cmo.com/features/articles/2018/9/7/adobe-2018-consumer-voice-survey.html

time information, such as *what time it is now?* and *what will the weather be like tomorrow?* When users talk to a voice assistant, their spoken words can be converted to text by APIs (e.g. Amazon's automatic speech recognition[5] and Google's Speech-to-Text API[6]). The diversity of expression in human language has posed enormous difficulties to language understanding. With the state-of-the-art natural language understanding (NLU) techniques, including syntax analysis (e.g. tokenization, identifying part-of-speech), entity recognition (e.g. organization, person, location), sentiment analysis, and intent and topic detection, machines are able to "understand" user questions. By detecting the given topic and intent, related information and potential answers are retrieved from various databases like Wikipedia, Google's knowledge graphs, and Microsoft's concept graphs. Retrieved information is then used to generate responses using different methods, such as rule-based and generative methods. The responses are then converted from text back to speech to answer user questions conversationally.

Voice assistant technology is now being leveraged to retrieve and reason on open public data through the development of skills (which are essentially micro tasks). In 2017, Esri prototyped an early government voice assistant application called Sonar[7]. It offered a chatbot that completed predefined tasks and addressed standard questions about a given community by leveraging open data available through Esri Hub. As shown in Figure 1, Sonar performs lookups on data matching the themes described in a user's query at a defined location. Users can ask about city services (e.g., trash pickup), safety (e.g., crimes), and transportation (e.g., bus routing). Sonar facilitates both open public data discovery and use by templatizing a set of intents designed to perform basic computations on geospatial data. In other words, Sonar provides a set of "core questions" that a community member would want to ask, and maps them to available, thematically relevant data, using location as context. Thus, governments can build additional skills upon Sonar's foundation.

Since the advent of Sonar, many U.S. cities have developed ad hoc voice assistant applications. Many are designed to reduce administrative burdens, such as "311 information" calls. For example, the Alexa skills developed for Albuquerque, New Mexico[8] allow residents to register complaints about graffiti, weeds, abandoned vehicles, and ask questions about city-owned facilities, like fee information for public parks. Raleigh, North Carolina[9] also allows residents to ask questions about the government, such as trash pickup days or elected representatives for a given neighborhood. Similarly, specific city departments, like New York City's Department of Environmental Protection[10], have created Alexa skills that allow residents to check their water usage and pay their bills. Los Angeles, California[11] has released several voice applications that provide residents with local information about recent earthquakes. The earthquake alert works on the Google Home system, which harvests USGS seismic data to notify residents of recent earthquake events based on the location of their device. The Alexa skills of Johns Creek, Georgia[12] are robust, continuously mining the city's open data portal to provide updated information about zoning and road closures.

These applications all work by knowing where to find open public data and how to use it in order to answer typical questions that people ask about government. Today, many

---

[5] https://developer.amazon.com/alexa-skills-kit/asr

[6] https://cloud.google.com/speech-to-text/

[7] https://github.com/Esri/sonar

[8] https://www.cabq.gov/alexa

[9] https://www.raleighnc.gov/home/news/content/CorNews/Articles/AlexaApp.html

[10] https://www1.nyc.gov/html/dep/html/customer_assistance/amazon-alexa.shtml

[11] https://assistant.google.com/services/a/uid/00000096ea087604?hl=en

[12] https://www.amazon.com/City-of-Johns-Creek-GA/dp/B07BHPGDR1

```
33
36    ### Amazon Echo
37
38    - Help help {Dataset}
39    - Hello Hello Sonar.
40    - Hal Open the pod bay doors HAL.
41    - Ping ping {Dataset}
42    - GetPopulation give me the population of {Location}
43    - GetPopulation how many people live nearby {Location}
44    - GetPopulation what is the population of {Location}
45    - GetPopulation population {Location}
46    - GetPopulation people in {Location}
47    - GetCrime safety of {Location}
48    - GetData what is the {Dataset} of {Location}
49    - GetData what is the nearest {Dataset} at {Location}
50    - GetData tell me about {Dataset} at {Location}
51    - GetData ask about {Dataset} at {Location}
52    - GetData when is {Dataset} for {Location}
53    - SummarizeData how many {Dataset} at {Location} since {TimePeriod}
54    - SummarizeData in past {TimePeriod} how many {Dataset} at {Location}
55    - AddNote add note {Note} at {Location}
56    - GetMap map of {Location}
57    - GetMap see {Dataset} at {Location}
58    - ExitApp stop
```

Figure 1 A list of the six intents (*ping, get population, get data, summarize data, add note,* and *get map*) accessible to users through the Sonar project's Alexa skill.

applications are built on top of voice assistant-accessible databases that contain standardized open public data (e.g. government data catalogs exposed through Esri's Open Data Hub). However, new trends such as the uptake of the schema.org Dataset standard[13] for the annotation of open public metadata (e.g. in Google Dataset Search[14]) enable the discovery of open public data through search engines [3]. As public data discoverability increases for the average user, it will also likely increase for the average voice assistant application. Thus, as it becomes easier for humans and machines to discover open public data, how can data be organized to facilitate use? We propose that time and space, inherent to geospatial data in particular, makes the themes that they are "about" more amenable to such curation.

3 Geospatial Limitations of Government Voice Assistants

The prospect for data discovery and question answering in the applications described in Section 2 is promising. Many municipalities are working to rapidly expand the skills that their voice assistants use to help answer questions and engage their communities. This is a reasonable tactic because it is likely that the efficacy of voice assistants will improve greatly over the next decade. These systems are synthesizing factual data with real-time computational abilities, using semantic technologies to answer increasingly complex questions.

However, we have observed two problems with this trend, which are even more evident when it comes to addressing geospatial questions: 1) voice assistant applications frequently bypass discovery, and 2) governments are building unsustainable skills. The first problem means that a system supplies an answer to a question without first allowing a user to explore available data. This may not seem like a problem when considering the alternative: a voice

---

[13] https://schema.org/Dataset
[14] https://toolbox.google.com/datasetsearch

assistant that would conversationally list available data. This mode of interaction would be tedious and far less efficient than exploring data by using a graphical browser. In a way, voice assistants are perceived to have abilities like those of a question-answering oracle. These question-answering systems bypass the process of manually discovering, manipulating, using, and reasoning on data themselves. In many cases, users often quickly accept the top suggestions by search engines[15]. However, much of the value of open data, especially open geospatial data, is the ability to explore and synthesize information, and conduct visual analysis. This is not possible with a voice assistant. What would be optimal for discovery is to make voice assistants more conversational. If a voice assistant application creates an index of datasets based on generic concepts that a user is familiar with, such as *objects* and *networks*, then the system could conversationally suggest relevant datasets.

The second problem is that if governments continue to build skills in their current manner, after a few years, they will likely have to maintain many heterogeneous (geospatial) tasks that will also be hard to improve. In other words, building skills in this manner is unsustainable. Furthermore, most of the aforementioned examples of applications in Section 2 are not explicitly geospatial. Those that could be considered geospatial work by retrieving pre-generated data from factual databases (e.g. water usage), and some leverage near real-time geospatial information (e.g. earthquakes). More complex geospatial questions, like those specific to a user's location, require more complex geospatial computing and cannot yet be answered. For instance, a question like *which hospitals are open now and are also within a 20-minute drive from home?*, cannot be answered simply by retrieving data from databases. Such questions require geospatial analysis and computing, which could be partially supported by leveraging existing APIs. We therefore believe that if skill building could leverage the organizational structure of data, and a corresponding conceptual model that humans have of these types of data, then perhaps computing with them could be easier as well.

4   A Vision for Geospatially-Enabled Voice Assistants

We propose a conceptual framework adopted from Cook and Daniels' software design methodology [4] as a means of facilitating geospatial data discovery and subsequent use to provide answers to users' geospatial questions. Our work formalizing this conceptual model for spatial data is ongoing and is applicable to both GIS and voice assistant environments. We are not proposing an implementation solution; rather, we are proposing a conceptual model to help organize the things that people want to ask about and the computations on geospatial data to answer those questions.

Cook and Daniels' software design methodology is comprised of an essential model, a formal model, and a system model. The essential model is a model of the world built by objects and events used to understand a situation. The formal model (also called the specification model) states what the software will do and formalizes the essential model by mathematical operations. The system model (also called the implementation model) specifies system-level behavior based on the formal model.

In our framework, the essential model specifies concepts about the real world. Since spatial questions are about things in the real world, they are cognitively represented by core concepts of spatial information like *fields* and *networks* [7]. Thus, the procedure to answer a question can be formalized as as a set of spatial operations with mathematical foundations (as a formal model). The information detected from user questions can be used as input for

---

the spatial operations. The spatial operations can then be implemented in a chosen software (as a system model). The results are finally computed by the chosen software and returned to the user as an answer. An example of this framework is shown in Figure 2.

> **Essential model**: a conceptualization of those aspects of the real world
> (user asks a question about the schedule of a public bus service)

> **Formal model**: formalizes the essential model by mathematical operations
> (graph theory formalizes the network of the public bus service)

> **System model**: specifies the behavior of the information system in the chosen software
> (software specifies the methods to compute the bus trip given the input itinerary)

Figure 2 The essential (blue), formal (green), and system (red) conceptual levels.

To operationalize this framework, we need a means of relating human concepts to formal operations and system level commands in a GIS [1]. Kuhn's core concepts of spatial information [7] provide a bridge, specifying concepts in user questions at the essential level and relating them to operations at the formal level. Previous work on question-based spatial computing used data abstraction to relate user questions to computations in a GIS [9].

Progress can be made on the essential and formal models for at least two core concepts: *fields* and *networks*. Fields as an essential model conceptualize continuous phenomena and are characterized by continuous functions from location to theme. Prototypical examples include elevation, temperature, and rainfall. Fields are formalized by map algebra. The field concept allows users to ask questions like *how much did it rain in my neighborhood last night?* Networks are a topological essential model, formalized by graph theory. They allow users to ask questions like *how many bus stops are between my house and downtown?* The system model could take the form of an existing geocomputation API (e.g. GDAL, ArcGIS Online, etc.). Today, architectures of many geospatially enabled portals (e.g. Socrata with QGIS[16], Hub with ArcGIS Online[17]) are already equipped to handle the system model specifications. By formalizing the operations that are to take place on the geospatial data in the portal, today's voice assistant applications move closer to the capabilities of conversational GIS.

The mathematical formalization of fields and networks suggests a manageable set of questions that users could ask of open government data. We surmise that these two concepts, their mathematical models, and the accompanying software packages, could provide an entry point for mapping between user questions and computations, following the architecture illustrated in Figure 2. In this vision, voice assistants serve as a kind of conversational GIS, answering a far broader range of geospatial questions about government.

---

[16] `https://dev.socrata.com/blog/2016/06/13/geospatial-analysis.html`
[17] `https://doc.arcgis.com/en/hub/sites/explore-data.htm`

Voice assistants following this framework would organize contents based on their spatial concepts, suggesting data and operations to perform on them based on the concepts in a user's question. For example, such a system could parse the previous question *which hospitals are open now and are also within a 20-minute drive from home?*, and recognize that "hospitals" are likely to be objects in a health care data set and determine that "a 20-minute drive" would require a road network data set. A computation would intersect currently open hospitals (stored as an attribute of the open dataset) and a 20-minute roadway service area from the user's home. If suitable open data sets do not exist, the voice assistant could suggest alternatives with similar themes based on the concepts present in the original question.

5 Conclusion

A vast amount of open public data is ready for discovery. Technological advances in voice assistant technology have the potential to actively connect users to developments in their communities. In this paper, we have explored voice assistant applications that governments are developing to improve open public data discovery and use. To address challenges that today's applications face, we have proposed a conceptual framework informed by core concepts of spatial information and structured as an essential, a formal, and a system model. Relating the language of user questions about the world to spatial computations is a step toward improving discovery and use of open public data for users and their communities.

───── **References** ─────

1 Guoray Cai, Hongmei Wang, Alan M. MacEachren, and Sven Fuhrmann. Natural Conversational Interfaces to Geospatial Databases. *Transactions in GIS*, 9(2):199–221, March 2005. `doi:10.1111/j.1467-9671.2005.00213.x`.

2 Wendy Carrara, Wae San Chan, Sander Fischer, and Evan Steenbergen. Creating value through open data: Study on the impact of re-use of public data resources. *European Commission*, 2015. `doi:10.2759/328101`.

3 Davide Castelvecchi. Google unveils search engine for open data. *Nature*, 561(7722):161–162, 2018. `doi:10.1038/d41586-018-06201-x`.

4 John Daniels and Steve Cook. *Designing Object Systems: Object-oriented Modelling with Syntropy*. Prentice Hall, Englewood Cliffs, NJ, September, 1994. ISBN: 0-13-203860-9.

5 George W. Furnas, Thomas K. Landauer, Louis M. Gomez, and Susan T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):964–971, 1987. `doi:10.1145/32206.32212`.

6 Peter A. Johnson, Renee Sieber, Teresa Scassa, Monica Stephens, and Pamela Robinson. The Cost(s) of Geospatial Open Data. *Transactions in GIS*, 21(3):434–445, 2017. `doi:10.1111/tgis.12283`.

7 Werner Kuhn. Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information Science*, 26(12):2267–2276, 2012. `doi:10.1080/13658816.2012.722637`.

8 Sara Lafia, Andrew Turner, and Werner Kuhn. Improving Discovery of Open Civic Data. *LIPIcs-Leibniz International Proceedings in Informatics*, 114(9):1–15, 2018. `doi:10.4230/LIPIcs.GIScience.2018.9`.

9 Behzad Vahedi, Werner Kuhn, and Andrea Ballatore. Question-based spatial computing—A case study. In *Geospatial Data in a Changing World*, pages 37–50. Springer, 2016. `doi:10.1007/978-3-319-33783-8_3`.

10 Anneke Zuiderwijk and Marijn Janssen. Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1):17–29, 2014. `doi:10.1016/j.giq.2013.04.003`.

Detecting the Geospatialness of Prepositions from Natural Language Text

Mansi Radke 🄳

Computer Science and Engineering Department, Visvesvaraya National Institute of Technology, Nagpur, India
mansiaradke@gmail.com

Prarthana Das

Computer Science and Engineering Department, Visvesvaraya National Institute of Technology, Nagpur, India
musiciselixir@gmail.com

Kristin Stock 🄳

Massey Geoinformatics Collaboratory, Massey University, Auckland, New Zealand
k.stock@massey.ac.nz

Christopher B. Jones 🄳

School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom
jonescb2@cardiff.ac.uk

──── **Abstract** ────

There is increasing interest in detecting the presence of geospatial locative expressions that include spatial relation terms such as *near* or *within <some distance>*. Being able to do so provides a foundation for interpreting relative descriptions of location and for building corpora that facilitate the development of methods for spatial relation extraction and interpretation. Here we evaluate the use of a spatial role labelling procedure to distinguish geospatial uses of prepositions from other spatial and non-spatial uses and experiment with the use of additional machine learning features to improve the quality of detection of geospatial prepositions. An annotated corpus of nearly 2000 instances of preposition usage was created for training and testing the classifiers.

1 Introduction

Automated recognition and disambiguation of geographic references in text documents has received considerable attention in recent years, often with the motivation of indexing the documents with regard to geographic space. The methods used to date have been dominated by a focus on identifying geographic names, i.e. toponyms, and using these directly as the basis for geographic footprints for text expressions or entire documents. The assumption however is that the references are absolute in the sense that the toponym provides the actual location referred to. While this is a reasonable default assumption, it is very common to refer to locations in an indirect manner using spatial relations, such as *near*, *at*, *close to*, *north of* etc., relative to a reference location. These expressions often take the form of triples of a subject (or located object), the spatial relation and an object (the reference location), as in "St Mary Church near Times Square." While some authors have proposed methods for modelling vague spatial relations such as *near* (e.g. [7, 10, 11]), relatively little work has been done on the basic, initial problem of reliably identifying the presence of relative

locational descriptions in natural language texts ([3, 5, 6, 8]). Effective methods for doing this are required as part of the process of extracting and interpreting indirect geographic references and to retrieve other geospatial facts that associate an event or some other object with a reference location, as for example in "Roald Dahl was born in Cardiff". Locational description detection methods are also required for automatic creation of test collections that can be used in developing and evaluating methods for spatial relation extraction and for modelling the use of individual spatial relations, e.g. [9]. In this paper, we present methods for automatic detection of spatial relational terms in sentences, in particular prepositions, that are used specifically in a geospatial sense and we distinguish these from prepositions that have other spatial senses and from prepositions that have no spatial meaning. We are interested in the ability to distinguish between spatial and geospatial senses of prepositions, as this is important for detecting text that can be georeferenced and thus mapped on a geographical scale (in contrast to text that describes a location inside a room, or on a person's body), a goal that is useful in a wide range of application areas.

The approach adopted is here applies the spatial role labelling method of [3]. That work aimed to detect all three components of spatial relational expressions which were referred to as the trajector, i.e, the located object, spatial indicator, i.e. the individual preposition that serves as spatial relation, and the landmark which is the reference location. Here we use their preposition disambiguation method, which was employed as part of a pipeline approach to detection of triples. The method was tested in [3] only for the purpose of detecting generic spatial prepositions, which might or might not be geospatial. Here we train the classifier on sentences containing a preposition that is used either in a geospatial sense, a spatial but not geospatial sense, or in a sense that is not spatial in any respect. We also experiment with modifying the classifier for geospatial prepositions to take account of other evidence that indicates the presence of place names and geographic feature types.

For the purpose of evaluating the approach, we have created a corpus of 1876 instances of preposition usage that have been manually labelled as geospatial, spatial (but not geospatial) and non-spatial. These prepositions occur within 674 sentences.

In the remainder of the paper Section 2 describes related work, Section 3 explains the methodology in detail, while Section 4 gives the details of the data set used and the experiments performed. Section 5 concludes the paper, pointing out some directions for future work.

2   Related work

A method specifically designed to detect whether a preposition has a spatial sense was presented by Kordjamshidi et al. [3] in a paper on spatial role labelling in the context of relation extraction. The paper focused on the three roles of trajector (located object), spatial indicator (spatial relation) and landmark (reference location). Two approaches to spatial role labelling were presented. In the first approach, called the pipeline approach, an input sentence is passed to the first stage of the pipeline which tokenizes the sentence and passes each token to a Part of Speech (POS) tagger. The sentence is also processed by a dependency parser and a semantic role labeller (the `LTH software` from [1]). If a preposition is identified by the POS tagger, a Naive Bayes classifier is used to make a decision on whether it is used in a spatial sense. The features used by the classifier are based on output from the POS tagger, the dependency parser and the semantic role labeller. For this stage of identifying the spatial sense of a preposition, an F1 score of .88 was achieved for the TPP dataset [4] with 10 fold cross validation. If the preposition is determined to have a spatial sense, then it is passed to

a second stage of the pipeline which identifies the trajector and the landmark with respect to the spatial indicator. This second stage uses probabilistic graphical models, in particular a Conditional Random Fields classifier, which again takes a variety of features generated by the initial parsing of the sentence. A triple of the form <Trajector, SpatialIndicator, Landmark> is returned as output by the pipeline. The second approach offered by Kordjamshidi et al. [3] uses joint learning in which all three of trajector, spatial indicator and landmark are detected simultaneously.

A method for detecting just the spatial relation and the reference object of spatial relations was described by Liu [5] where these partial relations were described as degenerate locative expressions (DLE). The approach is analogous to methods of Kordjamshidi et al., though they employed a smaller set of features for machine learning, that did not include dependency relations or semantic roles. An evaluation of the method in [6] obtained an F1 score of .76 when applied fully automatically to their TellUsWhere corpus on which it was trained. Note that no distinction was made in that work between geospatial and other spatial senses of prepositions. The method of [5] to extract DLEs was also exploited in Khan et al. [2] in which locative DLEs which explicitly encode spatial relations, with prepositions such as *near* and *in*, were distinguished from partial DLEs where a preposition such as *to* was not regarded as conveying explicit spatial information. A rule based approach was employed to extend the latter to an explicit spatial DLE when it was used as part of a spatial relation such as *next to*. This technique was part of a procedure to extract spatial triples by matching structures from the Stanford parser, of the form <governor, preposition, dependent>, with locative DLEs that used the same preposition. The governor would then serve as the located object of a spatial triple.

As part of a process of creating a corpus of geospatial sentences, Stock et al. [8] employed a set of language patterns to detect various ways in which geospatial information is described. This included a pattern to recognise when a place name or place type is preceded by a spatial relation which could be a preposition (though other parts of speech were also considered to represent spatial relations). They obtained a precision of 0.66 when applying these methods to detect geospatial expressions. A specialized collection of spatial relational expressions was created by Wallgrun, Klippel and Baldwin [9]. They used search patterns to query the web to find expressions that contained any of the three relations of *near*, *close* and *next to*. Their approach therefore constrained the results to include the specified spatial relation. They also confined the expressions to include specified types of located and reference objects. Our work differs from that in allowing any spatial relation that is classed as a preposition and in using a machine leaning approach to determine the geospatial or other spatial sense of the preposition.

3 Methods

3.1 What is a geospatial sense?

In order to distinguish here between geospatial, other spatial and non-spatial uses of prepositions, we employ a simple definition of a geospatial relation as one in which the preposition has a spatial sense and the reference object to which the preposition applies is a geographic feature, as in a named place or a geographic feature type. The reference object is normally expected to be outdoors. If it is part of a building it is expected to be an exterior part. We impose no constraint on the nature of the located object. If a preposition has a spatial sense but the reference object is not geographic then it is classed as spatial. If the preposition has no spatial interpretation then it is classed as neither geospatial nor spatial.

Examples of the kinds of expressions that appear on our corpus include the following, with preposition senses according to our annotation scheme (described above) shown in angular brackets:

- "And now on <non-spatial> a clear morning Graham Little and I are sitting at <geospatial> the bottom of (spatial) the wall fit and ready to go and the wall is plastered with <non-spatial> verglas."

- "In <non-spatial> a minute she had rushed from <geospatial> the house and was running down <geospatial> the garden"

3.2   Classifying prepositions as geospatial or spatial

In this work, we modify the first step of the spatial role labelling pipeline method of [3], i.e. their method for detecting the spatial sense of prepositions, by adding additional features for machine learning. The features used in the original classifier are listed in Table 1. As indicated above these are obtained from a combination of a POS tagger, a dependency parser and a semantic role labeller. The Part-Of-Speech Tagger (POS Tagger) assigns parts of speech to each word, such as noun, verb, adjective, etc. Dependency parsing assigns a syntactic structure to a sentence. The most widely used type of syntactic structure is a parse tree which can be useful in various applications such as grammar checking, but here it plays a critical role in the semantic analysis stage. In natural language processing, semantic role labeling (also called shallow semantic parsing) is a process that assigns labels to words or phrases in a sentence to indicate their semantic role, such as that of an agent, goal, or result. It consists of the detection of the semantic arguments associated with the predicate or verb of a sentence and their classification into their specific roles. We experiment with using just these features, but we also extend the method to add additional features that indicate whether a place name or a geographic place type is present in the expression that includes the target preposition. The presence of a place name is detected with the Geonames gazetteer, while the presence of a place type is detected with a dictionary of geographic place types. The `expat` application was used to generate these features (location and gnn patterns).

We used a Naive Bayes multi-class classifier with three output classes of geospatial, spatial but not geospatial, and neither geospatial nor spatial. We also used Naive Bayes binary classifiers for each one of these three classes *vs* the other two classes.

4   Experimental Set Up

4.1   Data set and its Annotation

Our dataset of 674 sentences was derived from two sources. 185 of the sentences came from the source of about 26,000 sentences that were used in the process of creating the Nottingham Corpus of Geospatial Language (NCGL) [8]. These sentences were harvested from the web using the algorithm described in [8], and was thus biased towards retrieving geospatial content, but also included spatial (but non-geospatial) expressions as well as some uses of prepositions that are non-spatial in any sense. The remainder of our collection is a sample of the TPP dataset of sentences produced for the preposition project (see Litkowski and Hargraves [4]). That dataset includes many examples of both spatial and non-spatial uses of prepositions, though relatively few of them have a geographical context.

Table 1 Features from [3] used in detecting the sense of a preposition.

preposition	the preposition itself
preposition	the lemma of the preposition
preposition	the POS tag of the preposition
preposition	the DPRL of the preposition
preposition	the semantic role label of the preposition
preposition	the sense of the preposition if assigned
preposition	the argument of the preposition in the SRL output
head1	the head1 itself
head1	the lemma of head1
head1	the POS tag of the head1
head1	the DPRL of the head1
head1	the semantic role label of the head1
head1	the sence of the head1 if assigned
head1	the argument of the head1 in the SRL output
head2	the head2 itself
head2	the lemma of head2
head2	the POS tag of the head2
head2	the DPRL of the head2
head2	the semantic role label of the head2
head2	the sence of the head2 if assigned
head2	the argument of the head2 in the SRL output

Many of the sentences include multiple prepositions and so in order to annotate the sense of the individual prepositions we created a distinct instance of a sentence for each preposition that it contained (as determined by a POS tagger). We considered a tuple <Sentence, Preposition> as a unique instance. So, if a sentence instance s had two prepositions $p1$ and $p2$, we created two instances from it, namely $< S, p1 >$ and $< S, p2 >$. This resulted in 1876 instances (indicating an average of just under three prepositions per sentence). These preposition-specific instances were then manually annotated as either geospatial, spatial (but not geospatial) or non-spatial.

Annotation was conducted through an iterative process that involved all four authors. In the case of the NCGL sentences, one person annotated all sentences, a subset of 100 of which were then checked by two others followed by a discussion of disagreements. A fourth person then re-annotated all of those sentences taking account of issues raised in the discussions. The TPP sentences were annotated by one person, after which one other checked them and highlighted disagreements. The first annotator then revised annotations to respect the result of this discussion. Finally a further stage of re-annotation of subsets of 100 of each of both groups of sentences was performed resulting in inter-annotator agreements of 0.89 for the larger TPP sourced data set and 0.75 for the NCGL sourced data set.

As an example of inter-annotator disagreement, consider the following sentence. "After 50m, you will reach a road with wide verges where you turn left toward Lambley." The first annotator marked *after* as non-spatial in sense. The second annotator noted that here *after* is used to represent the geospatial arrangement of different locations, and the latter sense was adopted for the final data set. In another example, in the phrase "Republic of China", the preposition *of* was marked spatial by one annotator, as "China" is a geographical place name, while the other annotator considered it as non spatial since "Republic of China" is an administrative entity. We adopted this latter annotation for the final data set.

4.2 Experiments performed

Before we present our results, we mention the balance of the classes in the datset used. Out of the total preposition instances (1877), the number of instances marked as non-spatial was 770, the number of instances marked as spatial was 773, and the number of instances marked as geospatial was 334.

Table 2 Features used in experiments.

Kord	All features used for preposition sense detection in [3]
Kord-Geo	The features from Kord plus the number of placenames and the number of geographic feature types found in the head words of the preposition
Kord-Geo-S	The features from Kord plus the number of place names and the number of geographic feature types found within the entire sentence in which the preposition occurs
Kord-Geo-All	The features from Kord-Geo-S plus the sum of the numbers of place names and a binary value of true if either a place name or a geographic feature type is present
Geo-Baseline-S	The number of place names and the number of geographic feature types found within the entire sentence in which the preposition occurs

Table 3 Results for 3-class classifier predicting geospatial, spatial (but not geospatial) or neither.

	Geospatial			Spatial			Neither		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Kord	0.442	0.578	0.501	0.747	0.744	0.745	0.763	0.664	0.710
Kord-Geo	0.514	0.614	0.559	0.751	0.762	0.757	0.772	0.696	0.732
Kord-Geo-S	0.566	0.638	0.600	0.732	0.802	0.765	0.783	0.665	0.719
Kord-Geo-All	0.600	0.692	**0.643**	0.749	0.797	**0.772**	0.796	0.692	**0.740**

Table 4 Results for three 2-class classifiers predicting geospatial, spatial (but not geospatial) and neither.

	Geospatial			Spatial			Neither			Spatial or Geospatial		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Kord	0.370	0.647	0.471	0.696	0.790	0.740	0.762	0.751	0.756	0.828	0.836	**0.832**
Kord-Geo	0.423	0.680	0.521	0.704	0.798	0.748	0.760	0.755	0.757	0.830	0.835	**0.832**
Kord-Geo-S	0.480	0.704	0.570	0.688	0.846	**0.759**	0.755	0.753	0.754	0.829	0.830	0.829
Kord-Geo-All	0.542	0.728	**0.621**	0.672	0.837	0.745	0.750	0.771	**0.761**	0.838	0.821	0.829
Geo-Baseline-S	0.625	0.419	0.502	0.494	0.889	0.635	0.422	0.326	0.368	0.595	0.689	0.639

Several experiments were conducted with a Naive Bayes classifier to evaluate the methods described above (note that the original method from [3] uses this classifier for determining the sense of a preposition). In the first experiment (Table 3) a multi-class Naive Bayes classifier was used to predict each of the three classes of geospatial, spatial (but not geospatial) and neither. There were several versions of the classifier that use different combinations of features (summarised in Table 2). One of these (Kord) just uses the features from [3] described above. It resulted in an F1 value of 0.50 for the geospatial class and better values of 0.745 for spatial and 0.710 for neither. This was extended by adding the two features of the number of place names and number of geographical features detected in the head words of the preposition that is being tested (Kord-Geo). Note that the head words are among the features generated by the procedure used in [3]. They correspond to the subject and object of the preposition. A further variation (Kord-GeoS) records these latter numbers at the sentence level, which was found to improve upon the performance when only observing

head words (though note that the quality of performance will depend upon the performance of the script to detect place names and geo-feature types). Experiments to employ features consisting of a binary value to record whether a place name or geo-feature were present and, separately, of a value that is the sum of the numbers of place names and geo-feature types, did not improve on sentence level performance and are not listed here. However, combining these latter data items with those in Kord-Geo-S did provide an improvement (referred to as feature set Kord-Geo-All) with an F1 for Geospatial of 0.643.

In addition to the three class classifiers we implemented several 2-class classifiers (see Table (4) with target classes of geospatial (*vs* spatial or neither), spatial *vs* (geospatial or neither) and neither (*vs* geospatial or spatial). Just as with the 3-class classifiers we used either just Kordjamshidi features (Kord), and place name and geographic features from the preposition's head words (Kord-Geo) and from the whole sentence in which the preposition occurred (Kord-GeoS). We also tested the method using Kord-Geo-All features, which gave the best 2-class performance for geospatial sense with an F1 of 0.621 but this did not improve on the result from the 3-class classifier. Output from the 2-class classifiers also included the complement of the Neither class, i.e. detection of prepositions that are either used in a spatial or a geospatial sense, which is equivalent to preposition classification task in [3]. We obtained an F1 value of 0.832 when using just the original features from [3].

As a baseline (Geo-Baseline-S) we implemented a Naive Bayes method for detecting whether a preposition has a geospatial sense, that uses, as machine learning features, just the presence of a place name and the presence of a geographic feature type. This was conducted at the preposition specific level, in which their presence was recorded only in the head words of the preposition, and at the level of whether they occurred anywhere in the sentence. The latter approach gave the better performance with an F1 of 0.502.

5 Conclusions and future directions

In this paper we have experimented with a method for detecting the geospatial nature of prepositions in sentences using a machine learning approach that was developed in [3] for generic spatial role labelling. Using a corpus of sentences annotated as either geospatial, spatial (but not geospatial) or neither geospatial nor spatial, we found that, when trained on this corpus, the original method was not able to detect geospatial prepositions with an F1 value greater than 0.50. However, it detected the spatial (but not geospatial) class with F1 of .745 and it detected prepositions that are used with either a geospatial or a spatial sense with an F1 of 0.832. We have adapted the method in an effort to improve its performance for detecting geospatial sense by adding features (for machine learning) that record whether a place name or a geospatial feature type is present in the head words that serve as subject and object of the preposition or, alternatively, whether they are present in the entire sentence. Using the sentence level features provided better performance with an F1 of 0.643 for geospatial sense. It also resulted in an improvement in detection of the spatial (but not geospatial) class with an F1 of 0.772. It may be noted that a classifier using only the presence of a place name or geographic feature type in the sentence provided better performance than the basic spatial role labelling method.

In future work we will investigate methods to make further improvements to the performance of the methods presented here. In particular we will address a limitation of the current method with regard to detection of place names and feature types by using a richer gazetteer and extending the dictionary of geographical feature types.

References

1   Richard Johansson and Pierre Nugues. LTH: Semantic Structure Extraction using Nonprojective Dependency Trees. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 227–230. Association for Computational Linguistics, 2007.

2   Arbaz Khan, Maria Vasardani, and Stephan Winter. Extracting Spatial Information From Place Descriptions. In *COMP@ SIGSPATIAL*, page 62, 2013.

3   Parisa Kordjamshidi, Martijn Van Otterlo, and Marie-Francine Moens. Spatial Role Labeling: Towards Extraction of Spatial Relations from Natural Language. *ACM Trans. Speech Lang. Process.*, 8(3):4:1–4:36, December 2011.

4   Ken Litkowski and Orin Hargraves. SemEval-2007 Task 06: Word-sense Disambiguation of Prepositions. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, SemEval '07, pages 24–29, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.

5   Fei Liu. *Automatic identification of locative expressions from informal text*. Masters by Coursework & Shorter thesis, University of Melbourne, Melbourne, Australia, 2013. URL: `http://minerva-access.unimelb.edu.au/handle/11343/38520`.

6   Fei Liu, Maria Vasardani, and Timothy Baldwin. Automatic Identification of Locative Expressions from Social Media Text: A Comparative Analysis. In *Proceedings of the 4th International Workshop on Location and the Web*, LocWeb '14, pages 9–16, New York, NY, USA, 2014. ACM.

7   V.B. Robinson. Individual and multipersonal fuzzy spatial relations acquired using human–machine interaction. *Fuzzy Sets and Systems*, 113(1):133–145, 2000.

8   Kristin Stock, Robert C. Pasley, Zoe Gardner, Paul Brindley, Jeremy Morley, and Claudia Cialone. Creating a Corpus of Geospatial Natural Language. In *Proceedings of the 11th International Conference on Spatial Information Theory - Volume 8116*, pages 279–298. Springer-Verlag New York, Inc., September 2013.

9   Jan Oliver Wallgrün, Alexander Klippel, and Timothy Baldwin. Building a Corpus of Spatial Relational Expressions Extracted from Web Documents. In *Proceedings of the 8th Workshop on Geographic Information Retrieval*, GIR '14, pages 6:1–6:8, New York, NY, USA, 2014. ACM.

10 M. Worboys. Nearness relations in environmental space. *International Journal of Geographic Information Science*, 15(7):633–651, 2001.

11 Xiaobai Yao and Jean-Claude Thill. How Far Is Too Far? – A Statistical Approach to Context-contingent Proximity Modeling. *Transactions in GIS*, 9(2):157–178, 2005.

Initial Analysis of Simple Where-Questions and Human-Generated Answers

Ehsan Hamzei 🄳
The University of Melbourne, Parkville, Victoria, Australia
ehamzei@student.unimelb.edu.au

Stephan Winter 🄳
The University of Melbourne, Parkville, Victoria, Australia
winter@unimelb.edu.au

Martin Tomko 🄳
The University of Melbourne, Parkville, Victoria, Australia
tomkom@unimellb.edu.au

— Abstract

Geographic questions are among the most frequently asked questions in Web search and question answering systems. While currently responses to the questions are machine-generated by document/snippet retrieval, in the future these responses will need to become more similar to answers provided by humans. Here, we have analyzed human answering behavior as response to *simple where questions* (i.e., *where* questions formulated only with one toponym) in terms of type, scale, and prominence of the places referred to. We have used the largest available machine comprehension dataset, MS-MARCO v2.1. This study uses an automatic approach for extraction, encoding and analysis of the questions and answers. Here, the distribution analysis are used to describe the relation between questions and their answers. The results of this study can inform the design of automatic question answering systems for generating useful responses to where questions.

1 Introduction

People frequently ask about geographic information in Web search [7, 13] and question answering systems [10]. Among many types of geographic questions, *where* (localization intention) and *how-to-get-to* (navigation intention) questions are dominant [3]. In everyday communication, these questions can be answered in terms of place and route descriptions, respectively. However, human-generated answers in a human-human dialogue are different from retrieved responses in human-computer interaction [1]. While in human-human question answering, one receives relevant responses with sufficient contextual information, current computer-based tools are not able to deliver answers of similar qualities [4]. In future, tools that provide responses similar to human-generated answers are envisaged instead of just retrieving documents and snippets [8]. For this, human answering behavior should be investigated as a major prerequisite.

As described in relevance theory of communication [15], people's answering behavior is based on the relevance of the answer to the question and to the context of communication. Relevance theory describes human-generated answers as simple, short, selective and cognitively

informative responses [15]. However, retrieved responses by computer-based tools and human-generated answers differ in both content and structure [6]. The retrieved documents/snippets may contain both relevant and irrelevant information regarding the question, and their structure (the flow of information) is not specifically designed to satisfy the inquirer's information need.

Geographic questions and specifically *where-questions* have special characteristics compared to other types of questions. The important influence of the inquirer's location on the relevance of answers has been noted [14]. Similarly, descriptive factors of places – their scale, type, and prominence – have a direct effect on the formation of the answers [14]. A thorough analysis of the relation between the questions and answers in terms of type, scale, and prominence of referred places is, however, still missing. Yet, this is an essential prerequisite for the understanding of the human question answering behavior. Here, we propose an approach to analyze the questions and answers based on type, scale, and prominence of places mentioned in their content. In short, we contribute:

- An encoding representation for the question and answers based type, scale, and prominence;
- Insights on the relation of type, scale, and prominence of places mentioned in the questions and answers by analyzing a large question answering dataset.

2   Data

MS-MARCO v2.1 [11] is a general purpose machine comprehension dataset provided by Microsoft. It contains question-answer pairs of the following types: (1) *numeric*, (2) *entity*, (3) *location* (including geographic questions [2]), (4) *person*, and (5) *description* [11]. Here, we focus on the MS-MARCO *location* records containing *simple where-questions* (questions with a single toponym) and their human-generated answers. Due to the lack of rich contextual information inside the *simple-where questions*, these questions are the base case and likely harder to be answered than *where-questions* with multiple toponyms.

3   Methodology

In this study, we defined a new representation encoding of the question and answer pairs by capturing their type, scale and prominence sequences, respectively. These sequences consist of values of the factors for the places referred to in the questions, followed by the values for the toponyms in the answers, ordered as they appear in the text. For example, the pair of question and answer *Where is Melbourne? In Victoria, Australia.* is encoded into a *type-sequence*: *{city, state, country}*.

Here, we first propose a process to extract, encode, and analyze the question/answer pairs. In the extraction step, toponyms from the questions and answers are extracted using both the Geonames and OpenStreetMap (OSM) Nominatim gazetteers. Next, the records extracted from the gazetteers are encoded to sequences of scale, prominence, and type. Finally, the relation between places in the questions and in their answers are investigated, using distribution analysis of the encoded sequences.

3.1   Extraction

The process starts by first filtering *location questions* that are started with *where* from the corpus. Then, the text is geoparsed for toponyms by matching against the gazetteers. Using parse tree information, noun phrases in the questions and their answers are checked against

the gazetteers starting from compound to simple noun phrases. Due to the characteristics of the extracted question/answer pairs (i.e., short texts, geographic where-questions, and localization information in the answers), every simple/compound noun phrase is considered as a toponym candidate. Finally, the ambiguity of toponyms in the pairs of corresponding questions and answers are resolved using map-based disambiguation techniques proposed in [9]. Consequently, the results of extraction are two gazetteers records (i.e., Geonames and OSM Nominatim records) for each extracted toponym in every pair of question and answer.

3.2 Encoding

To examine the relation between the structures of questions and their answers, three proxies have been defined for type, scale, and prominence of places, respectively. We have used toponym attribute information from gazetteers for this encoding. Sequence representations for each question and answer pairs are then generated based on these encoded values. To reduce the impact of gazetteers data incompleteness, only records which can have all extracted toponyms completely encoded into type, scale and prominence are further analyzed.

For type encoding, the Geonames schema of 667 place types (aka. feature codes) has been used without further changes[1]. The feature codes which are mentioned in the content of this paper are described in Appendix A.

A finite set of cognitively meaningful granularity levels is a prerequisite for encoding gazetteers records by scale. We have therefore adapted the seven-level schema from [12], with the granularity levels sequence of (1) furniture, (2) room, (3) building, (4) street, (5) district, (6) city, and (7) country. We have extended the schema to ten levels by adding coarser levels of scale: *county*, *state*, *country*, and *continent*. Nominatim records include an attribute (a number between 0–30) related to the OSM definition of scale (i.e., place_rank[2]). To convert the extracted gazetteers' records into the appropriate scale level, a look-up table linking OSM scale levels into the proposed scale schema has been devised manually.

Finally, we have used the *importance* attribute in the extracted Nominatim records as a proxy measure of place prominence. This value is estimated based on different factors, such as the frequency of the place appearances in Wikipedia[2]. The value ranges between 0 and 1, and it is designed to be used for ranking search results. To evaluate the prominence of places in questions and answers, we have classified these value into nine discrete levels of prominence by using the Jenks *natural breaks* method [5].

3.3 Distribution analysis

To investigate the relation between the questions and answers, we conducted distribution analysis of the encoded question/answer pairs. In distribution analysis, overall and sequence distributions are investigated and discussed. Overall distributions for questions and answers reveal the differences between places mentioned in questions, and places referred to in the answers. Sequence distributions show the distributions of values in each position of the encoded sequences (e.g., type sequences). The sequence distributions are used to investigate formation of the human-generated answers, in addition to their relations to the corresponding questions.

---

[1] `https://www.geonames.org/export/codes.html`
[2] `https://wiki.openstreetmap.org/wiki/Nominatim/Development_overview`
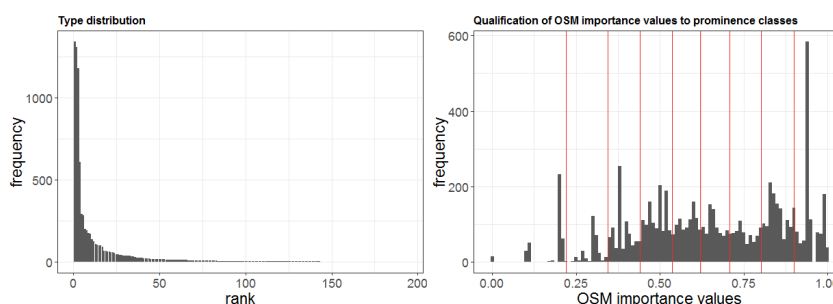
■ **Figure 1** Distributions of type and prominence of toponyms in the questions and answers.

4   Results and discussion

4.1   Extraction and encoding results

In the extraction process, 3238 simple where questions (from 31204 where questions) are found. Due to incompleteness of data in gazetteers in some cases the encoding into type, scale, and prominence cannot be done. Hence, during the encoding to type, scale, and prominence the number of records decreases by *22.5%* (2511 records out of 3238), *50.1%* (1587 records out of 3238), and *22.5%* (2511 records out of 3238).

After encoding the data, we find that only 185 unique place types out of 667 are referred to in the questions and answers. The frequency of these types forms a heavy long-tail distribution (Figure 1), where *81.6%* (i.e., 6072 out of 8218) of the extracted types belong to twenty unique types. This shows the reliance of people on few fundamental place types in the interpretation and answering of a large number of Web-based where-questions. In other words, the types in the corpus are biased in a way that a few types (e.g., states) are frequently observed, and a relatively large number of types (e.g., bridges) are found rarely in the dataset.

As shown in Figure 1, the *importance of places* extracted from Nominatim records are biased to medium and high values, which can be related to the geographic information people seek when they submit questions to search engines. The vertical lines in Figure 1 show the class breaks after classification of the continuous quantitative *importance* values into the nine levels of prominence.

4.2   Overall distribution analysis results

Figure 2 shows the distribution of the ten most-frequent place types in the questions and answers. Some types, such as *ADM1* (first-order administrative divisions), *PCLI* (independent political entities), and *RGN* (regions) are mostly used to formulate answers, while types such as *ADM3* (third-order administrative divisions), *ADM4* (fourth-order administrative divisions), *STM* (streams and rivers) and *PPL* (populated places, incl. villages and cities) are more frequently referred to in the questions. In other words, the distributions of type in the questions and their answers are systematically different. While lower-levels administrative divisions (e.g., ADM1) are frequently observed in human-generated answers, natural places (e.g., streams) and higher-levels administrative divisions (e.g., ADM4) are most frequently mentioned in the questions.

The type distribution is strongly related to the scale of the referents (Figure 3). While most of the questions are asked using place references at the *city* level of scale, they are answered at the *country*, *state*, and *county* levels. People are more searching for geographical-
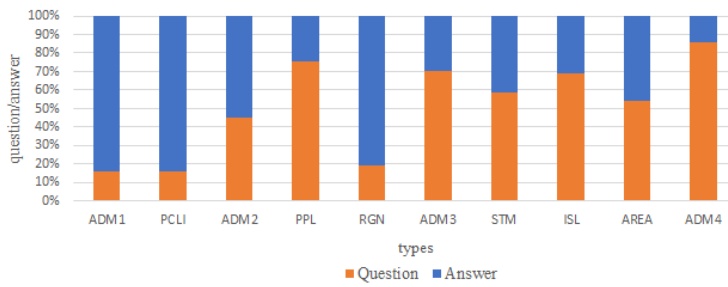
Figure 2 Distribution of types in the questions and answers, for the top ten most frequent types.
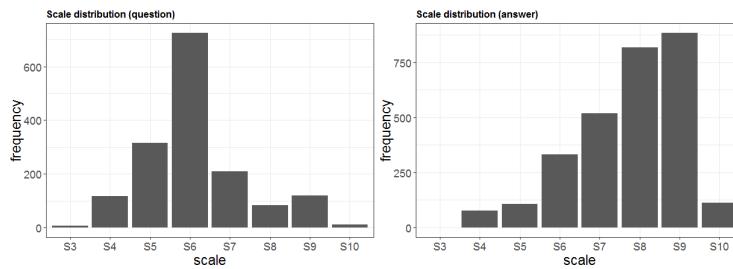
Figure 3 Scale distribution in the questions and answers.

scale places at the district and city levels of scale, while the answers to these questions are related to coarser levels such as country and state levels. We also note the lack of questions realting to fine-grained scale places. Similarly, Figure 4 shows the prominence distribution, centered around mid-range values for questions and biased to high-levels in answers. Two differences are, however, noticeable when comparing the distributions of scale and prominence. First, the coarsest level of scale (Level 10) is far less frequent than the highest level of prominence. The reason is that simple where question answers using continent level places references would be uninformative (i.e., of *low relevance*), while this is not the case for prominence (i.e., more prominent references are more relevant due to lower cognitive processing effort). Second, the overall distributions are similar in terms of skew (questions have positive skew, and answers have negative skew), however, their kurtosis is different (in both questions and answers, the distribution of scale is steeper than prominence). These patterns reveal that while scale and prominence may seem generally correlated, they capture distinct characteristics of places, with complex non-linear mapping between them. Evidently, the observed results are directly affected by the proxies used to capture type, scale and prominence.
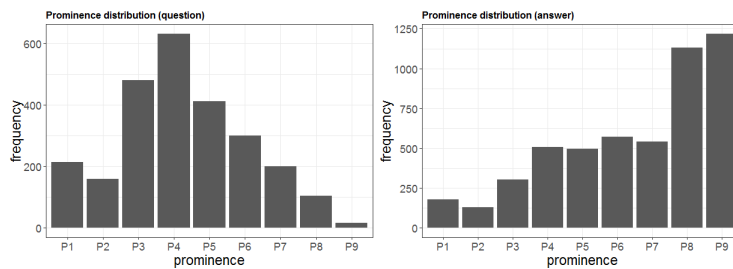
Figure 4 Prominence distribution in the questions and answers.
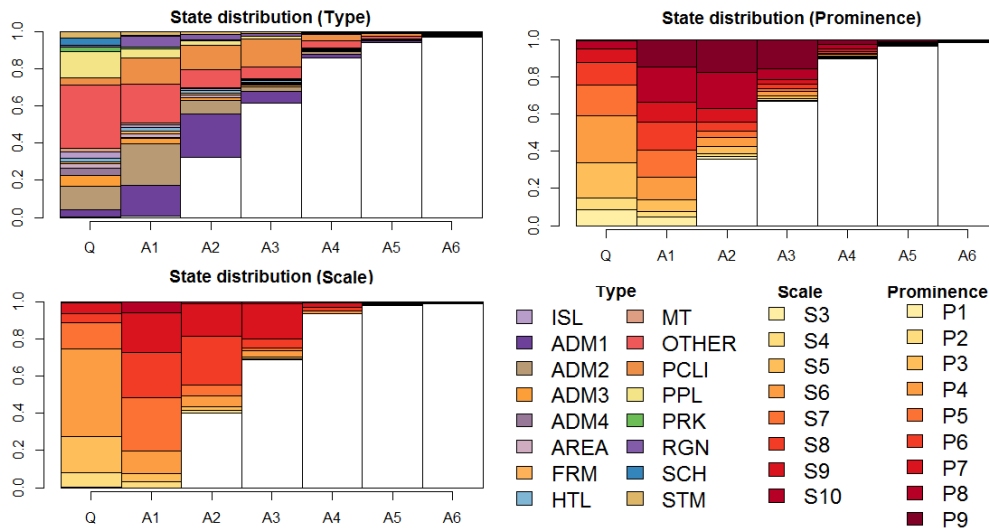
Figure 5 Sequence distribution of type, scale, and prominence.

4.3   Sequence distribution analysis results

Figure 5 shows the sequence distributions of type, scale, and prominence of places. In Figure 5 only the most frequent types are visualized in the sequence, and the rest are presented as *OTHER*. As the data contain only few answers with more than six toponyms (with a long tail distributions up to a maximum of 13 toponyms), we have focused only on the first six toponyms (capturing *94.3%* of the question-answer pairs). Most of the answers contain less than three toponyms. Figure 5 also reveals the differences between questions and answers in terms of type, scale, and prominence. Answers are formulated such that they start with lower values and end with higher values of both scale and prominence (fine to coarse, less to more prominent). In answers, certain type sequences are dominant: *ADM2* (e.g., Los Angeles County), *ADM1* (e.g., California), and *PCLI* (e.g., United States) are the most popular types in the first, second, and third positions of the answer-sequences, respectively. In general, the sequences of places which are mentioned in the answers are starting with less-known values in terms of type, scale, and prominence (i.e., low levels of scale and prominence, and particular types of places such as ADM2, and ADM3), and continue to well-known places in terms of these factors (i.e., higher levels of scale and prominence, and specific types of places such countries and political entities). In Appendix B, the patterns in scale and prominence sequences are investigated in more detail.

5   Conclusion

This paper presents a preliminary investigation of the relation of simple where questions and their human-generated answers. Type, scale, and prominence have been used as factors to investigate the human answering behavior of the simple where questions. We have proposed an approach for extracting, encoding and analyzing MS-MARCO question/answer records into type, scale, and prominence sequences. Later, we have discussed the relation based on overall and sequence distributions of these factors in the questions and their answers.

The results of this study show that human-generated answers to the questions follow a specific pattern starting from less-known values of type, scale, and prominence and continue to well-known places. This study reveals that type, scale, and prominence of places mentioned in questions has a direct relationship to formation of their answers. In summary, we have

shown that type, scale, and prominence are important factors which can be used to describe human answering behavior. Consequently, these factors can be used for mimicking human answering behavior to provide synthetic responses similar to human-generated answers.

This study shows the preliminary results of analyzing question answering data using type, scale, and prominence encoding. In future research, more research is needed to utilize and extend the proposed encoding approach to extract association rules from question answering datasets and to predict the structure of answers based on the encoding representation of the questions. In addition, the results of this study are limited to the context of Web search questions. Future work in other question/answering scenarios, especially contextualized human-human dialogue, lead to better understanding of human answering behaviour.

References

1   Christine Doran, John Aberdeen, Laurie Damianos, and Lynette Hirschman. *Comparing Several Aspects of Human-Computer and Human-Human Dialogues*, pages 133–159. Springer Netherlands, Dordrecht, 2003. `doi:10.1007/978-94-010-0019-2_7`.

2   Ehsan Hamzei, Haonan Li, Maria Vasardani, Timothy Baldwin, Stephan Winter, and Martin Tomko. Place questions and human-generated answers: A data analysis approach. In *Geospatial Technologies for Local and Regional Development*, pages 1–16, 2019.

3   Andreas Henrich and Volker Luedecke. Characteristics of Geographic Information Needs. In *Proceedings of the 4th ACM Workshop on Geographical Information Retrieval*, GIR '07, pages 1–6, New York, NY, USA, 2007. ACM. `doi:10.1145/1316948.1316950`.

4   Lynette Hirschman and Robert Gaizauskas. Natural language question answering: the view from here. *Natural Language Engineering*, 7(4):275–300, 2001.

5   G. F. Jenks. The Data Model Concept in Statistical Mapping. *International Yearbook of Cartography*, 7:186–190, 1967.

6   Karen Sparck Jones. *Retrieving information or answering questions?* London : British Library. (British Library Annual Research Lecture ; 8), 1990.

7   Emilia Kacprzak, Laura Koesten, Jeni Tennison, and Elena Simperl. Characterising Dataset Search Queries. In *Companion Proceedings of the The Web Conference 2018*, WWW '18, pages 1485–1488, Geneva, Switzerland, 2018. International World Wide Web Conferences Steering Committee. `doi:10.1145/3184558.3191597`.

8   Oleksandr Kolomiyets and Marie-Francine Moens. A survey on question answering technology from an information retrieval perspective. *Information Sciences*, 181(24):5412–5434, 2011. `doi:10.1016/j.ins.2011.07.047`.

9   Jochen L. Leidner, Gail Sinclair, and Bonnie Webber. Grounding Spatial Named Entities for Information Extraction and Question Answering. In *Proceedings of the HLT-NAACL 2003 Workshop on Analysis of Geographic References - Volume 1*, HLT-NAACL-GEOREF '03, pages 31–38, Stroudsburg, PA, USA, 2003. Association for Computational Linguistics. `doi:10.3115/1119394.1119399`.

10   Xin Li and Dan Roth. Learning question classifiers: the role of semantic information. *Natural Language Engineering*, 12(3):229–249, 2006.

11   Tri Nguyen, Mir Rosenberg, Xia Song, Jianfeng Gao, Saurabh Tiwary, Rangan Majumder, and Li Deng. MS MARCO: A human generated machine reading comprehension dataset. arXiv preprint arXiv:1611.09268, 2016.

12   Daniela Richter, Stephan Winter, Kai-Florian Richter, and Lesley Stirling. Granularity of locations referred to by place descriptions. *Computers, Environment and Urban Systems*, 41:88–99, 2013. `doi:10.1016/j.compenvurbsys.2013.03.005`.

13   Mark Sanderson and Janet Kohler. Analyzing geographic queries. In *SIGIR Workshop on Geographic Information Retrieval*, volume 2, pages 8–10, 2004.

14   Benny Shanon. Answers to where-questions. *Discourse Processes*, 6(4):319–352, 1983.

15   Deirdre Wilson and Dan Sperber. Relevance theory. In *Handbook of Pragmatics*. Blackwell, 2002.

A   Feature codes

Table 1 below shows the types which are mentioned in paper. The complete list can be found in the Geonames website.

Table 1 Feature codes used in the paper (extracted from Geonames documentation).

Code	Description	Example
ADM1	first-order administrative division (states, and provinces)	Oklahoma
ADM2	second-order administrative division (counties)	Brevard County
ADM3	third-order administrative division (cities)	City of Alhambra
ADM4	fourth-order administrative division (towns)	Newburgh
AREA	a part of land without homogeneous character/boundaries	Theresienwiese
FRM	a part of land dedicated to agricultural purposes	Branksome
HTL	hotels	The Carriage House
MT	mountains	Eagles Nest
PCLI	independent political entity	Paraguay
PPL	diverse type of populated places (e.g., cities, and villages)	El Granada
PRK	parks and recreational places	Franklin Square Park
RGN	an area with particular cultural character	Central Africa
SCH	schools and universities	Stuyvesant High School
STM	streams	Withlacoochee River

B   Differential scale and prominence sequences

Figure 6 shows the hierarchical (i.e., zooming in, zooming out), and non-hierarchical patterns in scale and prominence sequences using differential sequences. The differential sequences are created by subtracting values from their previous values in the scale and prominence sequences. Due to the fact that scale and prominence are ordinal values, the subtraction values are not valid, and consequently using *sign* function the values are translated into meaningful values – i.e., **0** (equal), **+** (greater than) and **−** (less than). Here, **0** values show the *non-hierarchical* pattern because the scale or prominence levels are not changed. The **+** values show the *zooming out* pattern, because the level of scale or prominence is increased compared to its previous level in the sequence. The **−** values show the *zooming in* pattern with same rationale. Figure 6 supports the discussion made in the paper, section 4.2, that values in the scale and prominence sequences are hierarchically structured starting with lower values (levels) followed by higher ones.
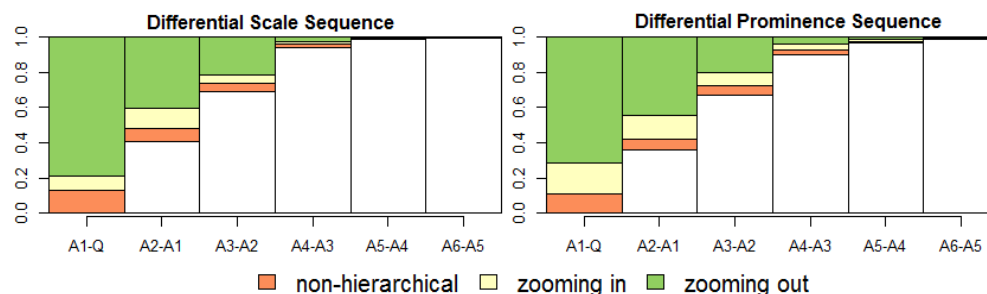
Figure 6 Sequence distribution of differential scale and prominence sequences.

Spatial Information Theory and Construction Informatics – a Fruitful Symbiosis

André Borrmann

Chair of Computational Modeling and Simulation, Technical University of Munich, Germany
https://www.cms.bgu.tum.de
andre.borrmann@tum.de

Abstract

Traditionally, Spatial Information Theory and Construction Engineering have been recognized as widely separated fields with only very little connections. However, in recent years the construction industry has undergone a substantial change: It is evolving from rather historic practices based on 2D drawings into modern digital processes based on information-rich 3D models that can be generated, analyzed and processed by means of computer technology. This progression, driven and fostered by the Construction Informatics Community, opens the possibility for innovative research in the fascinating area where Spatial Information Theory and Construction Engineering overlap. The paper gives an overview on ongoing activities in this area.

1 Introduction

Traditionally, Spatial Information Theory and Construction Engineering have been recognized as widely separated fields with only very little connections. However, in recent years the construction industry has undergone a substantial change: It is evolving from rather historic practices based on 2D drawings into modern digital processes based on information-rich 3D models that can be generated, analyzed and processed by means of computer technology. This progression has open the possibility for innovative research in the fascinating area where Spatial Information Theory and Construction Engineering overlap [1].

The paper covers a number of of research topics positioned at exactly this overlapping area.

2 Spatial analysis of 3D building and city models

The journey starts at presenting the 3D spatial analysis functionalities for building information models developed by the author [3, 4, 5]. Here, classical concepts of Spatial Information Theory originally developed for 2D Geographical Information Systems (GIS) have been transferred and applied to 3D building information models. As one example where spatial query functionalities are of great benefit, the automated generation of a precedence relationship graphs (PRG) from building information models is presented. The PRG forms an important component for automated construction progress monitoring, where point clouds are captured in regular time intervals and overlayed with 4D building models [23]. As however, many components remain invisible, the PRG helps to infer their existence in an indirect manner [7, 6]. Also in other scenarios spatial analysis provides powerful means for effective filtering [12, 9, 10]. This includes checking the spatio-semantic consistency of

building information models [8], but also the integrative analysis of both 3D city and building models [11] across heterogenous data models, such as CityGML and IFC (Figure 1).
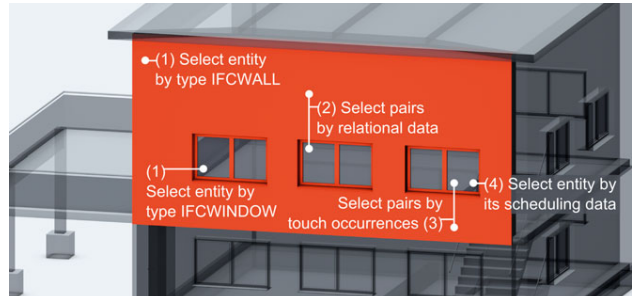
Figure 1 Querying building models by applying spatial operators, from [11].
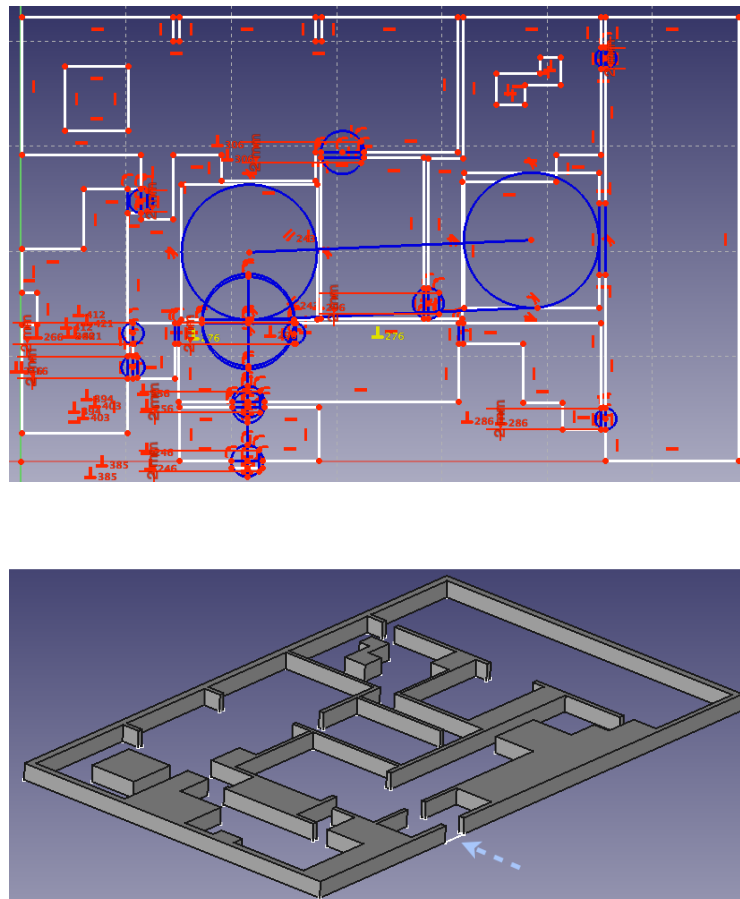
Figure 2 Interactive parametric FreeCAD sketch that maintains visibility, movement, and qualitative size constraints and the building model extruded from it. From [21].

3 Building design supported by spatial reasoning

The formal description of qualitative spatial relationships is helpful not only for analyzing completed designs and querying building information models, but also for supporting the architectural design process. A particularly powerful solution in this context is the integration of formal spatial reasoning with a feature-based parametric modeling engine [21]. We were able to demonstrate the proposed methodology by applying it to architectural floor plan layout design, where a number of spaces with well defined functionalities are automatically arranged such that particular functional design constraints are maintained (see Figure 2).

Recently, model synthesis, i.e. the automated creation of models on the basis of abstract engineering knowledge has received increasing attention [24, 22]. In this regard, the author's group has been successfully applying formal graph transformation techniques to realize the knowledge-driven detailing of building components [25]. To this end, parametric modelling engines were coupled with graph transformation systems. The talk will give an overview on the progress achieved and discuss the remaining challenges.
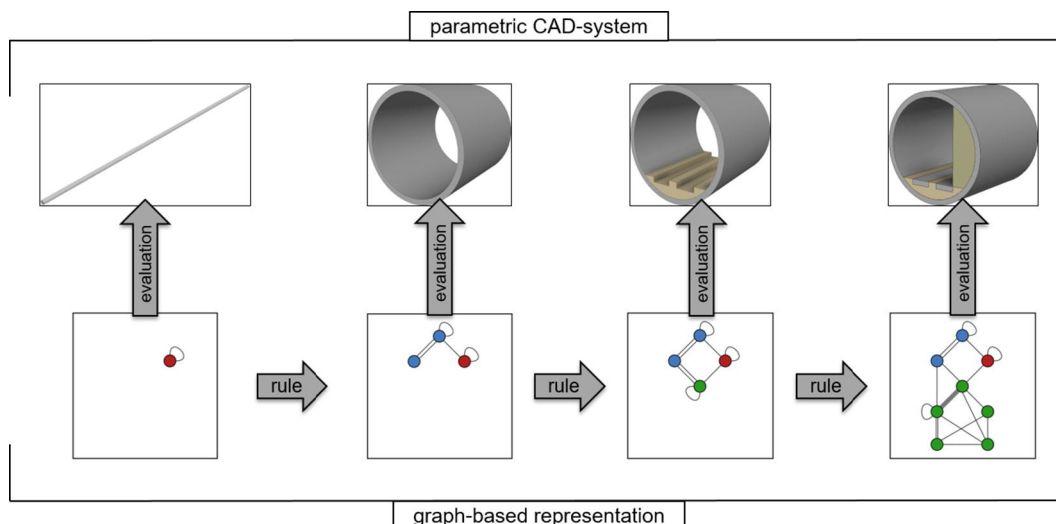
Figure 3 Application of formal graph transformation for automated detailing of tunnel models, from [25].

4 Semantic enrichment

An extremely important field of research is related to capturing the already built assets of the built environment. Here, the goal is to develop methods that allow to create a semantically rich digital representation from the raw data of point clouds and photographs in a largely automated manner. In this regard, the concept of semantic enrichment provides the possibility to assign volumetric 3D models with the respective semantics [20]. The applied rule-based approach for semantic enrichment heavily relies on spatial relationships between individual objects. In a related application field, rules with spatial semantics can be formulated to identify site equipment required for construction projects, for example [13].

	Element Group	Element description	#	obj1		Deck Superstructure			
				Bridge					
				Bridge	Primary Girder	Capping beam	Transverse beam	Deck slab panel
					111		201	301
	Bridge	Bridge		0000000000000yy00	xnxnnyxxnnnynxnynn	xxnnnnxxnnnnyxnyxn	nnnnnnxxnxnnxxnynn	xxxnnnxxnnnnynynnn
Deck Superstructure		Primary Girder	111	xnxnnyxxyy0000ynyy	nnnnnyxxxnynxnnnx	nnnnnnnxxnnyxnnnx	yynnnnxnnnnxxnnxx	yxxnynyynnnynynnn
		Capping beam		xxnnnnxxyy0000ynyy	nnnnnyxxnynxnnnx	nnnnnyyxxnnynxnnx	nnnnnxyynnnnxxnnnx	nnnnnnnynnnynnnn
		Transverse beam	201	nnnnnnxxyy0000ynyy	**yynnnnnxxyynynxnnxx**	nnnnxynyynnyxnnyx	nnnnnyyxxnxnxnnxx	nnxnxnnynxnynnnxn
		Deck slab panel	301	xxxnnnxxny0000ynyy	yxxynnynnynynxnnny	nnnnnnnnynnyxnnyx	nxnxnnnnnxnnxxnnxy	xxxnnnyxnxnynnnxx
obj2	Substructure	Bearing		nnnnnnxxny0100ynyy	ynnnnnxynynynxnnyy	xnnxnnxnnynnynxnny	nnnnnnxcnynnxxnnxy	nnnnnnxynynynynnnx
		Plinth (capping beam)		xnnnnnxxny0000ynyy	nnnnnyxynynynxnny	ynnynnxnnynynxnnxy	nnnnnnxnxnnxxnnxy	nnnnnnxynxnynnnxx
		Plinth (abutment)		xnnnnnxxny0000ynyy	nnnnnyxynynxnnnx	nnnnnnxnnynynxnnyx	nnnnnnxnxnnxxnnxy	nnnnnnynxnynnnxx
		Shear Key (capping beam)		xxnnnnxxny0000ynyy	nnnnnnxxnynynxnnnx	ynnxnnnxnynynxnnyx	nnnnnnxxnxnnxxnnxy	nxnnnnnynnnynnnyn
		Shear Key (abutment)	401	xxxnnnnxxny0000ynyy	nnnnnnxxnynynxnnnx	nnnnnnxxnynnynxnnyx	nnnnnnxxnxnnxxnnxy	nnnnnnynxnynnnyn
		Abutment	403	xxxnnnnxxyy0000ynxy	nnnnnnnxxnynnxnnnx	nnnnnnxyxxxnnyxnnxx	nnnnnyxnxnnxxnnxy	nnxnnnnynnnynnnn
		Column	405	xxnynnnxxyy0000ynxy	nnnnnnnxxnynynxnnx	nnnnnnxynyynynxnnyx	nnnnnnnynnnnxxnnxy	nnnnnnynnnnxnnnnn
Non-structural		Lamp post		xxxnynnxxxy0000ynyy	nnnnnnnxynynxnnnxn	nnnnnnnxynynynxnnyn	nnnnnnnxnxnnxxnnxy	nnnnnnnnxnynnnnn
		Safety barrier	507	xxxnxyxxyy0000ynyn	nnnnnynxynynxnnxx	nnnnnxnxynnyxnnyx	nnnnnnxxnnxxnnxy	nnnnnnynnxnynnnyn

Figure 4 Condition matrix for identifying the semantics of geometric objects of scanned bridges, from [20].

5 Code compliance checking

Another field of application is code compliance checking. In the design and engineering of buildings, a large number of building codes and regulations have to be taken into account. Today, the compliance of building designs with such regulations is checked manually; both by the responsible architects as well as the building permission authorities. The available commercial solutions for code compliance checking mainly follow a black-box approach where the rules that make up a certain regulation are implemented in a hard-wired fashion rendering their implementation in-transparent and non-extendable.

A number of researchers have tackled this problem and have proposed various ways that allow the user to define rules, either in a standard programming language or in a dedicated language. However, AEC domain experts usually do not have the required programming skills to use these languages appropriately. To overcome this issue, we developed the Visual Code Checking Language (VCCL), which uses a graphical notation in order to represent the rules of a code in a machine- and human-readable form [18, 19]. As spatial relationships play an important role in code compliance checking, VCCL provides dedicated operators with spatial semantics among its basic building blocks (Figure 5).
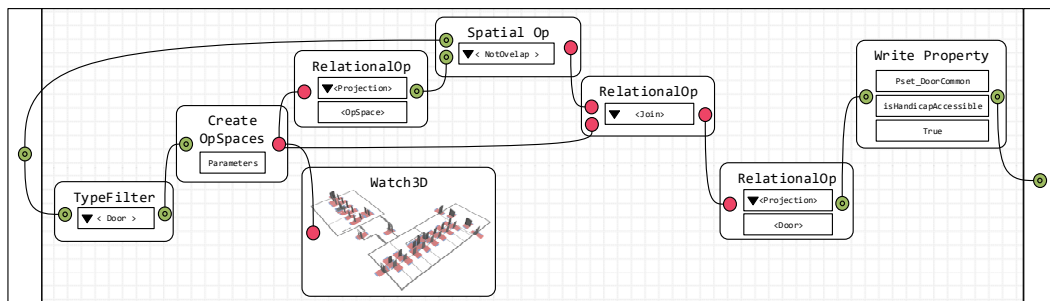
Figure 5 VCCL program for checking a building model for compliance with accessibility regulations, from [19].

6 Spatial cognition for pedestrian dynamics

Finally, we highlight the importance of spatial information for the simulation of pedestrians. In particular on the strategic and tactical levels of the simulation models, the proper modelling of wayfinding behavior is of utmost importance for achieving correct simulation results [2, 16, 17, 15, 14]. We demonstrate how elements of spatial cognition have been implemented in our pedestrian simulator MomenTUM (Figure 6) and illustrate their effect.
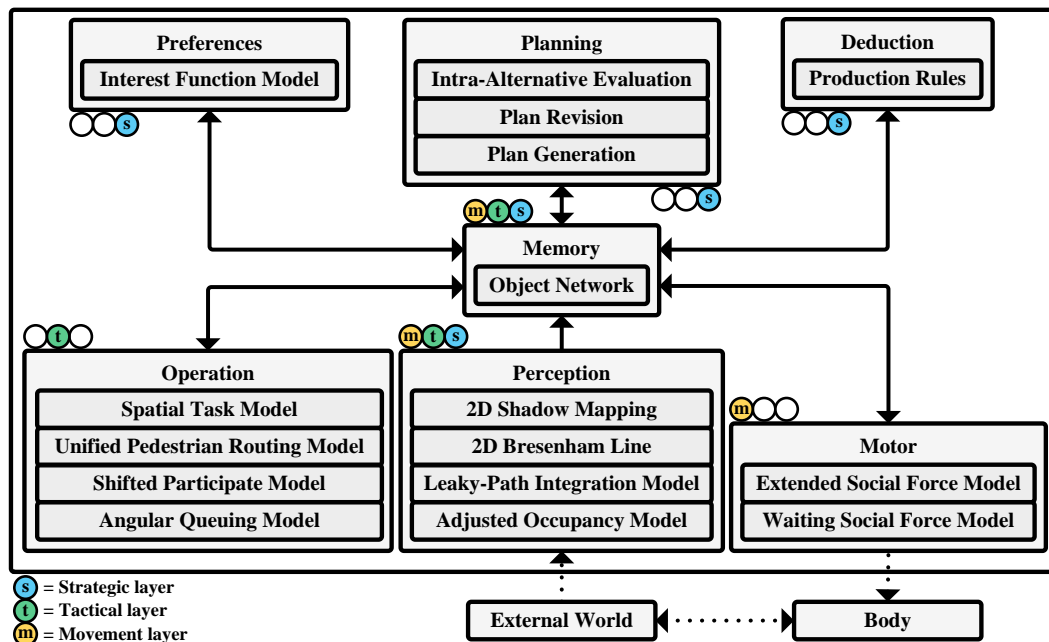
Figure 6 Architecture of the pedestrian dynamics simulator MomenTUM that relies on cognitive principles for modeling wayfinding behavior, from [14].

7 Conclusion

There are various fields of research where the combination of Spatial Information Theory and Construction Informatics results in strong synergies and enables new solutions for practical problems of the Architecture, Engineering, and Construction (AEC) industry. There is high potential for intensified research in this area.

References

1 M. Bhatt, A. Borrmann, R. Amor, and J. Beetz. Architecture, Computing, and Design Assistance. *Automation in Construction*, 32:161–164, 2013. `doi:10.1016/j.autcon.2013.01.001`.

2 D. H. Biedermann, P. M. Kielar, A. M. Riedl, and A. Borrmann. Oppilatio+ - A data and cognitive science based approach to analyze pedestrian flows in networks. *Collective Dynamics*, 1:1–30, 2016. `doi:10.17815/CD.2016.9`.

3 A. Borrmann and E. Rank. Specification and implementation of directional operators in a 3D spatial query language for building information models. *Advanced Engineering Informatics*, 23(1):32–44, 2009. `doi:10.1016/j.aei.2008.06.005`.

4   A. Borrmann and E. Rank. Topological analysis of 3D building models using a spatial query language. *Advanced Engineering Informatics*, 23(4):370–385, 2009. `doi:10.1016/j.aei.2009.06.001`.

5   A. Borrmann, S. Schraufstetter, and E. Rank. Implementing Metric Operators of a Spatial Query Language for 3D Building Models: Octree and B-Rep Approaches. *Journal of Computing in Civil Engineering*, 23(1):34–46, 2009. `doi:10.1061/(ASCE)0887-3801(2009)23:1(34)`.

6   A. Braun, S. Tuttas, A. Borrmann, and U. Stilla. A concept for automated construction progress monitoring using BIM-based geometric constraints and photogrammetric point clouds. *ITcon*, 20(8):68–79, 2015.

7   A. Braun, S. Tuttas, A. Borrmann, and U. Stilla. Automated progress monitoring based on photogrammetric point clouds and precedence relationship graphs. In *The 32nd International Symposium on Automation and Robotics in Construction and Mining (ISARC 2015)*, Oulu, Finnland, 2015.

8   S. Daum and A. Borrmann. Checking Spatio-Semantic Consistency of Building Information Models by Means of a Query Language. In *Proc. of the International Conference on Construction Applications of Virtual Reality*, 2013.

9   S. Daum and A. Borrmann. Simplifying the Analysis of Building Information Models Using tQL4BIM and vQL4BIM. In *Proc. of the EG-ICE 2015*, Eindhoven, The Netherlands, 2015.

10   S. Daum and A. Borrmann. Enhanced Differencing and Merging of IFC Data by Processing Spatial, Semantic and Relational Model Aspects. In *Proc. of the 23rd International Workshop of the European Group for Intelligent Computing in Engineering*, Kraków, Poland, 2016.

11   S. Daum, A. Borrmann, and T. H. Kolbe. A Spatio-Semantic Query Language for the Integrated Analysis of City Models and Building Information Models. In A. Abdul-Rahman, editor, *Advances in 3D Geoinformation*. Springer International Publishing, 2016.

12   S. Daum, A. Borrmann, C. Langenhan, and F. Petzold. Automated generation of building fingerprints using a spatio-semantic query language for building information models. In *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2014*, Vienna, Austria, 2014.

13   Katrin Jahr and André Borrmann. Semi-automated site equipment selection and configuration through formal knowledge representation and inference. *Advanced Engineering Informatics*, 38:488–500, 2018. `doi:10.1016/j.aei.2018.08.015`.

14   P. M. Kielar and Borrmann A. Spice: A Cognitive Agent Framework for Computational Crowd Simulations in Complex Environment. *Autonomous Agents and Multi-Agent Systems*, pages 1–30, 2018. `doi:10.1007/s10458-018-9383-2`.

15   Peter M. Kielar, Daniel H. Biedermann, Angelika Kneidl, and André Borrmann. A unified pedestrian routing model for graph-based wayfinding built on cognitive principles. *Transportmetrica A: Transport Science*, 14(5):406–432, 2017. `doi:10.1080/23249935.2017.1309472`.

16   Peter M. Kielar and André Borrmann. Coupling Spatial Task Solving Models to Simulate Complex Pedestrian Behavior Patterns. In *Proc. of the 8th Conference on Pedestrian and Evacuation Dynamics*, 2016.

17   Peter M. Kielar and André Borrmann. Modeling pedestrians' interest in locations: A concept to improve simulations of pedestrian destination choice. *Simulation Modelling Practice and Theory*, 61:47–62, 2016. `doi:10.1016/j.simpat.2015.11.003`.

18   C. Preidel and A. Borrmann. Towards code compliance checking on the basis of a visual programming language. *ITcon*, 21:402–421, 2016.

19   C. Preidel, S. Daum, and A. Borrmann. Data retrieval from building information models based on visual programming. *Visualization in Engineering*, 5(18):1–14, 2017. `doi:10.1186/s40327-017-0055-0`.

20   R. Sacks, L. Ma, R. Yosef, A. Borrmann, S. Daum, and U. Kattel. Semantic Enrichment for Building Information Modeling: Procedure for Compiling Inference Rules and Operators for Complex Geometry. *Journal of Computing in Civil Engineering*, Vol. 31(Issue 6 (November 2017)), 2017. `doi:10.1061/(ASCE)CP.1943-5487.0000705`.

21 C. Schultz, M. Bhatt, and A. Borrmann. Bridging qualitative spatial constraints and feature-based parametric modelling: Expressing visibility and movement constraints. *Advanced Engineering Informatics*, 2015. `doi:10.1016/j.aei.2015.10.004`.

22 Dominic Singer, Maximilian Bügler, and André Borrmann. Knowledge based Bridge Engineering - Artificial Intelligence meets Building Information Modeling. In *Proc. of the EG-ICE Workshop on Intelligent Computing in Engineering*, Krakow, Poland, 2016.

23 S. Tuttas, A. Braun, A. Borrmann, and U. Stilla. Acquisition and Consecutive Registration of Photogrammetric Point Clouds for Construction Progress Monitoring Using a 4D BIM. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(1):3–15, 2017. `doi:10.1007/s41064-016-0002-z`.

24 S. Vilgertshofer and A. Borrmann. A Graph Transformation Based Method for the Semi-Automatic Generation of Parametric Models of Shield Tunnels. In *Proc. of the EG-ICE Workshop on Intelligent Computing in Engineering*, Krakow, Poland, 2016.

25 S. Vilgertshofer and A. Borrmann. Using graph rewriting methods for the semi-automatic generation of parametric infrastructure models. *Advanced Engineering Informatics*, 33:502–515, 2017. `doi:10.1016/j.aei.2017.07.003`.

A Case for Geographic Masses

Brandon S. Plewe
Department of Geography, Brigham Young University, Provo, Utah, United States
http://carto.byu.edu
plewe@byu.edu

──── **Abstract** ────

Geographic masses, the stuff we deal with that cannot be categorized as geographic objects, comprise a crucial but largely unrecognized component of the core ontology of geographic information. Although masses have been rarely acknowledged in GIScience, they appear in geographic discourse just as often as objects. A concise but consistent formal definition of a geographic mass particular, which distinguishes a mass from an object, can be applied to any endurant phenomena, enabling a richer understanding of the geographic milieu, and more informed decision making during modeling and analysis processes.

1 Introduction

There appears to be a hole in the core ontology of Geographic Information Science. What has become the conventional wisdom in our understanding of the world and how to represent it, including objects and fields, space and time, processes and attributes, is missing a substantial class of phenomena.

In my introductory GIS courses, when I present the classic core principle of the object and field views of the world, I am sometimes asked, "what about *stuff* like water, petroleum, vegetation, suburbia, and so on? Are they objects or fields?" My typical answer, "neither," is not very satisfying to the average college student. What should we do with these? Do they matter?

In other fields, such as philosophy and linguistics, these phenomena are most often called *masses*, [29] and they have been extensively (if incompletely) studied. The purpose of this paper is not to reinvent the concept, but to answer the question, "are masses relevant to geographic inquiry and geographic information science?" If so, I will further develop an understanding of masses in a geographic context, and how they can be incorporated into the core ontology of geographic information.

Over the years, the GIScience community has flirted with mass phenomena. Couclelis comes close in mentioning "extensive entities" that do not fit into the classic object model, [5] and Peuquet acknowledges the difference between "continuous properties" (fields) and "continuous matter" (masses) without saying much about the latter [22]. Galton acknowledges the existence of masses as an aside, without incorporating them into his geo-ontology [6]: later, outside the GIScience realm, Galton does incorporate "material" (masses) into an ontological framework, but does not fully develop a theory thereof [7].

Smith and Mark mention "stuffs," but they express doubt that stuffs are relevant to geographic phenomena before they start their study, and they rarely appear in the resultant lists of typical geographic phenomenon [28]. However, this is at least partially due to the fact that the phrasing of their questionnaire is object-centric.[1]

2   What is a Mass?

The most common way the literature distinguishes phenomena as objects and masses is to simply say, "objects are things and masses are stuff," [4, 29] or to resort to examples: a building is an object, while the metal, wood, and concrete that comprise it are masses. The *Oxford English Dictionary* contains a number of definitions of *mass*, including some with closely related senses:

- A dense aggregation of objects having the appearance of a single, continuous body.
- A coherent body of matter of unspecified or indeterminate shape, and usually of relatively large bulk; a solid and distinct object occupying space.
- A large amount, number, or quantity of a thing or things, material or immaterial [1].

These and other definitions contain two basic characteristics of masses that differentiate them from prototypical objects: 1) they are *amorphous*, without regard for a defined shape or boundary; 2) they appear to be *continuous*, without regard for discrete parts [18].

This may seem precise, but definitions like the above have proven to be too vague to apply to many real phenomena, and contain two apparent contradictions that have vexed scholars for decades: 1) in the first definition, how can it simultaneously contain objects but appear continuous? and 2) how can a body/amount/quantity not have a shape or boundary? To develop the concept of a mass in geographic inquiry, these and other issues will need to be resolved.

A formal definition of a mass, if it can resolve these issues, should be more operational than these vague definitions. That is, it should help us more clearly distinguish objects and masses in edge cases, and know what we should subsequently do with them. While several formal general ontologies have been developed that include masses as a category [8, 27, 21, 13], I have yet to find one that fully formalizes the definition of a mass. So, I will try.

3   Ontological Framework

This work must fit within a general metaphysical and ontological framework. Given that multiple contradicting philosophies have been proposed and debated for hundreds of years, and none has emerged victorious, and because I have no interest in inventing yet another one, I will simply select the existing frameworks that make the most sense to me.

Firstly, and least controversially, we must distinguish between universals (kinds of phenomena) and particulars (individual phenomena); particulars may be thought of as instances of universals. Geospatial technology focuses on representing particulars, and that will be the primary focus here, but let us start by declaring a universal as follows:

---

[1] Their study gave each respondent one of five prompts to list geographic phenomena that came to mind. However, all five prompts were based on count nouns: "a feature," "an object," "a concept," or "something." They also note that fields appeared in the results even less often than stuffs. This is probably not intentional bias on the authors' part, because the prompts seem rather generic unless one were specifically looking for fields and masses. At best, the results of this study show that fields and masses are not so dominant in common-sense geography that participants woud think of them despite the wording of the prompts.

X (upper-case letter): A universal type of phenomenon. This could be as simple as a common noun, like "sheep" or "mountain;" but I relax the common definition to also include more specific concepts or forms of reference, like "those mountains" or "three sheep." In the latter case, note that *X* is not three actual sheep, just the notion of a group of three sheep.

Particular phenomena are more of a challenge, as there is some debate on the nature of their existence [25]. Poli, building on earlier theories, distinguishes three "strata of reality," with regards to a phenomenon being studied [23]: the *material* (the world as it actually exists independently of humanity), the *mental-psychological* (how an individual human conceptualizes the world), and the *social* (how people collectively organize the world through mechanisms such as language, institutions, or maps). Ontologists have tended to divide into camps according to which of these three strata they believe to be fundamental ("really real"), with the others seen as derivative, unstructured, or nonexistent: realists favor the material stratum, conceptualists the mental-psychological, and nominalists the social. We might call this tendency *stratum exclusivity*.

Furthermore, the scholarly dialogue concerning masses has followed three tracks that roughly correspond to Poli's strata: physics or metaphysics, cognition, and linguistics or semantics. By far, the latter approach has been the most common, probably because language structures are easier to access and study than mental ideas and physical reality. Frequently, studies have mixed the approaches, assuming a strong correspondence between physical masses, mass concepts, and mass terms, and using that assumption to make an argument or conclusion about one of the levels based on a characteristic of another level. We might call this *stratum conflation*. Laycock and Bunt both lament this tendency, and caution scholars to focus on one or the other [17, p.12] [2, p.49]. That is, evidence from any of the three realms may be suggestive of the nature of phenomena in another realm, but not proof.

I will attempt to be metaphysically neutral in keeping with Gracia, who acknowledges that all three strata, and all three lines of inquiry, can have equal validity, depending on the situation [11, p.199-205]. Such a neutral or hybrid stance is reminiscent of Lakoff's experiential realism, in which knowledge is equally influenced by reality, personal experience, and social experience [16]; as well as the post-positivist approach to scientific epistemology. It is also suggested in Herre's fourth *phenomenal* stratum of reality, consisting of phenomena that may or may not be "real" in the material-stratum sense, but are at the very least heavily motivated by real-world conditions; they may also be concepts, but are standardized by society and language to such a degree that we all recognize the same phenomenon, so they are indistinguishable from real [14, p.7]. For example, it may not be important whether a tree really exists as a distinct object, or as a conceptualization of sensory perceptions of an inherently unorganized reality, or as a term created by society to categorize an uncategorized reality, as long as I can point at something and we all agree that it is a tree. This all sounds like a lot of the phenomena we represent in GIS.

We can mediate the above strata and lines of inquiry by making a formal distinction between them:

a (lower-case letter): A particular geographic phenomenon in the material stratum. *Phenomenon* is defined very broadly to include anything that might be a subject of interest, with no restriction on its existence, nature, complexity, or spatiotemporal extent. *Geographic* limits the view to phenomena that occur or exist somewhere on Earth, at a medium to small scale (i.e., it could be shown on a map).

$a : X$ *a categorized phenomenon*, in which it is assigned to a universal through mental and/or social processes.[2] This can be read as "a–as–X." This is only a claim, which may or may not be valid, hence the following:

Xa categorizability as a predicate: "a is (a/an) X," or more precisely, "a can be considered (a/an) X." This is a predicate, not a set theoretic membership ($a \in X$), because there is no requirement that X have an extension (a set of individuals).

Representing both the real-world phenemonon, and our conceptualization(s) thereof, as separate but equal entities, may be able to resolve much of the ontological debate described above. One could say that a realist believes in an extremely strong correspondence between a and $a : X$, the latter being a simple derivative of the former. Conversely, the nominalist could be said to believe in a very weak correspondence, such that a is not attainable from a study of $a : X$, and subsequently focuses solely on the latter. In fact, there is a continuum of correspondence: there are likely phenomena that can only be considered a single way, while others can be categorized in a number of different ways, and there are some in which the concept is only loosely based on real-world conditions.[3]

In terms of spatio-temporal ontology, objects and masses are both *endurants* (hereafter $End(a : X)$), as opposed to occurrents (processes and events). An endurant is informally defined in most top-level ontologies [8, 27, 21][4] as a phenomenon that endures through time; it is recognizable as a complete entity at any time during its existence. For example, at any moment, a tree is still recognizable as a tree; over time, it may change, but it is the same tree.

Although I am not doing a full formalization of temporal nature, we do need some definitions that place endurants in space and time, based on Simons (but with my own symbols) [26, p.132]:

$S(a)$ the *footprint* of a, the minimal region of space in which it exists.

t any period of time, including intervals and moments of zero duration.

$T(a)$ The *lifespan* of a phenomenon, the period of time during which a exists.

$F_t a$ A *temporal predicate*, a claim that something about a is true throughout t [26, p.130].

a_t A *temporal restriction*, a as it exists at t, whether all of a, part of it, or none of it. This is formally defined by:

D1 $\forall t(F_t a_t \iff F_t a)$

Anything that is true of a during t is also true of a_t, and vice versa, including such predicates as part-whole relationships, attributes, and even existence. This does not require that t be part of $T(a)$; for any times outside its lifespan, a_t is empty.

3.1 Objects vs. Masses

The continuous and amorphous nature of a mass is manifested mereologically as *homogeneous reference*, perhaps first and best explained by Quine based on concepts from Goodman:

---

[2] I am not distinguishing here between the mental and social strata, or between concepts and words. I am definitely not asserting that they are indistinguishable, and a distinction may be useful in the future, but for the formal definitions presented in this paper, the difference did not turn out to be important.

[3] this acknowledgement provides a way to circumvent the oft-debated *coincidence problem*, the seeming paradox in which a single body of matter can be both "some gold" and "a ring" at the same time [12], which frequently arises in discussions of masses. The problem is solved by acknowledging that a single real phenomenon can be simultaneously categorized in two ways without having to be two phenomena.

[4] I have yet to find a formal definition of endurant or occurrent that isn't fraught with issues. Some have even argued that this demonstrates that the distinction doesn't exist [25]. I have developed a solution based on the $a/a : X$ distinction, but it is beyond the scope of this paper.

masses are *cumulative* or collective (the combination of two amounts of "some water" is still "some water") and generally *divisive* or dissective (half of "some water" is still "some water") [24, p.91,99] [10, p.38]. Objects are neither: half of "a car" is not "a car" but scrap metal, and "a car" and another car combined is not "a car" but "cars." Note that the concept of homogeneous reference is not evaluated over time, but at a moment: Over time, a rock (object) can be broken into multiple rocks, but then it is no longer a rock; at a single instant, a rock is not composed of multiple rock objects.

Spatial cumulativity and *spatial divisiveness*[5] can be formalized in mereology[6] as follows:

D2 $CumS(a:X) := \forall b[(T(a) \circ T(b) \ \wedge \ \forall t(t \leq T(a) \ \wedge \ t \leq T(b) \implies S(a_t) \neq S(b_t)) \ \wedge \ Xb \ \wedge \ \exists c(c = a + b)) \implies Xc]$

A categorized phenomenon is spatially cumulative iff for any other phenomenon with the same label, which exists at least in part during the same time, and is never spatially coincident, and such that the two phenomena have a meaningful mereological sum, then the sum is also of the same category. For example, if a is some sand categorized as "a volume of sand," and b is any different volume of sand that existed at the same time as a, such that it makes sense to consider their combination as a phenomenon, than that combination can also be categorized as a volume of sand. Conversely, a "country" fails this test because for any other distinct but contemporary country, it may be meaningful to collect them as a single phenomenon (i.e., the mereological sum exists), but that phenomenon is "two countries," a different universal.

D3 $DivS(a:X) := \exists b, c, t(t \leq T(a) \ \wedge \ t \leq T(b) \ \wedge \ t \leq T(c) \ \wedge \ \neg(S(b_t) \circ S(c_t)) \ \wedge \ a = b + c \ \wedge \ Xb \ \wedge \ Xc)$

A categorized phenomenon is spatially divisive iff at some time during its existence, it can be divided into two spatially disjoint parts that are each of the same category as the whole. For example, a typical volume of sand a can easily be divided into two volumes of sand b and c.[7] However, a country (as a sovereign state) cannot be composed of two countries.[8]

A mass can thus be defined as an endurant ($End(a:X)$) that is spatially amorphous, while an object is the opposite:

D4 $Mass(a:X) := End(a:X) \wedge CumS(a:X) \wedge DivS(a:X)$

A mass is any endurant that is spatially cumulative and divisive.

D5 $Object(a:X) := End(a:X) \wedge \neg(CumS(a:X) \wedge DivS(a:X))$

An object has one or neither of these characteristics. It is possible for a phenomenon to be cumulative but not divisive, such as two "horses." The opposite is common with

---

[5] Most existing definitions of masses do not distinguish homogeneous reference in space and time, but I have found this distinction to be crucial, because spatial parts and temporal parts have very different implications. Treatments of masses in space and time, such as Galton and Mizoguchi [7], would be more clear with this recognition. Spatial homogeneity distinguishes masses from objects, while temporal homogeneity (not discussed here) distinguishes occurrents into processes and events.

[6] In formal mereology, there sometimes seems to be as many notation systems and axiomatic systems as there are mereologists. I am using extensional mereology, **CEM** in the classification of mereological systems by Casati and Varzi [3], and the following notation: $\leq$ for part (proper part or equal), $\circ$ for overlap (having shared parts), $+$ for mereological sum. Note that mereology is employed only on the material-stratum phenomena, space and time, not on the categorized phenomena; this circumvents many of the issues with **CEM** pointed out by Simons and others [26].

[7] Yes, there is some very small volume of sand that can only be divided into two collections of a couple grains of sand, not a mass. More on this later.

[8] The United Kingdom is no exception; England is called a country, but it is still a different kind of entity from the UK as a whole.

■ **Figure 1** What is this? See Table 1.

linear and layer phenomena: a river can be cut into two parts, each called a river, but it is possible to find another river that combine to form "rivers."[9] In both these cases, it makes sense to classify them as objects.

How is this definition of objects and masses based on homogeneous reference equivalent to the earlier definition based on continuity and boundedness? The necessary boundary of an object clearly separates it from any neighboring object. Thus, when we consider them together, the intervening boundary makes us see them as two objects rather than one. Conversely, the boundary of a mass instance (say, a patch of water in the midst of the ocean) is at best arbitrary and inconsequential; so when two adjoining masses are considered together, their boundaries can be easily ignored (if they were ever recognizable to begin with) and the two considered as a single entity. Furthermore, the fact that a mass is divisive, able to be divided a number of ways without ontological change, suggests that the boundaries of each division are arbitrary and inconsequential.

The above definitions only apply to a single particular phenomenon categorized in one way. Each of the definitions could be extended to an entire universal category, iff every phenomenon that uses that category is classified the same way:

D4c $Mass(X) := \forall a(Xa \implies Mass(a : X))$
D5c $Object(X) := \forall a(Xa \implies Object(a : X))$

Likewise, they could be extended to a particular phenomenon in general, if every possible way of categorizing the phenomenon falls into the same ontological class:

D4p $Mass(a) := \forall X(Xa \implies Mass(a : X))$
D5p $Object(a) := \forall X(Xa \implies Object(a : X))$

The formal definitions can now be used to categorize actual phenomena. For example, the phenomena at the center of Figure 1 can be categorized in a number of ways, as shown in Table 1.

This example demonstrates the applicability of the formal definitions, but should not be taken as an inference of general patterns. For example, the last column is blank only

---

[9] This occurs because linear and layer objects are crucially bounded in one dimension, but not in the other(s).

Table 1 Categorizations of the phenomena shown in Figure 1.

X	CumS	DivS	$a : X$	X	a
A mountain	no	no	object	object	–
Limestone	yes	yes	mass	–	–
A geologic formation	no	yes	object	object	–
The mountains	yes	yes	mass	–	–
Mountainous landscape	yes	yes	mass	–	–
A mountain range	no	no	object	object	–
Terrain	yes	yes	mass	–	–
Elevation	yes	yes	mass	–	–

because I specifically chose a situation that could be classified many ways; there may be many phenomena that can only be classified one way. That said, it does show that categories and particulars may not be able to be universally classified as either mass or object. This demonstrates the real power of the formal distinction between a and $a : X$; previous attempts to define masses formally have generally either tried to define $\mathrm{Mass}(X)$ without reference to a (i.e., a nominalist approach), or to define $\mathrm{Mass}(a)$ without $a : X$ (i.e., a realist approach), or just conflate them, all of which have frequent exceptions, which have only served to strength opposition to the existence of masses.

4 Against Masses

Opinions are mixed on the existence of mass particulars [29]. Every few years since the 1960s, authors from various disciplines have proselytized the existence of masses, whereupon others have quickly responded to refute them. Each phase of the debate seems to repeat many of the same points, which I summarize here.

Arguments against the existence of mass particulars often involve two closely related assertions: 1) mass nouns (water, wood, metal) are strictly universals [30], because 2) any instance thereof (e.g., "the water in this lake") must have a boundary, and is therefore an object by definition. Perhaps the best refutation of these dates back to Chappell [4]. He accepts that instances of masses are different from their universals; he refers to the former as "parcels of stuff," using the most generic container term he could muster. However, he refutes the first argument by demonstrating that these parcels are still significantly different from object particulars (his "substances") in the same way that mass universals are different from object universals (i.e., having homogeneous reference), and should thus still be considered a separate kind of phenomenon.[10] The formal definition of a mass given above works just fine for these parceled particulars: if a is a parcel of water, and b is another parcel of water, then if $a \oplus b$ makes sense, it is a parcel of water.

---

[10] Laycock rejects Chappell's explanation on the grounds that requiring us to talk about plurals and masses in singular terms violates their inherent non-singularity [17]. In fact, Laycock doubts that mereology, set theory, or the entire predicate calculus can even apply to plurals and masses for this reason. However, he does not develop an alternative formalism, and those alternatives that have been published tend to have their own semantic and ontological problems; they may hold promise, though [19, 20]. Laycock's argument is compelling, but I believe Chappell's approach is still useful as long as the parcels are recognized as only temporary samples of the phenomenon, not the phenomenon of study (a solution mentioned in passing by Laycock). Simons follows a similar approach to fitting the predicate calculus and mereology to masses and plurals [26, pp.151–162].

On the second point, Chappell concedes that these parcels must have boundaries and a form, even if they are vague. However, he points out that they are "indifferent to form," that is, the boundaries and shape of a mass particular are not relevant to its identity and characteristics as a mass; as Jackendoff puts it, "one can think of the boundaries as outside the current field of view." [15, p.19] Furthermore, example mass particulars often used to argue against their specialness, such as the gold that constitutes a ring or the water contained in a cup, are clearly bounded and object-like, but they are straw man examples; it is just as easy to find masses that are practically impossible to bound and objectify, like the salt dissolved in the ocean or the moisture in soil.

It is not just that the boundary is vague; objects can also have indeterminate boundaries, but if so, their vaguely defined form is still crucial to their definition. For example, a mountain is usually vaguely bounded on the sides and bottom, but the form of its boundary (especially the profile shape of its upper surface) is absolutely crucial to its being a mountain. On the other hand, the rock that makes up the same mountain can be recognized, described and analyzed at length without ever referring to its boundary or shape.

Another issue that has been raised is that the common definitions of a mass, including the formal definition above, test positive for some phenomena that do not seem like prototypical masses. These include:

- Immaterial but not abstract phenomena (i.e., occurring at a location but having no mass), such as magnetism and field properties like temperature or population density.
- Phenomena that use mass terms, but are visibly discontinuous, such as vegetation or infrastructure.
- Uncounted plurals, such as "some people."

Each of these types of phenomena meet the formal definition of mass; do they meet the original intended definition, or is this a sign that the formalism is not faithful to the intent? All of them meet the requirement of being amorphous, because their boundaries are not relevant to their meaning.

Magnetism is continuous, and thus meets both of the requirements. As to its immateriality, note that none of the definitions require that a mass actually has mass: that is just an unfortunate coincidence of terminology, but every other term that has been proposed for this ontic category, such as substance or material, has the same problem. Immaterial continuous phenomena and field properties behave like masses, so I propose they should be considered a kind of mass.

The problem in the other two cases listed above is that they are not "really" continuous, but are composed of clearly visible individuals, unlike prototypical masses, such as water and metal; these are often called *collective nouns*. However, this distinction isn't as clear as it seems; it is just a matter of scale. Even most masses that appear continuous, such as water, are composed of objects at a sub-visible scale and are thus not infinitely divisive; this lower-limit mass decomposition is often called *Quine's minimal parts hypothesis* [24, p.99].

To talk about this scale effect, let use define the *support* of a phenomenon category as the smallest size that it can be and still be recognizable as that category; for a mass, it would be the approximate diameter of a collection of "several" constituent individuals that could be amassed.

It turns out that for almost any size support, one can come up with an example mass that is aggregated at that scale, as shown in Table 2. Where should we draw the line between a "true" mass and a collective? Yes, this argument rings of one of those classic Greek continuum paradoxes, but the point is that wherever we chose to draw the line would arbitrarily divide very similar phenomena. I have grouped them into four classes based on the relative perceptibility of the mass and its constituents, but even these have vague boundaries that depend on the particular phenomenon.

Table 2 Continuum of scales at which objects are aggregated into masses.

Mass	Constituent	Support	
magnetic force	truly continuous		MICRO-MASS
gold	atom	10^{-9}m	
water	molecule	10^{-8}m	only mass visible,
air	mix of molecules	10^{-6}m	object microscopic
clay	grain	10^{-5}m	
silt	grain	10^{-4}m	MINI-MASS
sand	grain	$10^{-2.5}$m	
grain	seed	10^{-2}m	object visible,
gravel	stone	$10^{-1.5}$m	mass common
grass	blade	10^{-1}m	
brick	brick	$10^{-0.5}$m	MESO-MASS
lumber	board	10^{0}m	object and mass
brush	plant	$10^{0.5}$m	equal
wildlife	animal	10^{1}m	MACRO-MASS
populace	person	$10^{1.5}$m	
woodland	tree	10^{2}m	
the desert	plant, rock, etc.	$10^{2.5}$m	object common,
the country	farm, house, road	10^{3}m	mass at distance
the mountains	mountain, valley	10^{4}m	or in abstract

Instead of trying to make the distinction at all, it seems more straightforward to just acknowledge that when we categorize these phenomena as masses, we are (temporarily) ignoring the individuals. As Bunt puts it, masses are treated "as if they did not consist of discrete parts," regardless of whether discrete parts physically exist or can be perceived [2, p.45]. This is much easier with the mini-masses than with the macro-masses, but it is the same cognitive leap. In fact, macro-masses have occasionally been acknowledged elsewhere. DOLCE, one of the general ontologies that have been published, has a category for "visual landscape," which includes phenomena such as The City, The Mountains, or The Desert, which are clearly macro-scale, and makes it a subcategory of "Amount of Matter" (its term for mass) [8].

Plurals are a little different in this regard; they acknowledge the existence of their constituent individuals, but they are still considered as less important than the collective mass.[11] Some are more mass-like than others, especially *pluralia tantum*, terms that occur only in plural form and cannot be counted [9, p.612], such as woods, outskirts, and suburbs. It occasionally works the other way too: some mass terms (in English) reflect universals that behave more like class aggregates than mass aggregates; they are a shortcut for talking about a variety of similar objects, but still recognizing them as distinct objects. The classic example is "furniture." It is entirely valid to refer to a single chair as "this furniture," which identifies it with a class, not a mass. A geographic-scale example would be a GIS layer called "infrastructure," which would likely consist of individual objects, not a single blob.

This intentional ignorance of boundaries and constituent individuals may seem offensive, especially if one is only concerned with things "as they really are," but it is not a problem

---

[11] The similarity, if not equivalence, of plurals and masses, is covered at length by Laycock and others [17]. Some, like Nicolas, have even attempted to combine them the other way by making masses look like plurals, but his model is a linguistic abstraction that makes little sense ontologically [19].

for our conceptual framework; it is just an example of a slightly weaker correspondence between a particular a and a choice of $a : X$. In fact, this framework minimizes the effect of Quine's minimal parts hypothesis by not expecting universal homogeneity. Considering a huge amount of water as a mass does not depend on what we would do with an unrelated microscopic collection of a few water molecules. Bunt recognizes this for linguistic analysis at least: "nothing in the use of a mass noun indicates a commitment ... to the existence of minimal parts." [2, p.45]. Essentially, the minimal parts hypothesis is an example of stratum conflation, imposing a material-realm expectation on a conceptual/social-realm entity.

One argument against assuming a strong correspondence is the fact that the mass/count term distinction varies from one language to another; some languages have more mass terms, some have less; very few have as many cases as English in which both are available to describe the same phenomenon (e.g., wildlife/animals). At the extreme, the Asian *classifier languages*, such as Indonesian and Cantonese, deal with almost all nouns in a very mass-like grammar, but this does not mean that they conceptualize everything as mass or that more masses exist in China than in England.

Another issue with assuming a direct correspondence between linguistic syntax and cognitive structure is that while languages evolve to express ideas, they are eventually standardized and regulated to a high degree. In English, some words are count and some are mass because that's what the OED says they are. In fact, I wonder if more could be learned about the linguistic/cognitive correspondence by studying bad grammar than by studying grammatical rules. For example, common mistakes by non-native speakers, and the perennial issue that students have with "data is/are," probably say more about their cognitive structures than their level of intelligence.

5   Do Masses Matter in Geography?

For whatever reason, masses have been occasionally mentioned in GIScience ontologies, but have never found their way into the common conceptual framework thereof. It is fair to ask, "Is that even an issue?" Perhaps they exist, but just do not matter enough to consider.

To investigate this, I took a look at what the field finds important, by surveying the subjects discussed in a spectrum of journals: the *International Journal of Geographic Information Science* (for a GIScience focus), the *Annals of the AAG* (for a broader geography perspective), *International Journal of Remote Sensing* (more of a remote sensing, physical geography, and raster focus) and *ArcUser* (for a GIS practitioner perspective).[12] In all, 91 articles were reviewed. The subject matter of each article was classified as either predominantly *physical geography*, *human geography*, or *environmental geography* (a mix of human and physical).

In each article, I recorded any endurant particular that was significantly discussed or studied. As best I could, I avoided occurents (especially processes) and universals (which typically meant skipping over the literature review and theoretical sections, and I generally skipped review articles), although one could probably debate the inclusion of a few of the items in my list. I documented 750 references to endurants (not unique; things like counties and cities were listed in many articles), an average of about 8 per article.

These were classified as either object or mass using the formal definitions above. Within the masses, I identified each as a field if it was clearly a property. Plural terms required

---

[12] Specifically, the issues mined were *IJGIS* V.22 #10, V.33 #1, V.33 #2 (21 total articles); *Annals* V.106 #1 January 2016, V.108 #3 March 2018 (28 total articles); *IJRS* V.40 #1 (20 articles); *ArcUser* V.19 #4 Fall 2016, V.21 #3 Summer 2018, #4 Fall 2018 (23 total articles). The sample was neither random nor strategic; these issues were at hand.

further consideration. They tended to fall into three conceptual groups, as evidenced by the narrative context: 1) the set of individuals was conceptualized as a single whole (e.g., "the group of people who were at the event"), in which case it was tagged as a single object; 2) the focus was on the individuals (e.g., "each of the 245 animals we saw..."), in which case they were tagged as objects; or 3) an uncounted plural that behaved as a mass concept (e.g., "farms lined the highway"), which was tagged as a type of mass. One test of the last type was whether the plural could be replaced by a synonymous mass term without changing the meaning (e.g., "farmland lined the highway").

This bibliometric study was not intended to be a rigorous analysis of the ontology of the entire discipline of geography or GIScience, only to get a feel for whether masses appear in geographic inquiry. They do.

Table 3 Phenomena listed in sampled geography, GIScience, and GIS articles.

	Objects	Masses	Field-Masses	Plural-Masses	Total
Human	50.9%	20.9%	12.4%	15.8%	387
Annals	79	42	15	29	165
IJGIS	38	11	18	9	76
IJRS	5	3	6	2	16
ArcUser	75	25	9	21	130
Environmental	48.2%	26.1%	12.2%	13.5%	245
Annals	60	33	3	23	119
IJGIS	10	5	5		20
IJRS	13	9	11	4	37
ArcUser	35	17	11	6	69
Physical	33.3%	29.9%	29.9%	6.8%	117
Annals	9	1	6		16
IJGIS	8	10	9		27
IJRS	22	19	25	8	74
Total	354	180	113	102	749
	47.3%	24.0%	15.1%	13.6%	

The listed phenomena are classified in Table 3. At least one mass particular was mentioned in almost every article, and overall, they were mentioned more often than fields. As one might expect, masses and fields were more common in physical geography papers than in human geography (with the latter tending to focus on human-built objects).

Although Mass phenomena were mentioned very frequently, they were never discussed more than in passing, often immediately being transformed into objects or fields for modeling. Very few articles reflected on the ontology of their subject matter at all, and of those that did, none acknowledged masses as an ontic category. For example, one paper discussed at length the ontology of terrain characteristics (slope, aspect, etc.) as fields, but never mentioned the ontology of terrain itself.

I also evaluated the mass and field-mass phenomena with regards to the earlier scale discussion, and encountered geography and GIS projects that were concerned with all of these scales. There were 73 mentions of micro-scale masses, 63 mini-scale, 18 meso-scale, and 139 macro-scale masses. One would expect micro-scale masses to be very common, simply because these are the prototypical masses that appear continuous to the naked eye. The large number of macro-scale masses is largely due to the fact that these are the scales at which geographic inquiry generally takes place, at which most individual objects are too small to be

considered separately. I did not fully classify the plural-masses, but in passing they appeared to be almost exclusively macro-scale. The relatively small number of meso-scale masses is likely due to a combination of its narrow range of scales, and the fact that these are the "table-top" scales at which prototypical objects are so prevalent. Also, many of the mass terms that do occur at these scales, such as silverware, are not very relevant to geography.

So, the fact is that masses of various aggregation scales are already represented in GIS, without overtly recognizing it. For example:

- A polygon layer representing trees (as opposed to digitizing every tree). Yes, the polygon shape suggests that an object is being represented, but this is an artifice of the vector data model. The phenomenon is a plural-mass, in which individual trees have been aggregated and then disregarded, and the polygon is merely a rough attempt to show where the trees-phenomenon is present. Common attributes of trees-as-mass, such as Percent Canopy Cover and Species Composition (themselves field-masses), would never make sense for individual trees.

- An isarithmic map of population density. Computing a continuous field of density using methods such as kernel density estimation is an overt act of aggregating and subverting individuals into a population mass, of which density is a property.

- A sensor network, such as a set of weather stations, which are essentially sample locations for measuring the characteristics of the air mass. Note that the network does not dictate a definition of the boundary of the air mass, and we can talk about the measured characteristics without any regard for such a boundary.

The "mass-ness" of some of these applications is greater than others. It may be that the mass concept is so embedded in a processing method that users don't need to know that it is there.

However, in other situations, a mass ontology could greatly benefit the design and implementation of data and procedures. For example, an impetus for this research was a paper submission I reviewed for a journal. The authors had obtained several layers of points representing different kinds of phenomenon, but they needed raster datasets for their model. Some of these were distinct objects and should have been rasterized by a mass-aggregation method such as kernel density estimation, while others were samples of a field, and should have been rasterized by an interpolation algorithm. However, it appeared that they chose their rasterization methods for each dataset randomly, and in almost every case, chose the wrong method. At best, they may have learned some sort of rubric in school to remember which is which, but forgot. At worst, they didn't care to learn which is which. Furthermore, the methods they used to combine and analyze the fields they generated were inappropriate for their ontological nature, resulting in a "solution" that had no real meaning. In this situation, a clear understanding of the phenomena their source data represented, and of the phenomena their model results represented, would have gone a long way toward making wiser design and analysis decisions, and a more coherent interpretation of the results.

This case is not alone in my experience. I'm not saying that all GIS analysis out there is being done incorrectly; the published work I collected above did not have glaring ontological errors, because adept GIScientists and GIS professionals have learned proper practice, even if they do not consciously think about masses. That said, just as a sound understanding of the nature of fields leads to better spatial analysis (especially using raster tools), an understanding of the role of masses in the geographic world and in geographic inquiry will lead to fewer mistakes of this type.

In a related way, perhaps one reason why masses are not widely recognized in geospatial theory and practice is because the intuitive pairing between the common dichotomies in

conceptual and data models (field with raster, object with vector) is so strong, to the point that exceptions to the correspondence (e.g., vector isolines representing a field) can cause confusion. I have seen such confusion lead to incorrect analyses, especially among students. Perhaps altering the ontology and/or conceptual model will break the assumed correspondence, helping people think more about what they are really trying to represent, and the data models and analysis tools they use.

Is there a need for new mass-based data models and analysis tools? Probably not. Because a mass does not have a strong identity like objects do, there is not much need for representing it as a unified entity. Typically, the study of geographic masses is the study of the geographic distribution of their attributes, which are nearly always fields. Since models, methods, and tools for representing and analyzing fields are mature, there is probably little need for more software.

That is, as long as we use what we have wisely. And being wise requires that you think about what you're doing. And part of what you're doing involves masses.

References

1 Oxford English Dictionary, 3rd Edition, 2000. URL: `http://oed.com/`.

2 Harry C. Bunt. *Mass Terms and Model-theoretic Semantics*, volume 42 of *Cambridge Studies in Linguistics*. Cambridge University Press, 1985.

3 Roberto Casati and Achille C Varzi. *Parts and places: The structures of spatial representation*. MIT Press, 1999.

4 Vere C Chappell. Stuff and things. In *Proceedings of the Aristotelian Society*, volume 71, pages 61–76. JSTOR, 1970–1971. URL: `https://www.jstor.org/stable/4544801`.

5 Helen Couclelis. A Typology of Geographic Entities with Ill-defined Boundaries. In Peter A. Burrough and Andrew U. Frank, editors, *Geographic Objects with Indeterminate Boundaries*, pages 45–55. London: Taylor and Francis, 1996. URL: `https://books.google.com/books?id=r-2i_hKbx9gC&pg=PA45`.

6 Anthony Galton. Desiderata for a Spatio-temporal Geo-ontology. In Werner Kuhn, Michael F. Worboys, and Sabine Timpf, editors, *Spatial Information Theory: Foundations of Geographic Information Science (COSIT 2003)*, pages 1–12. Berlin: Springer, 2003. `doi:10.1007/978-3-540-39923-0_1`.

7 Anthony Galton and Riichiro Mizoguchi. The Water Falls but the Waterfall does not Fall: New perspectives on Objects, Process and Events. *Applied Ontology*, 4(2):71–107, 2009. URL: `https://pdfs.semanticscholar.org/309d/2ba88406bff2f545eec8086d20673287070f.pdf`.

8 A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening Ontologies with DOLCE. In A. Gómez-Pére and V.R. Benjamins, editors, *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, 13th International Conference, EKAW 2002*, pages 166–181. Berlin: Springer Verlag, 2002.

9 Brendan S Gillon. Towards a common semantics for English count and mass nouns. *Linguistics and philosophy*, 15(6):597–639, 1992. `doi:10.1007/BF00628112`.

10 Nelson Goodman. *The Structure of Appearance, 3rd Edition*, volume 53 of *Boston Studies in the Philosophy of Science*. Dordrecht: D. Reidel, 1951/1977.

11 Jorge J.E. Gracia. *Metaphysics and Its Task: The Search for the Categorial Foundation of Knowledge*. State University of New York Press, 1999. URL: `https://books.google.com/books?id=_5JQKKsZIgUC`.

12 Richard E. Grandy. Stuff and things. *Synthese*, 21(3/4):479–485, 1975. `doi:10.1007/978-1-4020-4110-5_15`.

13 Heinrich Herre. Persistence, change, and the integration of objects and processes in GFO. In Vesselin Petrov and Adam C. Scarfe, editors, *Dynamic Being: Essays in Process-Relational Ontology*, pages 337–354. Cambridge Scholar Publishing, 2015.

14   Heinrich Herre.   Formal Ontology-A New Discipline Between Philosophy, Formal Logic, and Artificial Intelligence, 2016.   URL: `https://pdfs.semanticscholar.org/62af/ab4d0412df9519c1e7004ed3f5fc6fc59096.pdf`.

15   Ray Jackendoff. Parts and Boundaries. In B. Levin and S. Pinker, editors, *Lexical and Conceptual Semantics*, pages 9–45. Cambridge, MA: Blackwell, 1992.

16   George Lakoff. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* University of Chicago Press, 1987.

17   Henry Laycock. *Words without Objects: Semantics, Ontology, and Logic for Non-Singularity.* Oxford University Press, 2006.

18   Friederike Moltmann. Part structures, integrity, and the mass-count distinction. *Synthese*, 116:75–111, 1998. `doi:10.1023/A:1005046308299`.

19   David Nicolas. Mass nouns and plural logic. *Linguistics and philosophy*, 31(2):211–244, 2008. `doi:10.1007/s10988-008-9033-2`.

20   Alex Oliver and Timothy Smiley. *Plural Logic: Revised and Enlarged.* Oxford University Press, 2013.

21   Adam Pease. Suggested Upper Merged Ontology (SUMO). URL: `http://www.adampease.org/OP/`.

22   Donna Peuquet, Barry Smith, and B. Brogaard, editors. *The Ontology of Fields: Report of a Specialist Meeting held under the auspices of the Varenius Project.* Santa Barbara, CA: National Center for Geographic Information and Analysis, 1999. URL: `https://escholarship.org/uc/item/33d7c1pt`.

23   Roberto Poli. Ontology: The Categorial Stance. In Roberto Poli and Johanna Seibt, editors, *Theory and Applications of Ontology: Philosophical Perspectives*, pages 1–22. Springer, 2010. `doi:10.1007/978-90-481-8845-1_1`.

24   Willard V.O. Quine. *Word and object: An inquiry into the linguistic mechanisms of objective reference.* John Wiley, 1960.

25   Johanna Seibt.   Particulars.   In Roberto Poli and Johanna Seibt, editors, *Theory and Applications of Ontology: Philosophical Perspectives*, pages 23–55. Springer, 2010. `doi:10.1007/978-90-481-8845-1_2`.

26   Peter M. Simons. *Parts: a study in Ontology.* Oxford University Press, 1987.

27   Barry Smith and Pierre Grenon.   BFO: Basic Formal Ontology, 2002.   URL: `http://basic-formal-ontology.org/`.

28   Barry Smith and David Mark.   Geographical Categories: an Ontological Investigation. *International Journal of Geographical Information Science*, 15(7):591–612, 2001. `doi:10.1080/13658810110061199`.

29   Mark Steen. The Metaphysics of Mass Expressions. In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy.* Stanford University, 2012. URL: `http://plato.stanford.edu/entries/metaphysics-massexpress/`.

30   Peter F. Strawson. *Individuals.* London: Methuen, 1959.

Formal Qualitative Spatial Augmentation of the Simple Feature Access Model

Shirly Stephen

School of Computing and Information Science, University of Maine, Orono, ME 04469, USA
shirly.stephen@maine.edu

Torsten Hahmann

School of Computing and Information Science, University of Maine, Orono, ME 04469, USA
torsten.hahmann@maine.edu

─── **Abstract** ───

The need to share and integrate heterogeneous geospatial data has resulted in the development of geospatial data standards such as the OGC/ISO standard Simple Feature Access (SFA), that standardize operations and simple topological and mereotopological relations over various geometric features such as points, line segments, polylines, polygons, and polyhedral surfaces. While SFA's supplied relations enable qualitative querying over the geometric features, the relations' semantics are not formalized. This lack of formalization prevents further automated reasoning – apart from simple querying – with the geometric data, either in isolation or in conjunction with external purely qualitative information as one might extract from textual sources, such as social media. To enable joint qualitative reasoning over geometric and qualitative spatial information, this work formalizes the semantics of SFA's geometric features and mereotopological relations by defining or restricting them in terms of the spatial entity types and relations provided by CODIB, a first-order logical theory from an existing logical formalization of multidimensional qualitative space.

1 Introduction

The need to share and integrate the large amounts of heterogeneous geospatial data has resulted in the development of geospatial data standards, such as the OGC's GeoSPARQL standard [27] and the shared OGC/ISO standards Geography Markup Language (GML) [23] and Simple Feature Access [22]. All of these standards include some types of simple and complex geometric features – often simply referred to as *geometries* – for representing geographic objects. The most commonly used features include points, line segments and aggregations into polylines, and polygons and aggregations into polyhedral surfaces. Primarily concerned with interoperability across spatial databases and geographic information systems, these standards also prescribe a number of common spatial operators, e.g., for calculating intersections, differences, buffers, or distances between features.

Many of these standards have further incorporated a number of simple mereotopological relations (with Boolean values), such as intersects, contains, overlaps, meets, or crosses.

These are based on results from the Region Connection Calculus (RCC) [28] and the almost equivalent topological relations defined by the 9-intersection method [8, 9] and its dimensionally extended refinements (DE-9I) [5, 6] and further extensions [26, 29].

The Simple Feature Access (SFA) model [22], an OGC and ISO standard for vector-based encoding of 2D geometric data, is one of the most widely implemented standards for facilitating geospatial data interoperability. It is at least partially implemented by a wide range of geographic information systems and spatial database systems, including ArcSDE (the spatial database system that ArcGIS uses), PostGIS, and the spatial extensions of MySQL, Oracle, and IBM Db2. Other geospatial standards, like GeoJSON[1] and GeoSPARQL [27], also build on SFA.

However, the mereotopological relations provided by SFA and similar standards use them as query operators only[2]. This enables more natural access to geometric data but without formalizing the relationships between geometric representations and the mereotopological or other qualitative relations, these approaches cannot support qualitative reasoning over the queried information. Moreover, storing "native" topological information – for example as provided from textual sources where precise locations or spatial extents are unknown or unknowable – is currently not possible without having to *invent geometric objects*. For example, the spatial content of the two statements "Lot A is for sale and abuts Broadway." and "Lot B that does not border Broadway is not for sale." cannot be represented in GIS without assigning geometries to the named objects.

Frameworks for qualitative spatial representation and reasoning (see, e.g., the overview in [7]) such as the RCC support direct reasoning about topological and other kinds of qualitative spatial information (e.g., direction), but cannot easily mix geometric data sources (e.g., the precise location of "Broadway") and qualitative information (the fact that "Lot A" and Broadway are connected) to infer which lots on a property map may be for sale. Similar interpretation of qualitative spatial information on a geometric dataset is needed during natural disasters, when interpreting human reports (e.g., from social media or news reports) on road networks, elevation data, and hydrological data, to help answer simple queries, such as "is any part of the historic center flooded?".

The presented work is a step in this direction by developing a first-order logical theory[3] that treats geometric features (e.g., polylines, polygons) and relations between them as specializations of more general types of features (e.g., any kind of 2D regions or 1D features) and mereotopological relations between them. Key to this endeavour is the use of a *multidimensional* theory of space wherein, unlike traditional logical theories of mereotopology (including the RCC), spatial entities of different dimensions can co-exist and be related. We choose the theory CODIB (based on CODI [17, 16] with an extension by boundary/interior distinctions [15]) as suitable multidimensional theory of qualitative space and test to what extent geometric features from SFA [22] can be treated as specializations of CODIB's more general non-geometric spatial feature types from CODIB. For example, SFA's line segments or polylines should specialize the general one-dimensional spatial features, called "curves", from CODIB. Specifically, we want to leverage the detailed formal semantics encoded in CODIB to capture the semantics of SFA's various geometric feature types and mereotopological relations in greater detail. Currently, much of these semantics are described in natural language and mathematical notation in the standard, but are not accessible to automated

---

[1] http://geojson.org/

[2] Most GIS support the RCC or DE-9I relations, with recent progress on storing the computed relations more efficiently [24]. There has also been a call to extend this to a larger set of qualitative relations [11].

[3] The term "theory" refers throughout the paper to a logical theory. The terms "theory", "ontology" and "axiomatization" are used synonymously.

reasoning. Wherever possible, we logically define SFA's geometric features in terms of CODIB's spatial concepts and, where that is not possible, treat them as specializations with suitable constraints.

Our specific contributions are: (1) develop a first-order logic axiomatization, called SF-FOL, of SFA; (2) in the process, show that all of the geometric feature types from SFA specialize or map to types of spatial entities definable in CODIB; (3) fully define SFA's mereotopological relations in CODIB and thus provide computer-interpretable semantics of these qualitative relations; and (4) verify the consistency of SF-FOL. This makes both SFA's and CODIB's mereotopological relations applicable to geometric and qualitative data alike and allows using automated first-order logic theorem provers (ATPs) for integrated mereotopological reasoning over combinations of qualitative and geometric data from any sources that adhere to the SFA standard.

2 Background and Related Work

Mereotopological relations are among the most common qualitative spatial relations [25], and have been incorporated into virtually all upper ontologies [14]. They include purely topological relations such as contact/connection or disconnection, and purely mereological relations such as parthood, containment, or inside, as well as relations that describe the interaction of topology and mereology such as overlap (i.e., contact via sharing a part). Simple mereotopological relations have also been included in popular geospatial data standards thanks to seminal work on the 9-intersection method [8, 9], its dimension-extended refinement (DE-9I) [5] and extensions thereof [6, 26, 29]. However, the 9-intersection method determines these relations from geometric data by computing a matrix of values that indicate the pairwise intersections of two object's *interior* ($\circ$), *boundary* (∂), and *complement* ($'$). Each of the nine pairs have either Boolean values – empty nor non-empty intersection – as in the original 9-intersection framework, or have dimensional values – either -1 (empty intersection), 0, 1, or 2 – as in the dimension-extended method. This way of determining the qualitative relations requires an underlying geometric representation, with associated operations for determining their boundary and interior, for all involved objects. Moreover, the semantics of the mereotopological relations, especially their interaction (e.g. parthood specializes overlap or a whole is in contact with everything any of its parts is on contact with), are never explicitly captured and thus not available for qualitative reasoning with the underlying data. Moreover, the relations cannot be used for reasoning where geometric data models are not the only source of qualitative information.

This is in sharp contrast with axiomatic treatments of mereotopology, which constrain the interpretations of one or two primitive relations, such as contact and/or parthood, and define other relations, such as overlap or external contact, in terms of the primitive ones [3]. By explicitly formalizing relationships between the relations, axiomatic frameworks permit reasoning with qualitative information even in the absence of geometric information. The most well-known axiomatic theory is the RCC [28] that defines eight mereotopological relations similar to the ones from the basic 9-intersection model. The variety of existing axiomatic theories are more thoroughly reviewed in [20][4]. However, axiomatic theories of mereotopology

---

[4] Qualitative spatial calculi (see, e.g., the overview in [7]) are yet another approach to qualitative spatial reasoning, but they can only incorporate qualitative information and cannot make use of geometric information without first translating it to qualitative information. A hybrid reasoning system utilizing a constraint network reasoning approach for reasoning with both geometric and qualitative information has been presented in [10]. This work here goes a step further by explicitly formalizing the semantic

have, in the philosophical tradition of Whitehead, been often married to strict region-based conceptualizations of space wherein extended spatial entities – typically called regions – are the only first-class entities of the domain, while points and other lower-dimensional entities are not entities in the domain. This prevents full integration with geometric data standards, such as SFA, that permit entities of different dimensions. The idea of *multidimensional mereotopology* [12, 13, 17, 30] aims to overcome this restriction by axiomatically formalizing mereotopological relations not just between entities of equal dimensions but also between entities of different dimensions. This work utilizes the multidimensional mereotopology CODIB [17, 16, 15], which has been specifically developed to qualitatively generalize geometric data models, as basis for formalizing SFA's semantics. CODIB is based on the three primitive relations of **CO**ntainment, relative **DI**mension, and **B**oundary containment [15], which give the theory its name. CODIB builds on and extends the theory CODI (without any notion of boundaries) [17, 16] by the additional relation of boundary containment. Unlike other multidimensional theories [12, 30], CODIB allows entities of lower dimensions to exist independent of entities of higher dimension, similar to how such entities (e.g., polylines or points) are used in geometric data standards. [12, 30] require each line or curve to be part of the boundary of some 2D region and each point to be the endpoint of some curve in a model. The INCH calculus [13], on the other hand, does not model boundaries at all. Another alternative formalization of multidimensional mereotopology is provided by the space ontology (GFO space) [1] that is part of the General Formal Ontology (GFO). However, GFO space is primarily concerned with *physical, phenomenal space* (i.e., the space of material objects), which is different from the kind of *abstract, extensional space* that geometric data models describe[5][15, 1].

3 Preliminaries

We now review and formalize the relevant aspects of the SFA standard, namely its classes of geometric features and its qualitative relations. In particular, Section 3.1 formalizes the intrinsic semantics of the UML subclass hierarchy from the standards document in first-order logic as starting point for its semantic enhancement. Subsequently, Section 3.2 reviews key relations and concepts from the CODI and CODIB ontologies and provides definitions of novel concepts that are necessary to draw some of the distinctions that SFA makes. These concepts and relations will be used as basis for elaborating the SFA semantics and making its geometric features available for integration with purely qualitative information and for general qualitative reasoning.

All logical sentences throughout our exposition are assumed to be universally quantified. They are labeled in the format '[theory]-[type][number]' (e.g. SFC-T1) where the first letter(s) indicate the theory (e.g. SFC=simple features concept, SFR=simple features relation, PO=partial overlap, D=dimension), while the type distinguishes axioms (A), definitions (D: defining a concept or relation), theorems (T: a property provable from the axioms and definitions), and mappings (M: an axiom that establishes some relationship between SFA and CODIB). All theories are available in modularized form in the Common Logic syntax from the COLORE repository[6].

---

relationships between the two types of information for reuse with any logic-based reasoner.

[5] For example, in phenomenal space, any road would be a 3D object, whereas in abstract space it is typically modeled as a 1D spatial feature.

[6] In `https://colore.oor.net/`. Note that all of axioms are specified using only the classical first-order logic syntax of Common Logic and without use of any of Common Logic's specialized features such as restricted module import or use of sequence markers. This allows easy translation to pure first-order

3.1 Semantics of Simple Feature Concepts and Spatial Relations

SFA [21] is an OGC and ISO standard for vector-based encoding of 0-2D geometric data that aims to facilitate interoperability across GIS and spatial databases. SFA is at least partially implemented by ArcGIS, PostGIS, and the spatial extensions of MySQL, Oracle, and IBM Db2. Other standards, like GeoSPARQL [27] and GeoJSON, build on it.

3.1.1 Semantics of Concepts (Classes) from Simple Features

At the core of SFA lies a set of simple geometries such as individual points (*sf_point*), polylines (*sf_line_string*: a sequence of straight line segments), and polyhedral surfaces (*sf_polyhedral_surface*: a connected, possibly non-planar 2D area obtained by stitching polygons together). *Sf_line_string* and *sf_polyhedral_surface* specialize the general classes *sf_curve*, which may include non-straight segments, and *sf_surface*, which may include 2D areas with non-straight boundary segments, respectively (SFC-A1,A2). These two classes capture all kinds of 1D and 2D spatial objects. Note that at this point, we only formalize the relationships between the classes as we cannot capture their detailed semantics. Only later on, with the help of CODIB concepts and relations, can we formalize the classes in more detail.

In addition to the three classes of simple features, collections of simple features can be modeled using the *sf_geometry_collection* class. The four specializations of the abstract class *sf_geometry* are mutually disjoint (SFC-A3–A6) and jointly exhaustive (SFC-D1).

(SFC-D1) $sf\_geometry(x) \leftrightarrow$
  $sf\_point(x) \lor sf\_curve(x) \lor sf\_surface(x) \lor sf\_geometry\_collection(x)$

(SFC-A1) $sf\_line\_string(x) \rightarrow sf\_curve(x)$

(SFC-A2) $sf\_polyhedral\_surface(x) \rightarrow sf\_surface(x)$

(SFC-A3) $sf\_point(x) \rightarrow \neg sf\_curve(x) \land \neg sf\_surface(x) \land \neg sf\_geometry\_collection(x)$

(SFC-A4) $sf\_curve(x) \rightarrow \neg sf\_point(x) \land \neg sf\_surface(x) \land \neg sf\_geometry\_collection(x)$

(SFC-A5) $sf\_surface(x) \rightarrow \neg sf\_point(x) \land \neg sf\_curve(x) \land \neg sf\_geometry\_collection(x)$

(SFC-A6) $sf\_geometry\_collection(x) \rightarrow \neg sf\_point(x) \land \neg sf\_curve(x) \land \neg sf\_surface(x)$

Sf_line_string is further specialized into *sf_line* (SFC-A7), which represents a single straight line segment, and *sf_linear_ring* (SFC-A8), a linear feature that is closed, that is, its start and end points coincide and thus its boundary is empty. The intended semantics of *sf_line* and *sf_linear_ring* will be more fully formalized in Section 4.1 by establishing mappings to CODIB concepts that are more densely axiomatized. For example, SFC-M3, M4, M8, and M9 together with CODIB's formalization (including the definitions of AtomicS-D, SimpleS-D, BranchedS-D, ConS-D, and the formalization of the predicate *ICon* from [15]) entail that any *sf_line* is a connected curve with two distinct end points. Likewise, *sf_polygon* is a specialization of *sf_polyhedral_surface* (SFC-A9), capturing a planar 2D area with a single closed polyline as exterior boundary[7]. Another specialization of *sf_polyhedral_surface* is *sf_tin* (SFC-A10), a triangulated irregular network (TIN), which consists of triangles. A single triangle, described by *sf_triangle*, is a polygon and the simplest kind of a TIN (SFC-D2). It is bounded by a closed polyline (i.e., a *sf_linear_ring*) that consists of exactly three line segments (i.e., *sf_line*) – this will be formalized by SFC-M13 in Section 4.1.

---

logic representations such as the TPTP format [31] supported by many theorem provers and model finders.

[7] SFA models *sf_polygon* and *sf_polyhedral_surface* as separate specializations of *sf_surface*, but permits polyhedral surfaces to consist of a single polygon, in which case it is spatially a polygon.

(SFC-A7) $sf\_line(x) \to sf\_line\_string(x)$

(SFC-A8) $sf\_linear\_ring(x) \to sf\_line\_string(x)$

(SFC-A9) $sf\_polygon(x) \to sf\_polyhedral\_surface(x)$

(SFC-A10) $sf\_tin(x) \to sf\_polyhedral\_surface(x)$

(SFC-D2) $sf\_triangle(x) \leftrightarrow sf\_polygon(x) \wedge sf\_tin(x)$

Sf_multi_point, *sf_multi_curve* and *sf_multi_surface* specialize *sf_geometry_ collection* (SFC-A11); they are aggregations of only *sf_point*s, *sf_curve*s, or *sf_surface*s, respectively. *Sf_multi_curve* and *sf_multi_surface* are again abstract classes in SFA, with only the specializations *sf_multi_line_string* (SFC-A12) and *sf_multi_polygon* (SFC-A13) being instantiable. The latter two consist only of *sf_line_string*s and *sf_polygon*s, respectively, as axiomatically captured in Section 4.2.

(SFC-A11) $sf\_geometry\_collection(x) \leftrightarrow$
$\quad sf\_multi\_point(x) \vee sf\_multi\_curve(x) \vee sf\_multi\_surface(x)$

(SFC-A12) $sf\_multi\_line\_string(x) \to sf\_multi\_curve(x)$

(SFC-A13) $sf\_multi\_polygon(x) \to sf\_multi\_surface(x)$

The axioms SFC-A1 to SFC-A13 together with SFC-D1,D2 form the ontology SFC-Core[8] that serves as basis for our semantic elaboration of SFA in Section 4.

3.1.2 Spatial Relations in Simple Features

In addition to various geometric spatial operations (e.g., buffer, intersection, convexHull), which are only well-defined on geometric features (e.g., on polygons rather than general surfaces), SFA includes eight named qualitative spatial relations based on the dimensionally extended 9-intersection method [5] that can equally be applied to generalizations of geometric features such as arbitrary curves and surfaces. SFA's relations include the five primitive relations *disjoint*, *touches*, *within*, *overlaps*, and *crosses*, with three additional relations *contains* (inverse of *within*), *intersects* (negation of *disjoint*), and *equals* (conjunction of *within* and *contains*) being defined[9]. SFA expresses the semantics of these relations using the interior, boundary, and exterior of the related objects [22], but does not formally relate the relations to one another as we will do in Section 4.3. Three dimensional constraints are explicitly mentioned in SFA: touches does not apply to points (or *sf_multi_point*s), overlaps requires the involved entities to be of equal dimension, and crosses is not applicable to two surfaces (or *sf_multi_surface*s). These constraints will become provable as theorems of our CODIB-based formalization of these relations.

3.2 Dimensional Features and Qualitative Spatial Relations in CODIB

This subsection reviews CODIB by first reviewing its core CODI and then additional relation of boundary containment. A computer-readable encoding of the axioms are provided in the Common Logic syntax in the COLORE repository to facilitate automated verification and reasoning.

---

[8] Available from `https://colore.oor.net/simple_features`.
[9] See the definitions provided in SFR-M6–M8. We have only decided to map *contains* to CODIB's *Cont* relation and then define *within* as its inverse.

3.2.1 CODI

Core to the multidimensional mereotopology CODIB is the theory CODI[10] of containment and dimension that axiomatizes mereotopological relations in a dimension-independent way using two primitive relations: (1) the mereological notion of containment, $Cont(x, y)$[11], and a relation $\leq_{\dim} (x, y)$, read as "x has the same or a lower dimension than y", to compare the dimension of two entities [16, 17]. In addition, the primitive unary predicate $S(x)$ is used to denote spatial regions, which capture mathematical regions of space whose existence is independent of whether an actual physical object occupies a spatial region or not. $Cont$ is reflexive, symmetric, and transitive (Cont-A1–A3) and allows defining the zero (i.e., null) region denoted by the unary predicate ZEX (ZEX-D). Containment requires the contained entity to be of the same or a lower dimension than the entity it is contained in (CD-A1).

The relative dimension $\leq_{\dim} (x, y)$ alone can define additional relations of equal dimension $=_{\dim} (x, y)$, lesser dimension $<_{\dim} (x, y)$, minimal dimension $MinDim(x)$ (i.e., the dimension of a point; D-D6), and next-lower dimension $\prec_{\dim} (x, y)$ (D-D7). The relation $\leq_{\dim} (x, y)$ is axiomatized to form a discrete (i.e., there is a next-lower dimension for every non-minimal entity) and bounded (i.e., a lowest and highest dimension exists) pre-order over all spatial regions. That also implies that every spatial region must be of uniform dimension, i.e., all components (i.e. parts) thereof are of the same dimension, precluding objects such as a region consisting of a 2D region and a separate, isolated point or linear feature. Spatial regions can still *contain* lower-dimensional entities (e.g., a 2D region containing 1D features and points). Using the relative dimension of the involved entities, we can specialize containment to parthood (i.e., equidimensional containment; EP-D) and proper parthood (EPP-D). Minimal spatial entities have no proper parts (ME-D2), that is, they are indivisible. There can be minimal entities within each dimension.

(Cont-A1) $S(x) \wedge \neg ZEX(x) \leftrightarrow Cont(x, x)$

$\qquad\qquad\qquad\qquad\qquad\qquad$ (***containment is reflexive for all nonzero spatial regions***)

(Cont-A2) $Cont(x, y) \wedge Cont(y, x) \rightarrow x = y$ $\qquad\qquad$ (***containment is antisymmetric***)

(Cont-A3) $Cont(x, y) \wedge Cont(y, z) \rightarrow Cont(x, z)$ $\qquad\qquad$ (***containment is transitive***)

(ZEX-D) $ZEX(x) \leftrightarrow S(x) \wedge \forall y[\neg Cont(x, y) \wedge \neg Cont(y, x)]$ $\qquad\qquad$ (***zero region***)

(CD-A1) $Cont(x, y) \rightarrow x \leq_{\dim} y$ $\qquad\qquad$ (***interaction between $Cont$ and $\leq_{\dim}$***)

(D-D6) $MinDim(x) \leftrightarrow \neg ZEX(x) \wedge \forall y\,[\neg ZEX(y) \rightarrow x \leq_{\dim} y]$ (***minimal-dimensional entities***)

(D-D7) $x \prec_{\dim} y \leftrightarrow (\leq_{\dim} y \wedge \neg(y \leq_{\dim} x) \wedge \forall z\,[z \leq_{\dim} x \vee y \leq_{\dim} z]$ (***next-lower dimension***)

(EP-D) $P(x, y) \leftrightarrow Cont(x, y) \wedge x =_{dim} y$ $\qquad\qquad$ (***parthood: equidimensional containment***)

(EPP-D) $PP(x, y) \leftrightarrow P(x, y) \wedge x \neq y$ $\qquad\qquad$ (***proper parthood***)

(ME-D2) $Min(x) \leftrightarrow \neg ZEX(x) \wedge \forall y\,[\neg PP(y, x)]$ $\qquad\qquad$ (***minimal entities within a dimension***)

Contact, $C(x, y)$, as the most general topological relation is definable as x and y sharing some contained object (C-D) and is provably reflexive and symmetric. Specialized types of contact can be distinguished based on the relative dimension: partial overlap $PO(x, y)$ holds only between entity of equal dimension and requires them to share a part (PO-D); incidence $Inc(x, y)$ holds between entities of different dimension and requires a part of the lower-dimensional entity to be shared with the higher-dimensional entity (Inc-D); and superficial contact $SC(x, y)$ requires the shared entity to be of a lower dimension than both of the entities in contact (SC-D).

---

[10] `colore.oor.net/multidim_mereotopology_codi/codi.clif`
[11] The relation $Cont$ is the qualitative generalization of SFA's *contains* relation.

(C-D) $C(x,y) \leftrightarrow \exists z[Cont(z,x) \wedge Cont(z,y)]$ $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ (***contact***)

(PO-D) $PO(x,y) \leftrightarrow \exists z[P(z,x) \wedge P(z,y)]$ $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ (***overlap in a part***)

(Inc-D) $Inc(x,y) \leftrightarrow \exists z[(Cont(z,x) \wedge P(z,y) \wedge z <_{dim} x) \vee (P(z,x) \wedge Cont(z,y) \wedge z \prec_{\dim} y)]$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ (***incidence***)

(SC-D) $SC(x,y) \leftrightarrow \exists z[Cont(z,x) \wedge Cont(z,y)] \wedge \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow z \prec_{\dim} x \wedge z \prec_{\dim} y]$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ (***superficial contact***)

While CODI does not distinguish different primitive types of entities, they can be defined: *PointRegion*s (which encompass individual points and sets of points) are of minimal dimension, *Curve*s are of next higher dimension, and so forth [19]. All of these primitive classes specialize the class S of abstract spatial regions.

(PR-D) $PointRegion(x) \leftrightarrow S(x) \wedge MinDim(x) \wedge \neg ZEX(x)$ $\quad\quad\quad\quad\quad\quad\quad$ (***point sets***)

(Point-D) $Point(x) \leftrightarrow PointRegion(x) \wedge Min(x)$ $\quad\quad\quad\quad\quad\quad\quad\quad$ (***individual points***)

(Curve-D) $Curve(x) \leftrightarrow S(x) \wedge \forall y[PointRegion(y) \rightarrow y \prec_{\dim} x]$ $\quad\quad\quad$ (***curves as 1D entities***)

(AR-D) $ArealRegion(x) \leftrightarrow S(x) \wedge \forall y[Curve(y) \rightarrow y \prec_{\dim} x]$ $\quad$ (***areal regions as 2D entities***)

3.2.2   CODIB

CODIB[12] is a logical extension of CODI that introduces an additional primitive relation of *boundary containment*, $BCont(x,y)$. *BCont* specializes containment by requiring the contained entity to be of a lower dimension than the containing entity (BC-A1), though the contained entity does not need to be of the next-lower dimension. For example, an areal (i.e., 2D) region can contain both curves and points in its boundary. Additional axioms (BC-A2–A5) that constrain the interaction of *BCont* with other relations, including incidence, parthood, partial overlap and containment are not shown here, they are documented in [15]. *BCont* is primitive because it cannot be defined in CODI, that is, in some models of CODI it cannot be determined whether a contained entity is actually contained in the boundary or interior of some containee.

(BC-A1) $BCont(x,y) \rightarrow Cont(x,y) \wedge x \prec_{\dim} y$

3.2.3   Refined Spatial Region Concepts in CODIB

In order to express the SFA concepts in detail, we further refine the basic dimensionally defined classes from CODIB based on whether and how their parts are connected, resulting in the subclass hierarchy shown in Figure 1. A connected region is one that is internally connected (ConS-D), while a region that is not internally connected is called a multipart region (MS-D). The property of *Internal connectedness* (ICon-D) from CODI requires each proper part y of an entity x to be connected to its complement $x - y$ such that the shared entity (denoted by the intersection of y and $x - y$) is of exactly one dimension lower than x[13]. For example, two polygons that share a line segment as boundary are internally connected, but if they only share a point, they are not.

A connected region that contains at least three non-overlapping proper parts that share an entity of lower dimension is called a branched region (BranchedS-D). A simple region is one that is connected and not branched (Simple-D). An atomic region is a simple region without any proper parts (Atomic-D).

---

[12] `colore.oor.net/multidim_mereotopology_codib/codib.clif`
[13] See [16] for the full axiomatization of the intersection and complement operations in CODI.
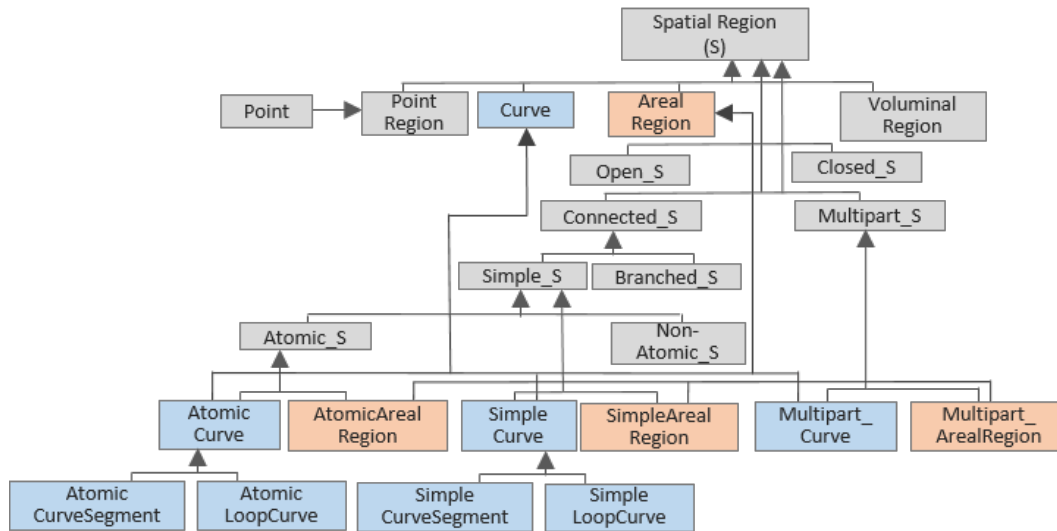
■ Figure 1 Taxonomy of refined CODIB spatial region concepts classified based on presence/absence of boundaries, connectedness, branching and parts.

(ICon-D) $ICon(x) \leftrightarrow \forall y[PP(y,x) \rightarrow C(y,x-y) \wedge y \cdot (x-y) \prec_{\dim} x]$ (*internally connected*)

(ConS-D) $Connected\_S(x) \leftrightarrow S(x) \wedge ICon(x)$ (*connected spatial region*)

(MS-D) $Multipart\_S(x) \leftrightarrow S(x) \wedge \neg Connected\_S(x)$ (*multipart spatial region*)

(BranchedS-D) $Branched\_S(x) \leftrightarrow Connected\_S(x) \wedge \exists p,q,r,s[PP(p,x) \wedge PP(q,x) \wedge PP(r,x) \wedge$ $\neg PO(p,q) \wedge \neg PO(p,r) \wedge \neg PO(q,r) \wedge s \prec_{\dim} p \wedge s \prec_{\dim} q \wedge s \prec_{\dim} r \wedge Cont(s,p) \wedge Cont(s,q) \wedge$ $Cont(s,r)]$ (*A branched spatial region is a connected region that has three distinct non-overlapping parts p,q,r that all share a common lower-dimensional entity s. For example, a branched curve has three non-overlapping segments that all share a point.*)

(SimpleS-D) $Simple\_S(x) \leftrightarrow Connected\_S(x) \wedge \neg Branched\_S(x)$ (*simple spatial region*)

(AtomicS-D) $Atomic\_S(x) \leftrightarrow Simple\_S(x) \wedge Min(x)$ (*an atomic spatial region is a simple spatial region that is minimal, i.e., has no proper parts*)

These properties are now used to define specialized classes of curves and areal regions.

(SCS-D) $SimpleCurveSegment(x) \leftrightarrow Curve(x) \wedge Simple\_S(x) \wedge \exists p,q[BCont(p,x) \wedge$ $BCont(q,x) \wedge p \neq q]$ (*Simple curve segment has two distinct end points*)

(SLC-D) $SimpleLoopCurve(x) \leftrightarrow Curve(x) \wedge Simple\_S(x) \wedge \forall y[Point(y) \rightarrow \neg BCont(y,x)]$

      (*Simple loop curve is closed: it does not contain any point in its boundary*)

(ACS-D) $AtomicCurveSegment(x) \leftrightarrow SimpleCurveSegment(x) \wedge Atomic\_S(x)$

(ALC-D) $AtomicLoopCurve(x) \leftrightarrow SimpleLoopCurve(x) \wedge Atomic\_S(x)$

(SAR-D) $SimpleArealRegion(x) \leftrightarrow ArealRegion(x) \wedge Simple\_S(x)$

(MC-D) $Multipart\_Curve(x) \leftrightarrow Curve(x) \wedge Multipart\_S(x)$

(MAR-D) $Multipart\_ArealRegion(x) \leftrightarrow ArealRegion(x) \wedge Multipart\_S(x)$

4   Axiomatization of Simple Feature as Extension of CODIB

In this section we present the core of our formalization that elaborates the semantics of the concepts in the skeleton axiomatization of SFA from Section 3.1 using qualitative concepts and relations from $CODI(B)$. This results in two new ontologies that logically extend SFC-Core and CODIB: SFC-FOL, which includes the more detailed axiomatization of SFA's concepts, and SFR-FOL, which axiomatizes SFA's mereotopological relations. Figure 2 summarizes the taxonomic relationships between the SFA and $CODI(B)$ concepts, but the real contribution are the detailed axiomatic mappings.

4.1   Axiomatization of Simple Feature's Simple Geometric Features

SFA's most general spatial entity is the class *sf_geometry*, which can be mapped to CODI's (and CODIB's) most general class of a *spatial region S* (SFC-M1). *Sf_point* and *sf_surface* map one-to-one to CODI's *Point* and *ArealRegion* (SFC-M2,M6), respectively. CODI's *Curve* captures any kind of one-dimensional features, that may be bounded segments (e.g., a *CurveSegment*), closed (e.g., a *LoopCurve*), infinite (e.g., a ray or a line in the mathematical sense), or branching with more than three endpoints. *Sf_curve* is much more restricted in scope in that it explicitly requires a start and an end point, though the points may coincide as in a closed curve. SFA's definition of *sf_curve* rules out infinite or branching curves. Thus, *sf_curve* maps to the union of *SimpleCurveSegment* and *SimpleLoopCurve* (SFC-M3). SFC-M4 and SFC-M5 elaborate the two cases in more detail. A *sf_curve* that is a *SimpleCurveSegment* has distinct start and end points (SFC-M4), while one that is a *SimpleLoopCurve* has identical[14] start and end points (SFC-M5) and does not contain any points in its boundary (SFC-T1). The axioms SFC-M1 to M6 tie in SFA's simple features with the qualitative spatial ontologies CODIB and allows using CODIB's mereotopological relations in conjunction with SFA features.

(SFC-M1) $sf\_geometry(x) \leftrightarrow S(x)$

           (**sf_geometry is equivalent to CODIB's Spatial Region class**)

(SFC-M2) $sf\_point(x) \leftrightarrow Point(x)$ (**sf_point is equivalent to CODIB's Point**)

(SFC-M3) $sf\_curve(x) \leftrightarrow SimpleCurveSegment(x) \lor SimpleLoopCurve(x)$

       (**an sf_curve is either a SimpleCurveSegment or SimpleLoopCurve in CODIB**)

(SFC-M4) $sf\_curve(x) \land SimpleCurveSegment(x) \rightarrow \exists p1, p2[sf\_point(p1) \land sf\_point(p2) \land$
    $sf\_start\_point(p1, x) \land sf\_end\_point(p2, x) \land BCont(p1, x) \land BCont(p2, x) \land p1 \neq p2]$

   (**an sf_curve that is a simple curve segment has distinct start and end points that are boundary contained**)

(SFC-M5) $sf\_curve(x) \land SimpleLoopCurve(x) \rightarrow \big[\exists p1, p2[sf\_point(p1) \land sf\_point(p2) \land$
    $sf\_start\_point(y, x) \land sf\_end\_point(z, x)]\big] \rightarrow y = z$

        (**an sf_curve that is a simple loop curve has the same start and end point**)

(SFC-T1) $sf\_curve(x) \land SimpleLoopCurve(x) \rightarrow \neg\exists y[sf\_point(y) \land BCont(y, x)]$

      (**an sf_curve that is a loop curve does not contain any point in its boundary**)

(SFC-T2) $sf\_curve(x) \rightarrow \forall y[PP(y, x) \land Min(y) \rightarrow AtomicCurveSegment(y)]$

         (**any sf_curve has AtomicCurveSegments as only minimal parts**)

(SFC-M6) $sf\_surface(x) \leftrightarrow ArealRegion(x)$

             (**sf_surface is equivalent to CODIB's ArealRegion**)

---

[14] Note that in CODIB, two points are identical if they are co-located.

The SFA concepts at the next, more refined level of the hierarchy in Figure 2 use CODIB's distinctions between (1) open and closed, (2) atomic, simple (atomic or not), and branched. For example, the SFA concept *sf_line_string* refines the union of CODIB's *SimpleCurveSegment* and *SimpleLoopCurve* and *sf_line* refine *AtomicCurveSegment*, respectively (SFC-T3,M7). The only added constraints are that each segment is a linear approximation between two points – a fact that cannot be expressed within a qualitative representation of space. *Sf_linear_ring* is a *sf_line_string* that is closed and thus a *SimpleLoopCurve* (SFC-M8).

(SFC-T3) $sf\_line\_string(x) \rightarrow SimpleCurveSegment(x) \vee SimpleLoopCurve(x)$

$$(\textbf{\textit{from SFC-A1, SFC-M3}})$$

(SFC-M7) $sf\_line(x) \rightarrow AtomicCurveSegment(x)$

$$(\textbf{\textit{sf\_line specializes CODIB's AtomicCurveSegment}})$$

(SFC-M8) $sf\_linear\_ring(x) \rightarrow SimpleLoopCurve(x)$

$$(\textbf{\textit{sf\_linear\_ring specializes CODIB's SimpleLoopCurve}})$$

*Sf_polygon*s are simple areal regions with a single exterior boundary and with each boundary piece being a *sf_linear_ring* (SFC-M9). A *sf_polyhedral_surface* is a simple areal region formed by "stitching" together *sf_polygon*s along their common boundaries (SFC-M10). Such surfaces in a 3-dimensional space may not be planar as a whole. An *sf_triangle* is a *sf_polygon* (SFC-M11) with exactly three non-overlapping lines forming their boundary. The exterior boundary defines the "top" of the surface which is the side of the surface from which the exterior boundary appears to traverse the boundary in a counter clockwise direction. The interior boundary will have the opposite orientation, and appear as clockwise when viewed from the "top". *Sf_tin* is a *sf_polyhedral_surface* whose minimal parts are *sf_triangle*s (SFC-M12).

(SFC-M9) $sf\_polygon(x) \rightarrow SimpleArealRegion(x) \wedge \exists y,z[sf\_linear\_ring(y) \wedge BCont(y,x) \wedge boundary(z) = y \wedge P(x,z)] \wedge \forall v\big[BCont(v,x) \rightarrow \exists w[P(v,w) \wedge BCont(w,x) \wedge sf\_linear\_ring(w)]\big]$

(***sf_polygon** specializes CODIB's **SimpleArealRegion** such that some linear ring in its boundary bounds a region z of which x is part. This accommodates polygons with and without holes. For polygons with holes, some linear ring describes the polygons "outer boundary", whereas for polygons without holes z = x can be chosen such that z is the entire boundary of x. The second condition expresses that every entity v in the boundary of x must be part of some linear ring that that describes a continuous piece of internal or external boundary of x's entire boundary.*)

(SFC-M10) $sf\_polyhedral\_surface(x) \leftrightarrow SimpleArealRegion(x) \wedge ICon(x) \wedge \forall y[P(y,x) \wedge Min(y) \rightarrow sf\_polygon(y)]$ (***sf_polyhedral_surface** is equivalent to CODIBs **SimpleArealRegion** that is internally-connected and is an aggregation of **sf_polygon**s*)

(SFC-M11) $sf\_triangle(x) \leftrightarrow sf\_polygon(x) \wedge \exists p,q,r[\neg PO(p,q) \wedge \neg PO(p,r) \wedge \neg PO(q,r) \wedge sf\_line(p) \wedge sf\_line(q) \wedge sf\_line(r) \wedge BCont(p,x) \wedge BCont(q,x) \wedge BCont(r,x) \wedge \forall s(sf\_line(s) \wedge BCont(s,x) \rightarrow s = p \vee s = q \vee s = r)]$

(***sf_triangle** is a **sf_polygon** with exactly three non-overlapping lines bounding it*)

(SFC-M12) $sf\_tin(x) \leftrightarrow sf\_polyhedral\_surface(x) \wedge \forall y[Min(y) \wedge PP(y,x) \rightarrow sf\_triangle(y)]$

(***sf_tin** is a polyhedral surface consisting only of **sf_triangle**s as minimal parts*)
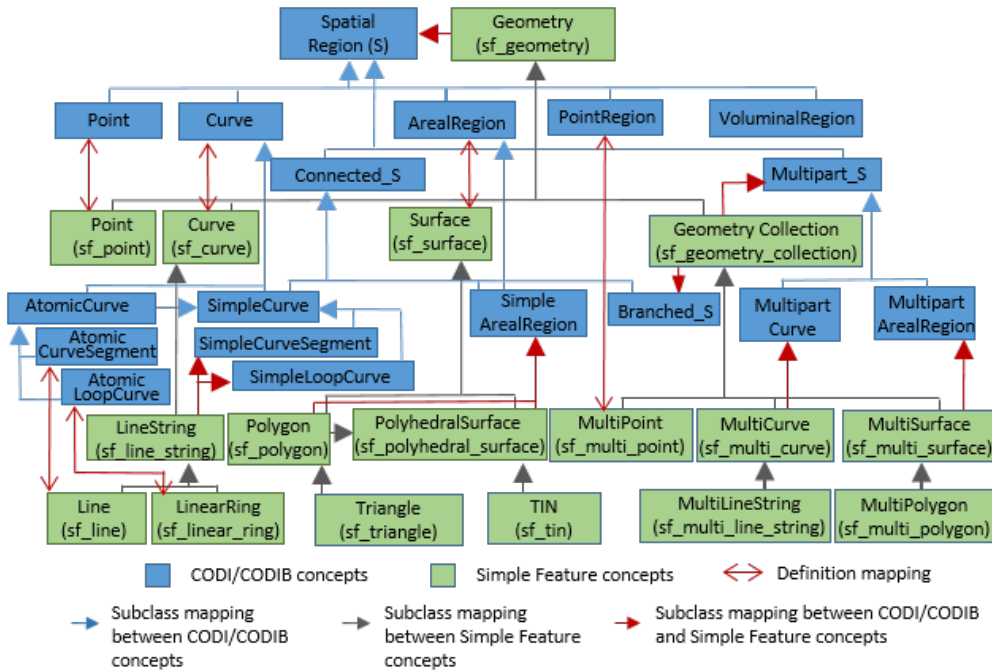
Figure 2 Hierarchy of SF-FOL indicating subclass relationships among SFA concepts, among CODI and CODIB concepts and between SFA and CODI(B) concepts.

4.2 Axiomatization of Simple Feature's Simple Feature Collections

Sf_geometry_collection includes all multipart or branched spatial regions (SFC-M13). Its subclasses map to CODIB's *PointRegion* (SFC-M14) or refine its *Multipart_Curve* (SFC-M15) or *Multipart_ArealRegion* (SFC-M16), respectively, which exhaustively classify *sf_geometry_collection* (SFC-T4). These mappings are not one-to-one because unlike the corresponding CODIB concepts, the SFA concepts restrict how the components can be spatially configured. For example, SFA does not include "branching", non-planar constructions consisting of multiple 2D regions (e.g., three 2D regions meeting in a single line segment) or non-planar arrangements of points. *Sf_multi_line_string* and *sf_multi_polygon* refine *sf_multi_curve* and *sf_multi_surface* (SFC-M17,M18) in that they are constituted only from line strings (i.e., linearly approximated curves) and polygons (i.e., surfaces with linear approximated boundaries).

(SFC-M13) $sf\_geometry\_collection(x) \rightarrow Multipart\_S(x) \vee Branched\_S(x)$

(***sf_geometry_collection specializes CODIB's multipart or branched spatial region***)

(SFC-M14) $sf\_multi\_point(x) \rightarrow PointRegion(x)$

(SFC-M15) $sf\_multi\_curve(x) \rightarrow Multipart\_Curve(x)$

(SFC-M16) $sf\_multi\_surface(x) \rightarrow Multipart\_ArealRegion(x)$

(SFC-T4) $sf\_geometry\_collection(x) \rightarrow PointRegion(x) \vee Multipart\_Curve(x) \vee$
$Multipart\_ArealRegion(x)$ (***SFA's geometry collection is either a PointRegion, Multipart_Curve or Multipart_ArealRegion***)

(SFC-M17) $sf\_multi\_line\_string(x) \leftrightarrow sf\_multi\_curve(x) \wedge \forall y[P(y,x) \wedge Min(y) \rightarrow$
$sf\_line\_string(y)]$

(***sf_multilinestring is a sf_multicurve with minimal parts that are sf_line_strings***)

(SFC-M18) $sf\_multi\_polygon(x) \leftrightarrow sf\_multi\_surface(x) \wedge \forall y[P(y,x) \wedge Min(y) \rightarrow sf\_polygon(y)]$

(***sf_multipolygon is a sf_multisurface with minimal parts that are sf_polygons***)

Table 1 SFA's mereotopological relations, their equivalent *Egenhofer* relations, and the developed mappings to CODIB's relations. The relations in the bottom part are all defined in terms of the top five relations.

SFA	9IM	Definition in terms of CODIB relations and additional theorems
disjoint	disjoint	**(SFR-M1)** $sf\_disjoint(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \neg C(x,y)$
touches	meet	**(SFR-M2)** $sf\_touches(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow BCont(z,x) \wedge BCont(z,y)]$
		(SFR-T1) $sf\_touches(x,y) \rightarrow SC(x,y)$
		(SFR-T2) $sf\_touches(x,y) \rightarrow sf\_point(x) \wedge \neg sf\_point(y)$
crosses	-	**(SFR-M3)** $sf\_crosses(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \big[[Inc(x,y) \wedge \neg Cont(x,y) \wedge \neg Cont(y,x)] \vee \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow Curve(x) \wedge Curve(y) \wedge (z <_{\dim} x \wedge z <_{\dim} y \wedge \neg BCont(z,x) \wedge \neg BCont(z,y))]\big]$
		(SFR-T3) $x <_{\dim} y \wedge sf\_crosses(x,y) \rightarrow Inc(x,y) \wedge \neg Cont(x,y)$
		(SFR-T4) $x =_{\dim} y \wedge sf\_crosses(x,y) \rightarrow SC(x,y)$
		(SFR-T5) $sf\_crosses(x,y) \wedge sf\_curve(x) \wedge sf\_curve(y) \rightarrow SC(x,y)$
overlaps	overlap	**(SFR-M4)** $sf\_overlaps(x,y) \leftrightarrow S(x) \wedge S(y) \wedge PO(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)$
contains	contains $\vee$ covers	**(SFR-M5)** $sf\_contains(x,y) \leftrightarrow S(x) \wedge S(y) \wedge Cont(x,y)$
within	inside $\vee$ coveredBy	**(SFR-M6)** $sf\_within(x,y) \leftrightarrow sf\_contains(y,x)$
equals	equal	**(SFR-M7)** $sf\_equals(x,y) \leftrightarrow sf\_contains(x,y) \wedge sf\_within(x,y)$
intersects	$\neg$ disjoint	**(SFR-M8)** $sf\_intersects(x,y) \leftrightarrow \neg sf\_disjoint(x,y)$
		(SFR-T6) $sf\_intersects(x,y) \leftrightarrow sf\_touches(x,y) \vee sf\_crosses(x,y) \vee sf\_overlaps(x,y) \vee sf\_contains(x,y) \vee sf\_within(x,y)$
		(SFR-T7) $sf\_intersects(x,y) \leftrightarrow S(x) \wedge S(y) \wedge C(x,y)$
relate (any)	-	**(SFR-M9)** $sf\_relate(x,y) \rightarrow sf\_intersects(x,y) \vee sf\_disjoint(x,y))$ **(SFR-T8)** $sf\_intersects(x,y) \leftrightarrow S(x) \wedge S(y)$

The axioms of SFC-Core together with the mappings SFC-M1 to SFC-M18 form the ontology SFC-FOL[15]. The theorems SFC-T1 to SFC-T4 can be proved from SFC-FOL.

4.3 Axiomatization of Simple Feature's Qualitative Spatial Relations

So far we have focused on elaborating the semantics of SFA's feature types using CODIB. But SFA's mereotopological relation can, likewise, be expressed using CODIB's relations as summarized in Table 1, similar to the mapping between the DE-I9 relations and $CODIS$ [18]. All SFA relations, except for *sf_disjoint*, are specializations of contact (C). *Sf_disjoint* is the negation of contact (SFR-M1), which places no dimensional restriction on the involved entities. The relation *sf_touches* relates two connected features who share parts of their boundaries (i.e., $\partial x \cap \partial y \neq \emptyset$) but no parts of their interiors ($x^\circ \cap y^\circ = \emptyset$). This specializes CODIB's superficial contact relation SC that holds for objects that are in contact but do not share a part of either object. But SC is not sufficient as it allows the lower-dimensional entity to share part of its interior with the higher-dimensional entity (e.g., a curve segment tangential

---

[15] Available from `https://colore.oor.net/simple_features`.

to a region). Instead, *sf_touches* needs to express that any shared entities are boundary contained in both of the participating entities (SFR-M2). Then, *SC* becomes provable from it (SFR-T1). From the definition of *SC* it can further be inferred that *sf_touches* applies to entities of any dimension except between two points (SFR-T2).

Sf_crosses is a specialization of one of two of CODIB's relation: (1) incidence *Inc* for two entities of different dimension, where a part of the lower-dimensional entity is contained in the higher-dimensional one (e.g., a curve being incident with a polygon by a segment of the curve being contained in the polygon), or (2) superficial contact *SC* for two entities of equal dimension that share only a lower-dimensional entity (e.g., two curves intersecting in a point) (SFR-M3).

Sf_overlaps is a stronger contact relation that only applies to two equidimensional entities and is equivalent to CODIB's partial overlap *PO* when neither entities is a part of the other (SFR-M4). Full containment of an entity inside another entity of the same spatial dimension is represented in CODI by its primitive containment relation, which maps to *sf_contains* (SFA-M5) and to *sf_within* for its inverse (SFR-M6). The special case of spatial equality is captured by *sf_equals* (SFR-M7). *sf_intersects* is the negation of *sf_disjoint* (SFR-M8), which means it generalizes *sf_touches*, *sf_crosses*, *sf_overlaps*, *sf_contains*, *sf_within*, and, indirectly, *sf_equals* (SFR-T6) and is logically equivalent to CODIB's contact relation (SFR-T7). *sf_relate* describes any of SFA's mereotopological relations (SFR-M9), which maps to any pair of spatial entities in CODIB no matter how they are spatially related (SFR-T8).

The axioms of SFC-Core together with the mappings SFR-M1 to SFR-M9 form the ontology SFR-FOL[16]. The theorems SFR-T1 to SFR-T8 can be proved from SFR-FOL.

4.4 Logical Verification

Our primary tool for evaluating the developed first-order ontology SF-FOL are different variants of consistency checking summarized in Table 2. In its simplest form, consistency checking verifies that an ontology is free of internal contradiction. This typically involves constructing some small finite model using a finite model finder. A known problem with this approach is that it aims to construct the smallest models, which are often trivial in the sense that the extension of many classes and relations therein are empty or universal. For example, one trivial model for CODIB consists of a set of isolated points, but without any curves or areal regions. Moreover, most of the CODIB relations, such as *BCont*, *SC*, or *Inc*, may not be used at all in a trivial model whereas other relations, such as *Cont* or *P*, may relate objects only to themselves. Such a model does not prove that all classes may indeed be instantiated (i.e., some curve, areal region, or more specialized defined subclasses such as a branched curve) and all relation may apply to pairs of distinct entities. One can force the creation of non-trivial models by adding existential axioms of the form $\exists x P(x)$ and $\exists x, y[R(x,y) \wedge x \neq y]$ to the theory. This approach has been implemented in the Macleod suite of tools[17] and previously been utilized to prove CODI's and CODIB's nontrivial consistency with the help of the finite model finder Paradox3 [4]. Here, the same approach is used to prove SF-FOL's nontrivial consistency.

An additional way to verify an ontology is to prove its consistency with some sample datasets. Rather than constructing an arbitrary model that satisfies certain constraints, this external verification ensures that the ontology is actually consistent with the kind of

---

[16] Available from `https://colore.oor.net/simple_features`.
[17] `https://github.com/thahmann/macleod`

Table 2 Overview of the employed consistency checking methods for verification of the developed first-order logic ontology.

Type	Task	Description
Internal verification	Consistency checking	Ascertains the ontology is free of internal contradictions
	Non-trivial consistency checking	Ascertains that a model exists that instantiates each class and each relation positively and negatively by pairs of distinct objects
External verification	Consistency checking with data	Ascertains that the ontology is consistent with a set of assertions describing a dataset

model encountered in the domain. This has not been done previously for CODI or CODIB as real-world purely qualitative information is hard to come by. However, by mapping SFA concepts to CODIB as qualitative generalization thereof, we can now exploit the abundance of geometric data already stored in GIS or geospatial databases.

In this work SF-FOL is verified internally, nontrivially and externally with Paradox3. Proving nontrivial consistency of SF-FOL ensures that instantiation of all the axiomatically defined or restricted Simple Feature types and SFA's mereotopological relations is possible and the new mappings and axioms do not contain any contradictions. In addition, we employed small subsets of data, consisting of samples of 20 to 40 geometric features, to externally verify SF-FOL. The data is extracted from publicly hosted shapefiles[18] that includes polygon representations of counties and subdivisions, polyline representations of major roads, and point representations of schools and other civic buildings within the state of Maine. Only the type of geometry and the SFA relations to other, nearby geometries are stored as assertions. The extracted assertions (i.e., the ABox) were added to SF-FOL (i.e. the TBox) and handed to the model finder to construct a model. As an additional step, we encoded sample queries, such as '*What are the areal regions within Penobscot county that intersect I-95?*', which can be expressed logically in CODIB as $ArealRegion(s) \land sf\_within(s,'PenobscotCounty') \land sf\_intersects(s,'I95')$. This allows retrieving possible instantiations of s, which were manually inspected to identify any unintended models, such as schools being returned as possible solutions, that helped refine the axiomatization.

Generally, the utilized ontology verification techniques are somewhat similar to software testing techniques: they can help identify problematic models of an ontology that require changing or adding axioms but do not prove that the ontology is fully correct. This would require a full representation theorem describing the structure of *all the models* of SF-FOL, which is beyond the scope of this paper. The completeness of SF-FOL is not verified as this would require alternative characterization of all models.

5 Conclusion and Future Work

A core component of many geospatial data models and standards used to store and analyze conventional GIS data are taxonomic classifications of geometric feature types and basic mereotopological relations to support qualitative querying of the geometric data. However, the semantics of the mereotopological relations are not explicitly formalized and thus not

---

[18] https://www.maine.gov/megis/catalog/
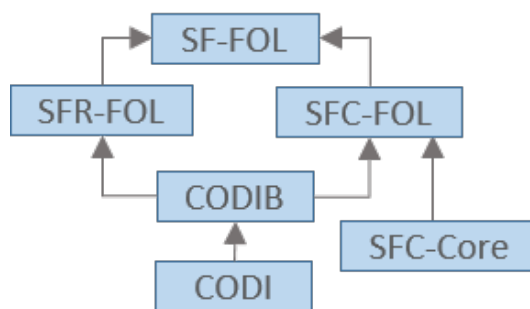
Figure 3 The relationships between the developed and reused axiomatic theories.

accessible for further automated reasoning. Because of this limitation, purely qualitative spatial information, i.e. spatial information that relates objects for which no geometric information is available in the data store, cannot be easily reasoned over in conjunction with existing geometric data. To address this challenge, this paper presents a semantically augmented formalization, SF-FOL, of the basic geometric feature types (axiomatized in SFC-Core) and qualitative spatial relations (axiomatized in SFR-Core) of the Simple Features Access (SFA) standard. This augmented formalization is provided as an extension of the CODIB theory, a qualitative axiomatization of mereotopological space in first-order logic. The relationships between the developed theories is illustrated in Figure 3.

It is shown that all of SFA's geometric features specialize the more general, only dimensionally-constrained, classes of spatial entities from CODIB and its subtheory CODI. The distinctions between "straight line segments" and "curve segments" and, analogously, between "fully bounded regions" and "polygons" are the only ones that are not fully definable in CODIB because they are inherently geometric[19]. But because these distinctions are irrelevant to mereotopological relations, all of CODIB's spatial relations can be evaluated over geometric features in SF-FOL. Likewise, all of SFA's mereotopological relations are fully defined in the SFR-FOL module of SF-FOL and thus can be employed for querying over both geometric and qualitative data.

Future Work: While the mereotopological approach of describing geometric concepts and spatial relations enhances spatial reasoning capabilities, formalization in a language such as first-order logic and relying on general-purpose automated theorem provers and model finders for reasoning comes with the cost of intractability of reasoning. The number of first-order logic (FOL) assertions explodes even when reasoning with a very small spatial dataset. Preliminary experiments with Paradox, one of the best performing FOL model finders, show that reasoning with data against a fairly complex ontology such as CODIB often terminates without success except for the tiniest datasets. In ongoing work, we systematically test how to improve model finding performance by explicitly using the qualitative abstractions and "throwing away" geometric information and by converting data into logically equivalent formats that are less taxing on a model finder.

---

[19] One cannot distinguish a straight line from a curve without a metric in the space that defines the shortest segment between two points, see the discussion of such issues in [2, 20]

References

1   Ringo Baumann, Frank Loebe, and Heinrich Herre. Towards an Ontology of Space for GFO. In *Conf. on Formal Ontology in Inf. Systems (FOIS-16)*, pages 53–66, 2016.

2   Stefano Borgo and Claudio Masolo. Full mereogeometries. *Rev. Symb. Logic*, 3(4):521–567, 2010.

3   Roberto Casati and Achille C. Varzi. *Parts and Places*. MIT Press, 1999.

4   Koen Claessen and Niklas Sörensson. New techniques that improve MACE-style finite model building. In *Workshop on Model Computation at CADE 2003*, 2003.

5   Eliseo Clementini and Paolino Di Felice. A comparison of methods for representing topological relationships. *Inf. Sci.*, 3(3):149–178, 1995.

6   Eliseo Clementini and Paolino Di Felice. A model for representing topological relationships between complex geometric features in spatial databases. *Inf. Sci.*, 90(1):121–136, 1996.

7   Anthony G. Cohn and Jochen Renz. Qualitative Spatial Representation and Reasoning. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation*. Elsevier, 2008.

8   Max J. Egenhofer. Reasoning about binary topological relations. In *Symp. on Large Spatial Databases (SSD'91)*, LNCS 525, pages 141–160. Springer, 1991.

9   Max J. Egenhofer and John Herring. Categorizing binary topological relations between regions, lines, and points in geographic databases. Technical report, Department of Surveying Engineering, Univ. of Maine, 1991.

10 Giorgio De Felice, Paolo Fogliaroni, and Jan Oliver Wallgrün. A Hybrid Geometric-Qualitative Spatial Reasoning System and Its Application in GIS. In *Conf. on Spatial Inf. Theory (COSIT-09)*, 2009.

11 Paolo Fogliaroni, Paul Weiser, and Heidelinde Hobel. Qualitative Spatial Configuration Search. *Spatial Cognition & Computation*, 16(4):272–300, 2016. `doi:10.1080/13875868.2016.1203327`.

12 Anthony Galton. Taking dimension seriously in qualitative spatial reasoning. In *Europ.Conf. on Artif. Intell. (ECAI-96)*, pages 501–505, 1996.

13 Nicholas M. Gotts. Formalizing commonsense topology: the INCH calculus. In *Int. Symp. on Artif. Intell. and Math.*, pages 72–75, 1996.

14 Michael Gruninger, Carmen Chui, and Megan Katsumi. Upper Ontologies in COLORE. In *Proc. of the Joint Ontology Workshops (JOWO 2017)*, 2017.

15 Torsten Hahmann. *A Reconciliation of Logical Representations of Space: from Multidimensional Mereotopology to Geometry*. PhD thesis, Univ. of Toronto, 2013.

16 Torsten Hahmann. On Decomposition Operations in a Theory of Multidimensional Qualitative Space. In *Int. Conf. on Formal Ontology in Inf. Syst. (FOIS 2018)*, pages 173–186, 2018.

17 Torsten Hahmann and Michael Grüninger. A naïve theory of dimension for qualitative spatial relations. In *Symp. on Logical Formalizations of Commonsense Reasoning (CommonSense 2011)*. AAAI Press, 2011.

18 Torsten Hahmann and Michael Grüninger. A Theory of Multidimensional Qualitative Space: Semantic Integration of Spatial Theories that Distinguish Interior from Boundary Contact (Extended Abstract). In *Proc. of the Int. Conference on Spatial Information Theory (COSIT 2011), Belfast, Maine, September 12-16, 2011*, 2011.

19 Torsten Hahmann and Michael Grüninger. Multidimensional mereotopology with betweenness. In *Int. Joint Conf. on Artif. Intell. (IJCAI-11)*, pages 906–911, 2011.

20 Torsten Hahmann and Michael Grüninger. Region-based Theories of Space: Mereotopology and Beyond. In Shyamanta M. Hazarika, editor, *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions*, pages 1–62. IGI, 2012.

21 J Herring. OpenGIS Implementation Standard for Geographic information - Simple feature access-Part 1: Common architecture, 2011.

22 International Electrotechnical Commission (ISO/IEC). ISO 19125:2004 geographic information – simple feature access, 2004.

23 International Electrotechnical Commission (ISO/IEC). ISO 19136:2007 geographic information – Geography Markup Language (GML), 2007.

24 Zhiguo Long, Matt Duckham, Sanjiang Li, and Steven 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`.

25 David M. Mark and Max J. Egenhofer. Modeling spatial relations between lines and regions: Combining formal mathematical models and human subjects testing. *Cartogr. Geogr. Inf. Sci.*, 21(3):195–212, 1994.

26 Mark McKenney, Alejandro Pauly, Reasey Praing, and Markus Schneider. Dimension-refined topological predicates. In *Conf. on Advances in Geographic Information Systems (GIS-05)*, pages 240–249. ACM, 2005.

27 Matthew Perry and John Herring. OGC GeoSPARQL – a geographic query language for RDF data, 2012.

28 David A. Randell, Zhan Cui, and Anthony G. Cohn. A spatial logic based on regions and connection. In *KR'92: Principles of Knowledge Representation and Reasoning*, pages 165–176, 1992.

29 Markus Schneider and Thomas Behr. Topological relationships between complex spatial objects. *ACM Trans. Database Systems*, 31(1):39–81, 2006.

30 Barry Smith. Mereotopology: a theory of parts and boundaries. *Data Knowl. Eng.*, 20(3):287–303, 1996.

31 G Sutcliffe. The TPTP Problem Library and Associated Infrastructure. From CNF to TH0, TPTP v6.4.0. *Journal of Automated Reasoning*, 59(4):483–502, 2017.

Why Classificatory Information of Geographic Regions Is Quantum Information

Thomas Bittner

Departments of Philosophy and Geography, State University of New York,
135 Park Hall, Buffalo, NY 14260, USA
bittner3@buffalo.edu

⎯ Abstract ⎯

This paper gives an information - theoretic argument in support of the claim that there is geographic quantum information. Quantum information is information in the sense of Shannon's information theory, that, in addition, satisfies two characteristic postulates. The paper aims to show that if the density of information (bits per unit of space) that is possible for classificatory geographic qualities is limited, then it follows that the two characteristic postulates of quantum information are satisfied for information about those geographic qualities.

1 Introduction

The aim of this paper is it to give an information - theoretic argument in support of the claim that there is geographic quantum information. For this purpose it is necessary to specify what geographic quantum information is, and how it can be distinguished from classical geographic information. In this context it is assumed that the notion of geographic quantum information arises from the notion of quantum information which in turn arises from the notion of information in Shannon's information theory [13].

At it's core, Shannon's information theory abstracts from what information is about by (a) understanding processes that yield information, i.e., measurements/observations, as interactions that establish correlations between observed and observing systems; and (b) by quantifying the amount of information that one system has of another system as the number of the elements of a set of alternatives out of which the correlation arises [13].

Within this abstract framework the distinction between classical and quantum information arises because classical information satisfies Postulate 1:

▶ **Postulate 1** (Unlimited amount of classical information). *There is an unlimited amount of information that an observing system can obtain from an observed system.*

By contrast, quantum information satisfies Postulates 2 and 3 [7]:

▶ **Postulate 2** (Limited information [11]). *There is a maximum amount of (relevant[1]) information that an observing system can obtain from an observed system.*

▶ **Postulate 3** (Unlimited information [11]). *It is always possible for an observing system to acquire new information about an observed system.*

---

[1] There is some discussion on what "relevant information" is. For details see [11, 7].

Since Postulates 1 and 2 cannot be jointly true, information can be analyzed either within the framework of classical information or the framework of quantum information but not both frameworks simultaneously. Although Postulates 2 and 3 "look" somewhat contradictory, they are known to be consistent [7].

What classical or quantum information is and how the former can be distinguished from the latter can be described in an abstract, information - theoretic framework in the context of the above postulates. Once this is understood, one can inquire why the information that an observing system can have of an observed system is classical or quantum in nature. It is the hypothesis of this paper that (at least in some domains) this question can be answered by analyzing the interrelations of given observing and observed systems and by determining whether or not certain features of the two systems and their interaction *entail* either (a) that the information that the observing system can obtain of the observed system satisfies Postulate 1 and thereby renders the systems and their information-theoretic interrelation classical; or (b) that the information that the observing system can obtain of the observed system satisfies Postulates 2 and 3 and thereby renders the systems and their interrelation non-classical.

For those who accept this analytical and information - theoretic methodology, to argue that information that is obtained by geographic classification and delineation is quantum information, is to argue why it follows from the nature of geographic regions and their qualities that information obtained by geographic classification and delineation satisfies Postulates 2 and 3.

2 Amounts of information

Information is a *discrete* quantity, i.e., there is a minimum amount of information exchangeable: a single bit, or the information that distinguishes between two alternatives. Therefore, the process of acquisition of information (a measurement/ an observation) can be framed as a question that an observing system asks an observed system [18]. Since information is discrete, any process of acquisition of information can be decomposed into acquisitions of elementary bits of information by elementary (i.e., yes/no) questions.

▶ **Proposition 1.** *Only yes/no questions that in a given domain can at least in principle be answered in both ways are in a position to provide information about that domain.*

Proof. Consider the formula $\phi = P \vee \neg P$. A yes answer to the yes/no question Q="Is the formula ϕ true?" does not yield information in any domain in which the law of the excluded middle holds. This is because in those domains there cannot be a no answer to Q. Therefore a yes answer to Q does not exclude any alternatives/possibilities and is void of information. Similarly for questions that cannot have a yes answer.   ◀

Any observed system S is characterized by the elementary yes/no questions that can be asked to it. The answers to a sequence of elementary yes/no questions $(Q_1, Q_2, Q_3, \ldots)$ to S, can be represented as a binary string $(e_1, e_2, e_3, \ldots)$, where each e_i is either 0 or 1 (no or yes) and represents the response of the system S to the elementary question Q_i [18]. In what follows, the focus is on binary strings of length L that represent finite sequences of answers to L elementary yes/no questions. Combinatorially, there are 2^L binary strings of length L. One can identify 2^L *complete* questions $Q^1, \ldots, Q^{2^N}$ that consist of sequences of L elementary yes/no questions such that the complete question Q^i corresponds to the bit string s^i if and only if there is a yes answer to Q^i iff the yes/no answer to the elementary question Q_j is recorded in the bit s_j^i for $1 \leq j \leq L$. This is illustrated in Table the left part of Table 1 for the specific case of two yes/no questions Q_1 and Q_2 which give rise to the set Q_c of $2^2 = 4$ combinatorially possible *complete* questions of length 2 [2, 8, 11].

■ **Table 1** Left: The set $Q_c = \{Q_c^{(i)} \mid 0 \leq i < 4\}$ of 2^4 combinatorially possible complete questions $Q_c^{(i)}$ formed by two yes/no questions Q_1, Q_2; Right: Two sets of complete questions Q_S and Q_R for 2 bits of information.

	$s_c^{(i)}$				$s_S^{(i)}$				$s_R^{(i)}$			
i	Q_1	Q_2	$Q_c^{(i)}$	$Q_S^{(i)}$	Q_1	Q_2	$\bigwedge_i Q_i$	R_1	R_2	$\bigwedge_i R_i$	$Q_R^{(i)}$
0	0	0	$\neg Q_1 \wedge \neg Q_2$	Q_S^1	1	0	$Q_1 \wedge \neg Q_2$	1	0	$R_1 \wedge \neg R_2$	Q_R^1
1	0	1	$\neg Q_1 \wedge Q_2$	Q_S^1	1	0	$Q_1 \wedge \neg Q_2$	0	1	$\neg R_1 \wedge R_2$	Q_R^2
2	1	0	$Q_1 \wedge \neg Q_2$	Q_S^2	0	1	$\neg Q_1 \wedge Q_2$	1	0	$R_1 \wedge \neg R_2$	Q_R^1
3	1	1	$Q_1 \wedge Q_2$	Q_S^2	0	1	$\neg Q_1 \wedge Q_2$	0	1	$\neg R_1 \wedge R_2$	Q_R^2

The fact that there are 2^2 combinatorially possible pattern of answers to two yes/no questions does not guaranty that all the combinatorial possibilities are also kinematically possible, i.e., possible when domain-specific constraints are in place. This is illustrated in the table in the right of Table 1.

Suppose that a string of L bits of information of a system S has been obtained by an observer O. According to Postulate 1 for classical information the fact that only L bits were obtained does not mean that $L + 1$ bits of information could not be obtained by posing complete questions that arise from $L + 1$ yes/no questions. That is, in domains in which Postulate 1 holds, it is always possible to acquire more information by asking longer and longer complete questions.

3 The Wheeler-Sinton paradigm

The systematic investigation of the nature of geographic information and geographic information processing from an information-theoretic perspective was pioneered by Sinton [15]. In the language of Wheeler's [18] information-theoretic view of measurement/observation processes, Sinton's conception of the nature of geographic information can be expressed as follows:

▶ **Postulate 4** (Wheeler-Sinton paradigm). *For a string **s** of L bits of information to count as geographic information of a system S for the observer O, **s** must have the following properties:*

*(WS1) **s** is constituted by bits that result from a yes answer to a* complete *question that is formed by* elementary *yes/no questions that fall in three broad classes: (q) elementary yes/no questions about measurable/observable (geographic) qualities of S; (s) elementary yes/no questions about the location of S in geographic space; and (t) elementary yes/no questions about the temporal location of S.*

*(WS2) The complete question that gives rise to **s** is such that:*

(1) *One type of* elementary *yes/no questions (q), (s), (t) is* fixed*. That is, only one elementary yes/no question of this type has a yes answer and thereby picks out what is fixed.*

(2) *One type of* elementary *yes/no questions (q), (s), (t) is* controlled*. That is, there is a fixed number of yes/no questions of this type – control questions – that must have a yes answer. Yes answers to control questions pick out the cells of a* fiat *subdivision (Def 2).*

(3) *One type of* elementary *yes/no questions (q), (s), (t) is* measured*. That is, every yes answer to a control question is complemented by a yes answer to at least one of the yes/no questions of this type – elementary yes/no questions in Wheeler's standard understanding.*

To control some domain via a fiat subdivision means that this domain is partitioned by a set of cells which boundaries are fiat in nature:

▶ **Definition 2** (Fiat subdivision). *Let x be a (region of) some domain (e.g., a region of geographic space, a range of temperatures, etc.). A set of regions $X = \{x_1, \ldots x_n\}$ is a* subdivision *or a* **partition** *of x iff (1) jointly, the $x_1, \ldots, x_n \in X$ mereologically [14] sum up to x and (2) the members $x_i, x_j \in X$ are pair-wise disjoint [14]. In a* fiat *subdivision X the boundaries of the $x_i \in X$ are not aligned with physical discontinuities of the domain x that is subdivided [16, 17].*

In what follows the Wheeler-Sinton paradigm is employed in the context of fixing time, controlling space via fiat spatial subdivisions, and measuring/observing qualities in cells of fiat subdivisions.

▶ **Example 3** (adapted from [4]). Consider Fig. 1 and suppose that (a) the information that is obtained by an observer O via measurement/observation of some portion of the surface of the Earth (the observed system S) is information about the quality of elevation across S; (b) the information about spatial location is controlled by projecting a fiat [16, 17] raster-shaped partition onto S as indicated in the top left part of the figure; and (c) information about temporal location is fixed by allowing for a single time stamp.

In the context of the Wheeler-Sinton paradigm (a–c) mean: (1) the yes answer to one of the yes/no questions $Q_t^1, \ldots, Q_t^{10}$ picks out a particular time stamp; (2) the yes answers to the yes/no questions $Q_c^1, \ldots, Q_c^{36}$ pick out particular cells in the grid structure projected onto S; (3) for every control region picked out by a control question Q_c^i there is a yes answer to at least one of the yes/no questions $Q_m^1, \ldots, Q_m^{110}$. This is displayed in the table in the top right of Fig. 1. Jointly, the yes/no questions of (1-3) give rise to 2^L possibilities for strings of L bits of information and the associated complete questions. This is partly displayed in the table in the middle of Fig. 1. This table "implements" Sinton's methodology of measure/control/fix as laid out in (a-c) in Wheeler's information-theoretic framework of modeling information obtained by measurement/information using answers to yes/no questions. The constraints imposed by Sinton's paradigm of measure/control/fix allows for a more efficient encoding of L bits of information. The constraints reduce the combinatorial possibilities of 2^L strings of length L to the members of set $\mathsf{S_S}$.

Consider the complete question $Q_S^{img} \in Q_S$ as depicted in the bottom of Fig. 1. A yes answer to Q_S^{img} yields L bits of information. This information is encoded in the string $s_S^{img} \in \mathsf{S_S}$. The same information is encoded (more implicitly) in the image in the top left of Fig. 1.   ◀

In the context of fixing time via the yes answer to the question $Q_f = \{Q_t^i\}$, controlling space via yes answers to the questions $Q_c = \{Q_l^1 \ldots Q_l^j\}$ and yes/no answers to classificatory questions $Q_m = \{Q_m^1 \ldots Q_m^k\}$ a family of complete questions (Sec. 2) arises and is denoted by $\mathcal{Q}(Q_f, Q_c, Q_m)$. A yes answer to a complete question in $\mathcal{Q}(Q_f, Q_c, Q_m)$ yields L bits of information. In the context of Example 3 one has: $Q_S = \mathcal{Q}(Q_f, Q_c, Q_m)$ and $Q_S^{img} \in Q_S$.

Example 3 satisfies Postulate 1: There does not seem to be a limit to the amount of information about elevation across S that can be had by an observer O. More information can be obtained by refining the partition cells and asking yes/no questions about the elevation in these refined cells. Similarly, more information can be obtained by allowing for more precise elevation measurements, i.e., by enlarging the set Q_m. The elevation information of Example 3 is a prototypical example of *classical* geographic information.

645	650	654	658	653	648
664	666	670	672	668	659
678	682	684	693	689	680
703	708	714	721	719	716
728	732	738	744	745	732
730	739	744	749	748	735

time	Q_t^1	Time is 11/20/2018?
	$\dots$	$\dots$
	Q_t^{10}	Time is 11/29/2018?
location	Q_l^1	Location is cell 1?
	$\dots$	$\dots$
	Q_l^{36}	Location is cell 36?
quality	Q_q^1	Elevation is 640?
	$\dots$	$\dots$
	Q_q^{110}	Elevation is 750?

$s_S \backslash Q_S$	Q_t^1	$\dots$	Q_t^{10}	Q_l^1	$\dots$	Q_l^{36}	$(Q_q)_1^1$	$\dots$	$(Q_q)_1^{110}$	$(Q_q)_2^1$	$\dots$	$(Q_q)_2^{110}$	$\dots$	$(Q_q)_{36}^{110}$
s_S^1	1	$\dots$	1	1	$\dots$	1	1	$\dots$	1	1	$\dots$	1	$\dots$	1
s_S^2	1	$\dots$	1	1	$\dots$	1	1	$\dots$	1	1	$\dots$	1	$\dots$	0
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
s_S^L	0	$\dots$	0	0	$\dots$	0	0	$\dots$	0	0	$\dots$	0	$\dots$	0

where $L = 1 + 36 + (36 * 110)$

$$\mathsf{S_S} = \left\{ s_S^i \in s_S \left| \begin{array}{l} (\Sigma_{j=1}^{10} s_S^i[j]) = 1 = s_S^i[1] \ \& \\ (\Sigma_{j=11}^{46+1} s_S^i[j]) = 36 \ \& \\ (\bigwedge_{j=0}^{36-1} (\Sigma_{k=k_1}^{k_m} s_S^i[k] = 1)) \geq 1 \\ k_1 = 10 + 36 + (j * 110) + 1 \\ k_m = 10 + 36 + (j * 110) + 110 \end{array} \right. \right. \begin{array}{l} \text{(fix time)} \\ \text{(36 yes-bits of control information)} \\ \text{(at least one of 110 measurement bits is 1} \\ \text{for each control location)} \end{array} \left. \right\}$$

$Q_S^{img} = \text{Y}$ iff $Q_t^1 = \text{Y} \wedge \neg(Q_t^2 = \text{Y}) \wedge \dots \wedge \neg(Q_t^{10} = \text{Y}) \wedge Q_c^1 = \text{Y} \wedge \dots \wedge Q_c^{36} = \text{Y} \wedge \neg((Q_q)_1^1 = \text{Y}) \wedge \dots \wedge ((Q_q)_1^6 = \text{Y}) \wedge \dots \wedge \neg((Q_q)_1^{110} = \text{Y}) \wedge \neg((Q_q)_2^1 = \text{Y}) \wedge \dots \wedge ((Q_q)_2^{11} = \text{Y}) \wedge \dots \wedge \neg((Q_q)_{36}^{110} = \text{Y})?$

Figure 1 The Wheeler-Sinton paradigm of geographic information: temporal location is fixed (11/20/2018), spatial location is controlled by fiat, and a geographic quality is measured. Q_S^{img} is a complete yes/no question the yes answer to which yields L bits of information. The same information is encoded in the image in the top left. [4] (The image in the top left is from [5].)

4   Geographic regions and their qualities

The purpose of this and the next two sections is it to argue why it follows from the nature of geographic regions and their qualities that information about geographic classification and delineation satisfies Postulates 2 and 3 and thereby entails the quantum nature of classification and delineation information. The focus of the discussion is on Bailey's classification and delineation of ecoregions [1]. In this discussion the distinction between quality determinables and quality determinates [12] is taken for granted. Quality determinables are qualities such as energy, temperature, climate type, climate regime, land-surface quality, etc. Those quality determinables subsume the quality determinates such as 100 Joule, 70 degrees Fahrenheit, Irregular planes, etc.

▶ **Definition 4** (Classificatory quality determinables). *The determinable ϕ is a classificatory quality determinable if and only if ϕ has the following properties: (i) the quality determinates of ϕ are the leafs of a finite isA tree; (ii) ϕ is the immediate parent of its determinates in the isA tree; (iii) no instance of ϕ can fail to instantiate one of ϕ's determinates; and (iv) distinct quality determinates cannot be instantiated in partially overlapping entities/regions.*

Examples of classificatory quality determinables include land-surface qualities such as land-surface form, and climate qualities such as climate types and climate regimes (Fig. 2).
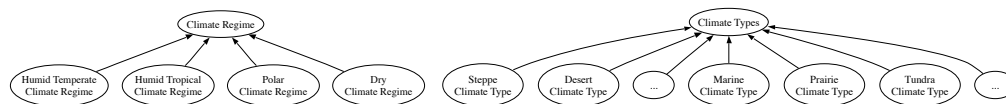
■ **Figure 2** Climate regime qualities (left) and Climate type qualities (right).

▶ Remark 5. Condition (iv) of Def. 4 is consistent with the classical intuition that distinct classificatory determinates cannot be co-instantiated in the same region. It is also consistent with the possibility of the superposition of multiple qualities in the same region. This is because no entity/region can only partially overlap itself. (If x partially overlaps x then there is a part of x that is not a part of x.)

4.1   Land-surface and climate qualities

There are at least four land-surface determinables [10, Table 1]: *Land-surface form*, potential natural vegetation (*Climax Vegetation*), *Land use*, and *Soil type*. Determinates such as *Irregular Plains* inhere in specific regions such as the Central Great Plains (CGP). In the body of Table 2 a number of quality determinates that inhere in CGP are listed.

■ **Table 2** Land-surface qualities of geographic regions – Central Great Plains as an example [10, Table 1].

Geographic region	*Land-surface form* (*LF*)	*Climax vegetation* (*CV*)	*Land use* (*LU*)	*Soil type* (*S*)
. . .	. . .	. . .	. . .	. . .
Central Great Plains	Irregular plains	Bluestem / grama prairie, bluesteam prairie, buffalo grass	Cropland, cropland with grazing land, some irrigated agriculture	Dry Mollisols
. . .	. . .	. . .	. . .	. . .

Climate qualities fall into at least two major groups [1]: climate types and climate regimes [9, 1]. Relevant climate types are listed in the right of Fig. 2. Relevant climate regimes are listed in the left of Fig. 2.

4.2   Taxonomic structure

In Bailey's system geographic regions that are characterized by a given climate regime determinate are called *domains*. The geographic regions that count as domains and that overlap the territory of the contiguous US are displayed in the top left map of Fig. 3. Geographic regions that are characterized by climate types (in the context of a given climate regime) are called *divisions*. The geographic regions that count as divisions and that overlap the territory of the contiguous US are displayed in the bottom left map of Fig. 3. Finally, geographic regions that are characterized by Land-surface types (in the context of a given climate type) are called *provinces*. In Fig. 4 the solid arrows represent the taxonomic hierarchy for ecoregions. A dotted line connecting a quality determinable to a kind of ecoregion indicates that a quality determinate of this determinable is instantiated at the respective ecoregion and thereby determines the kind of this ecoregion. For details, see [3].
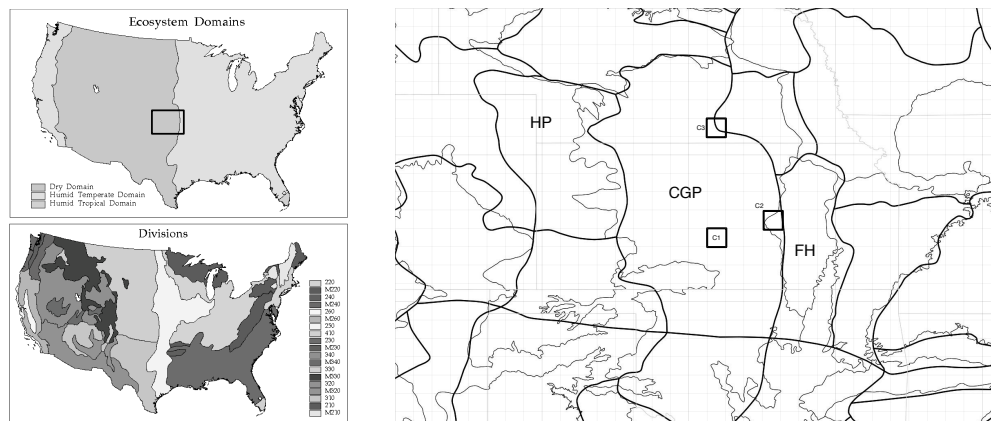
Figure 3 Ecoregion domains of the US (top left); Ecoregion Devisions of the US (bottom left); right: Ecoregions of regional scale in Bailey's framework (bold black boundaries) and in the framework of the EPA (gray boundaries).

4.3 Partonomic structure and scale

In Bailey's system land-surface qualities and climate qualities also give rise to a partonomic nesting of the geographic regions in which they inhere – the partonomic hierarchy of ecoregions is depicted in Fig. 4 using dashed arrows. The partonomic structure ensures that regions in which climate type qualities inhere (divisions) are (proper) parts of regions in which climate regime qualities inhere (domains) and that regions in which land-surface qualities inhere (provinces) are (proper) parts of regions in which climate type qualities inhere (divisions).

In conjunction with the partonomic structure Bailey also identifies scales of regions in which certain kinds of qualities inhere: climate regime qualities mainly inhere in geographic regions of global scale, climate type qualities mainly inhere in geographic regions of continental scale, and land-surface form qualities mainly inhere in geographic regions of regional scale (Fig. 4).

▶ **Postulate 5** (Qualities and scale)**.** *The relation between classificatory quality determinables such as* Land-surface form*,* Climate regime*,* Climate regime *and regions of the surface of the Earth is* scale dependent*: Every quality determinable has an associated scale which constrains the size of the regions in which its determinates can be instantiated.*

Consequently, classificatory qualities such as *Land-surface form*, *Climate regime*, *Climate regime* are non-dissective [6].

5 Classificatory geographic information

In Bailey's ecoregion framework geographic regions are classified according to their climate regimes, climate types and land-surface qualities (classificatory quality determinables). In the Wheeler-Sinton paradigm classificatory information that can be obtained by measurement/observation at given controlled locations/regions is information about instantiated quality determinates that fall under those classificatory quality determinables (Fig. 2, Tab. 2). In this context the choice of controlled locations/regions is *constrained* by the scale of the regions in which the respective quality determinates can inhere.

▶ **Example 6.** Suppose that information is sought about the instantiation of the quality determinates (e.g. prairie climate type, Q_m^i) of a given classificatory quality determinable (e.g., climate type, Q_m). In the Wheeler-Sinton framework this information is obtained via a
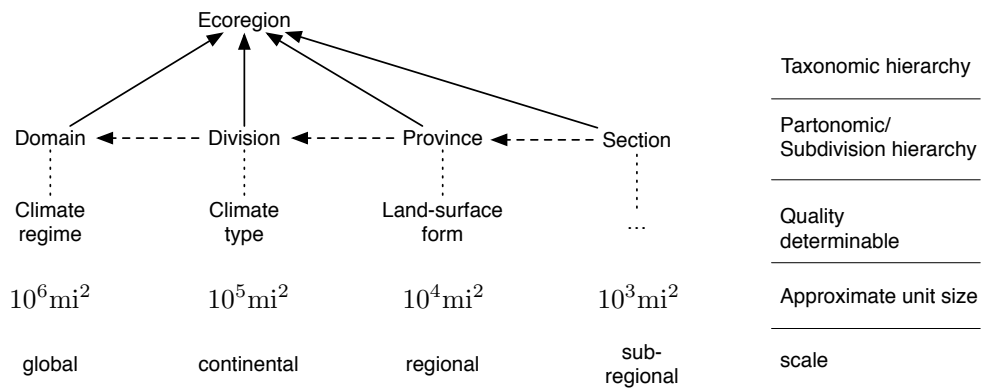
Figure 4 Taxonomic and partonomic hierarchy in Bailey's [1] classification and delineation system for ecoregions (adapted from [3]).

yes answer to a yes/no question Q_m^j that is posed to the control region identified by a yes answer to a control question Q_l^i. (As above, the same symbol, Q_m^j, is used for the yes/no question and the quality determinate of which information about instantiation is sought. Similarly the same symbol is used for the control questions and the control regions they pick out.) A yes answer to a question Q_m^j posed to a control region Q_l^i is possible only if Q_l^i is of a scale (continental scale, order of 10^5 mi^2) that permits the instantiation of a quality determinate Q_m^j of which information is to be acquired. ◀

This example motivates the following granularity postulate:

▶ **Postulate 6** (Granularity [4]). *If the Wheeler-Sinton scheme is applied in contexts in which (a) information about the instantiation of the quality determinates $Q_m^1 \ldots Q_m^h$ of the classificatory quality determinable Q_m is to be obtained by answers to yes/no questions $Q_m^1 \ldots Q_m^h$; (b) time is fixed; and (c) space is controlled via control cells referenced by yes answers to control questions of the form $Q_l^1 \ldots Q_l^n$, then there is a minimal size of control cells – a finest level of resolution/granularity – for which elementary yes/no questions of the form:*

"Does the control cell Q_l^i instantiate the quality determinate Q_m^j?"

still can have a yes answer.

As pointed out in Proposition 1, only yes/no questions that *can* have a yes answer as well as a no answer when posed to a system S are in a position to yield information about S. This is because only if a question can be answered in either way, then a specific yes or no answer rules out possibilities and thereby provides information in Shanon's sense. From this information-theoretic perspective in conjunction with Postulate 6 it follows:

▷ **Claim 7.** There is a minimal size of control cells – a finest level of resolution/granularity – below which no information about the instantiation of classificatory quality determinates that are associated with this level of resolution can be obtained.

Support for this claim is provided in what follows in the specific context of climate and land surface form qualities in Bailey's ecoregion framework.

5.1 Information about climate regime qualities

Consider information that can be obtained by the measurement/observation of climate regime qualities in the context of the Wheeler-Sinton paradigm applied to Bailey's ecoregions of global scale. As displayed in the left of Fig. 2 there are four climate regime determinates $Q_{\text{CRQ}}^1 \ldots Q_{\text{CRQ}}^4$ which inhere, according to Bailey, in regions of global scale at the order of 10^6 mi^2 (Fig. 4). A lower bound to regions in which climate regime qualities can inhere arises from the order of the size of regions at which climate types are instantiated in conjunction with the partonomic nesting of ecoregions (Fig. 4). This lower bound on the size of regions in which the climate regime qualities $Q_{\text{CRQ}}^1 \ldots Q_{\text{CRQ}}^4$ can inhere is somewhere between the order of 10^5 mi^2 and 10^6 mi^2.

Consider the map in the right of Fig. 2. The area covered by this map as a whole corresponds to the black rectangle in the top left map of the figure. This region, CGP-Map, – which is larger than the state of Kansas – may still be too small for a climate regime determinate to be instantiated in it. If the region CGP-Map is indeed too small for climate regimes to inhere in this region, then the question

$$Q_{\text{CRQ}}^{\textsf{CGP-Map}} \quad = \quad \text{``Does the region CGP-Map referenced by a yes answer to the}$$
$$\text{question } Q_l^{\textsf{CGP-Map}} \text{ instantiate the quality determinate } Q_{\text{CRQ}}^i?\text{''}$$

cannot have a yes answer for any of the four climate regime determinates (Q_{CRQ}^i with $1 \le i \le 4$). Since a yes answer is impossible, a no answer does not provide information.

▶ Remark 8. The lack of information that can be obtained from a no answer to the yes/no question $Q_{\text{CRQ}}^{\textsf{CGP-Map}}$ does NOT mean that similar information cannot be *inferred* from yes/no questions that collect information of qualities that can inhere in CGP-Map. The partonomic structure of ecoregions (Fig. 4) in conjunction with information about qualities that are instantiated at a part of a region may make it possible to *infer* information about qualities of that larger embedding region [3]. See also Remark 9.

5.2 Information about climate type qualities

Now consider information that can be obtained by the measurement/observation of climate type qualities in the context of the Wheeler-Sinton paradigm applied to geographic regions of continental scale. As displayed in the left of Fig. 2, in Bailey's framework there are more than five climate type determinates $Q_{\text{CTQ}}^1 \ldots Q_{\text{CTQ}}^{\#\text{CTQ}}$ ($\#\text{CTQ} > 5$) which inhere in regions of continental scale at the order of 10^5 mi^2 (Fig. 4). Again, a lower bound arises from the order of the size of regions at which land-surface qualities are instantiated in conjunction with the partonomic nesting of ecoregions (Fig. 4). This means that information of the classificatory qualities $Q_{\text{CTQ}}^1 \ldots Q_{\text{CTQ}}^{\#\text{CTQ}}$ can be obtained by measurement/observation of those qualities in (control) cells of subdivisions in which the cell size is between the order of 10^4 mi^2 and 10^5 mi^2. In this case the region CGP-Map is of the right scale and the question

$$Q_{\text{CTQ}}^{\textsf{CGP-Map}} \quad = \quad \text{``Does the region CGP-Map referenced by a yes answer to the}$$
$$\text{question } Q_l^{\textsf{CGP-Map}} \text{ instantiate the quality determinate } Q_{\text{CTQ}}^i?\text{''}$$

can have a yes answer for all of the $1 \le i \le \#\text{CTQ}$ climate type determinates. In particular, there is a yes answer to the question

$$Q_{\text{Steppe}}^{\textsf{CGP-Map}} \quad = \quad \text{``Does the region CGP-Map referenced by a yes answer to the}$$
$$\text{question } Q_l^{\textsf{CGP-Map}} \text{ have the quality } Q_{\text{CTQ}}^{\text{Steppe}}?\text{''}$$

The yes answer to $Q_{\text{Steppe}}^{\text{CGP-Map}}$ does yield information to the effect that the climate type *Steppe climate* (as opposed to the other possible climate type qualities) is instantiated in the region CGP-Map.

▶ **Remark 9.** In Bailey's system climate types are defined in the context of given climate regimes. From a yes answer to the question $Q_{\text{Steppe}}^{\text{CGP-Map}}$ it can be *inferred* in Bailey's system that CGP-Map is a (proper) part of a region of continental scale in which the quality determinate *temperate climate regime* inheres. As pointed out in Remark 8, this kind of inference is based on structural properties within Bailey's framework and not on the information-theoretic principles of Wheeler and Sinton.

5.3   Information about land surface qualities

Finally, consider information that can be obtained by the measurement/observation of land-surface qualities in the context of the Wheeler-Sinton paradigm applied to geographic regions of regional scale. Suppose that there are $\#_{\text{LSQ}}$ land-surface determinates $Q_{\text{LSQ}}^1 \ldots Q_{\text{LSQ}}^{\#_{\text{LSQ}}}$ which can inhere in regions of regional scale at the order of 10^4 mi^2 (Fig. 4). Again, a lower bound arises from the order of size of regions at which sub-regional qualities are instantiated in conjunction with the partonomic structure that is entailed in Bailey's system (Fig. 4). This means that the classificatory qualities $Q_{\text{LSQ}}^1 \ldots Q_{\text{LSQ}}^{\#_{\text{LSQ}}}$ are to be measured/observed in control cells which size is between the order of 10^3 mi^2 and 10^4 mi^2. In this case the region $cell_1$ – the cell with bold boundaries labeled c_1 that is part of the region CGP in the right of Fig. 3 – is of a scale that is compatible with the measurement/observation of land-surface qualities. Therefore, the question

$$Q_{\text{LSQ}}^{cell_1} = \text{“Does the region } cell_1 \text{ referenced by a yes answer to the question } Q_l^{cell_1}$$
$$\text{have the quality (pattern) } Q_{\text{LSQ}}^i?\text{”}$$

can have a yes (as well as a no) answer for all Q_{LSQ}^i with $1 \leq i \leq \#_{\text{LSQ}}$. In fact, there is a yes answer to the question:

$$Q_{\text{LSQ}}^{cell_1} = \text{“Does the region } cell_1 \text{ referenced by a yes answer to the question } Q_l^{cell_1}$$
$$\text{have the land surface form irregular plains and (some of) the climax}$$
$$\text{vegetation listed in row three of Tab. 2, (some of) the land uses listed in}$$
$$\text{row three of Tab. 2, and (some of) the Soil type listed in row three of}$$
$$\text{Tab. 2?”}$$

The yes answer to $Q_{\text{LSQ}}^{cell_1}$ encodes the information about the respective land surface qualities in the region $cell_1$.

The region $cell_1$ is a cell in a fiat subdivision of the area that is depicted in the right of Fig. 3. This fiat subdivision arises from the control of space that is imposed in the context of the Wheeler-Sinton paradigm. The control by fiat ensures that all the cells are of the same size and therefore all the control questions that are associated with each of the control cells have an answer that yields information. The fiat subdivision in the figure gives rise to $32 * 22 = 704$ control questions. In what follows the symbol $\#_{\text{LSC}}$ is used for the number of control cells which size is between the order of 10^3 mi^2 and 10^4 mi^2. Similar but coarser subdivisions arise in the context of the application of the Wheeler-Sinton paradigm to climate regime and climate type qualities.

The sets $\mathcal{Q}_{\text{LSQ}}$, $\mathcal{Q}_{\text{CTQ}}$, and $\mathcal{Q}_{\text{CRQ}}$ of complete questions about the measurement/observation of land-surface qualities as well as climate regime and climate type qualities arise in the same way as discussed in Example 3. Yes answers to those complete questions give rise

respectively to the set $\mathsf{S}_{\mathrm{LSQ}}$ of bit strings of information about land-surface qualities, the set $\mathsf{S}_{\mathrm{CTQ}}$ of bit strings of information about climate type qualities, and the set $\mathsf{S}_{\mathrm{CRQ}}$ of bit strings of information about climate regime qualities.

6 Quantum information

In contrast to the unlimited amount of classical information that is possible (Postulate 1, Example 3), the amount of quantum information that an observing system O can obtain about an observed system S is limited as specified in Postulate 2. Any quantum system (S, O) has a maximal information capacity L, where L is an amount of information in bits. L bits of information exhaust the amount of information an observing system O can have about the observed system S.

6.1 Multiple families of complete questions

Given the limited information capacity that characterizes a quantum system (S, O), the question arises how Postulate 3 can possibly be true of (S, O). In the context of the example illustrated in Table 1 the set $Q_S = \{Q_S^1, Q_S^2\}$ of kinematically possible complete questions for two bits of information was introduced. Limiting the amount of information to two bits does not entail that $Q_S = \{Q_S^1, Q_S^2\}$ is the only kinematically possible set of complete questions. That is, there may be a second set of two yes/no questions $\{R_1, R_2\}$ which give rise to two bits of information via the set $Q_R = \{Q_R^1, Q_R^2\}$ of kinematically possible complete questions. The logical structure of the questions in Q_R mirrors the structure of the questions in Q_S as illustrated in Table 1.

Postulate 3 captures what happens if, after having obtained $L = 2$ bits of information by asking Q_S^1, O asks another question, say Q_R^1, as permitted by Postulate 3. Jointly, Postulates 2 and 3 can be understood as follows [11]: Since the amount of information that O can have about S is limited by Postulate 2, it follows that, if O has a maximal amount of information about S, then, if new information about S is acquired by O, O must loose information. In particular, if a yes answer to the question Q_S^1 is followed by a yes answer to the question Q_R^1, then the information obtained by O via a yes answer to Q_R^1 overwrites the information obtained by O via a yes answer to Q_S^1. In virtue of the information that O obtains by a yes answer to Q_R^1, O looses all of its two bits of information that was obtained by a yes answer to Q_S^1. If the yes answer to the question Q_R^1 is in turn followed by a yes answer to the question Q_S^1, then genuinely new information about S is obtained by O. And so on.

6.2 Complete questions, control and resolution

Consider the information of a system of geographic regions (the observed system S) where, in the framework of the Wheeler-Sinton paradigm, the observer O fixes time, controls space (by partitioning S into raster cells), and measures/observes classificatory geographic qualities that can inhere in S. Let $\mathcal{Q}(Q_t, Q_l, Q_m)$ be a set of complete question that O can pose to S. Q_m is the set of yes/no questions about measurable/observable quality determinates of the classificatory quality determinable Q_m. Here and in what follows, the symbol Q_m is used for the quality determinable, as well as for the set of quality determinates that fall under the determinable Q_m, as well as for the set of yes/no questions that gather information about the instantiation of the determinates that fall under Q_m. The context will disambiguate.

In the framework of the Wheeler-Sinton paradigm and its families of complete questions that obtain information about the instantiation of classificatory quality determinates, families of *maximally* complete questions are defined as follows:

▶ **Definition 10** (Maximally complete questions). *Let $\mathcal{Q}(Q_f, Q_l, Q_m)$ be a family of complete questions that is associated with the classificatory quality determinable Q_m. The subdivision that is picked out by the control questions in $Q_l = \{Q_1, \ldots, Q_{\#C}\}$ is of* maximal *resolution if and only if (1) Q_l is the set of control questions with associated control cells of (roughly) equal size in the range of order in which all the complete questions in $\mathcal{Q}(Q_f, Q_l, Q_m)$ can have a yes answer; (2) there is a set of control questions $Q'_l = \{Q'_1, \ldots, Q'_{2*\#C}\}$ that arises when refining the subdivision picked out by $Q_l = \{Q_1, \ldots, Q_{\#C}\}$ by replacing every cell of Q_l by two cells of (roughly) equal size; and (3) the complete questions in $\mathcal{Q}(Q_f, Q'_l, Q_m)$ with the measurable/observable classificatory quality determinates in Q_m and the control questions $Q'_1, \ldots, Q'_{2*\#C}$ fail to yield information because the Q'_i are too small to instantiate the determinates in Q_m.*

A family $\mathcal{Q}(Q_f, Q_l, Q_m)$ of complete questions that is associated with the classificatory quality determinable Q_m is maximal if and only if the resolution of the subdivision that is picked out by Q_l is of maximal resolution.

One can prove the following proposition:

▶ **Proposition 11** (adapted from [4]). *Let $\mathcal{Q}(Q_f, Q_l, Q_m)$ be a set of complete questions that is associated with the classificatory quality determinable Q_m. The members of $\mathcal{Q}(Q_f, Q_l, Q_m)$ have a maximal information capacity (and therefore satisfy Postulate 2), only if the control questions in Q_l pick out cells of* maximal *resolution in the sense of Def. 10.*

Proof. Since every question in $\mathcal{Q}(Q_f, Q_l, Q_m)$ is complete and adheres to the Wheeler-Sinton scheme, a yes answer to any of the complete questions in $\mathcal{Q}(Q_f, Q_l, Q_m)$ yields the amount of L bits of information about the instantiation of quality determinates in Q_m in the cells that are picked out by the control questions in Q_l. Since Q_m is a classificatory quality determinate, it has a finite and fixed number of quality determinates (Def. 4), more information about the instantiation of quality determinates in Q_m in (parts of) the observed system S can be had only by further subdividing the control cells in Q_l. But such a refinement would render the questions that are associated with the cells of the refined subdivision void of information because those questions cannot have a yes answer. This is because, by assumption, the cells that are associated with the control questions in Q_l are already of maximal resolution in the sense of Def. 10. Thus, the amount of information of complete questions in $\mathcal{Q}(Q_f, Q_l, Q_m)$ associated with control questions in Q_l that acquire information about instantiation at cells of maximal resolution is maximal. Hence Postulate 2 is satisfied. ◀

6.3 Multiple families of maximally complete questions

Proposition 11 demonstrated that there is a limit to the amount of information about the instantiation of certain kinds of classificatory quality determinates that is possible. The question now arises whether Postulate 3 is true for such systems with limited information capacity.

Let $\mathcal{Q}(Q_t, Q_l, Q_m)$ be a set of maximally complete question that O can pose to S and suppose that (1) the subdivision that is picked out by Q_l is raster-shaped, and (2) a yes answer to a question in $\mathcal{Q}(Q_t, Q_l, Q_m)$ delivers L bits of information to O about S. Limiting the amount of information that O can have of S to L bits does not entail that $\mathcal{Q}(Q_t, Q_l, Q_m)$ is the only set of maximally complete questions that O can pose to S in the context of the Wheeler-Sinton paradigm. There may be a second set $\mathcal{Q}'(Q_t, Q'_l, Q_m)$ of maximally complete questions which give rise to L bits of information when posed by O to S. Let Q'_l be the set of yes/no questions that pick out the cells of a subdivision of S that arises when the

(raster-shaped) subdivision that is associated with Q_l is translated by half of a cell size either along the rows or along the columns of the partition. From the construction of $\mathcal{Q}(Q_t, Q_l, Q_m)$ and $\mathcal{Q}'(Q_t, Q_l', Q_m)$ it follows that: (a) the family of complete questions $\mathcal{Q}(Q_t, Q_l, Q_m)$ has an information capacity of L bits if and only if $\mathcal{Q}'(Q_t, Q_l', Q_m)$ has an information capacity of L bits; and (b) $\mathcal{Q}(Q_t, Q_l, Q_m)$ is a family of maximally complete questions if and only if $\mathcal{Q}'(Q_t, Q_l', Q_m)$ is.

▶ **Proposition 12.** *Let $\mathcal{Q}(Q_t, Q_l, Q_m)$ and $\mathcal{Q}'(Q_t, Q_l', Q_m)$ be as above. For every maximally complete question $Q \in \mathcal{Q}(Q_t, Q_l, Q_m)$ and for every maximally complete question $Q' \in \mathcal{Q}'(Q_t, Q_l', Q_m)$: A yes answer to the question Q posed by O to S that is (immediately) followed by a yes answer to the question Q' leaves O with a total of L bits of information about S.*

Proof. Assume that O has zero bits of information of S before asking Q. A yes answer to Q then gives O exactly L bits of information about S. Now suppose that a yes answer to Q' after a yes answer to Q leaves O with $L + 1$ bits of information about S. From the construction of the questions Q and Q' and the underlying raster-shaped subdivisions, it follows, that there must be two control cells $c_1, c_2 \in Q_l$ and one control cell $c' \in Q_l'$ such that c_1 and c_2 both partially overlap c' and jointly contain c' in their mereological sum. The following cases are relevant: (a) $q_1 = q_2 = q'$ where q_1, q_2, q' are the qualities that are instantiated respectively at c_1, c_2 , c' according to the information provided by S to O. In this case Q' cannot yield any new information. Thus, a yes answer to Q' following a yes answer to Q leaves O with exactly L bits of information. This contradicts the assumption that $L + 1$ bits of information were obtained; (b) The second case is $q_1 \neq q'$ or $q_2 \neq q'$. Focus on $q_1 \neq q'$ and the regions $c_1 \cap c'$ and $c_1 \setminus c'$. It follows that if $L + 1$ bits of information can be obtained by O from a yes answer to Q' following a yes answer to Q, then this information originates from the instantiation of q' in the region $c_1 \cap c'$ and q_1 in $c_1 \setminus c'$. This is because, by Def. 4, distinct determinates of a classificatory quality determinable cannot inhere in partially overlapping regions. Thus, via the additional bit of information, jointly Q and Q' provide information about instantiation at regions that are of roughly half the size of cells of maximal resolution. This contradicts the assumption that the questions Q and Q' are maximally complete questions. Thus, a yes answer to Q' after a yes answer to Q does not leave O with $L + 1$ bits of information about S. Similarly for the sub-case $q_2 \neq q'$ of (b).

Since the thesis that a yes answer to Q' after a yes answer to Q leaves O with $L + n$ bits of information about S can be ruled out for $n = 1$, the thesis can also be ruled out for $n > 1$. ◀

From Proposition 12 it follows that Postulate 3 is satisfied for families of maximally complete questions of classificatory quality determinates:

▶ **Corollary 13.** *It is always possible to obtain new classificatory information.*

Proof. Consider the two families of maximally complete questions $Q \in \mathcal{Q}(Q_t, Q_l, Q_m)$ and $Q' \in \mathcal{Q}'(Q_t, Q_l', Q_m)$ that O can ask S. Suppose that O obtains L bits of information via a yes answer to the question Q' after having obtained L bits of information via a yes answer to the question Q. On those assumptions the L bits obtained via a yes answer to Q' must overwrite/erase/destroy the L bits of information previously obtained via a yes answer to Q. This is because the information capacity of L bits cannot be exceeded (Proposition 12). Therefore a yes answer to the question Q following a yes answer to the question Q' yields genuinely new information: information that was not available (anymore) to O after obtaining L bits of information via the yes answer to Q'. ◀

The questions $Q \in \mathcal{Q}(Q_t, Q_l, Q_m)$ and $Q' \in \mathcal{Q}'(Q_t, Q'_l, Q_m)$ are geographic examples of what in quantum information theory are called *complimentary* questions.

7   Information about geographic delineation

Now suppose that there is a third set, $\mathcal{Q}(Q_t, Q_l, Q^{i/b})$, of complete questions which enable the observing system O to obtain information about discontinuities in the distribution of classificatory qualities across the observed system S. Those questions obtain information about the location of bona fide boundaries [17] that separate regional parts of S in which distinct classificatory determinates are instantiated. In what follows the members of $\mathcal{Q}(Q_t, Q_l, Q^{i/b})$ are called *delineatory* questions while the members of $\mathcal{Q}(Q_t, Q_l, Q_m)$ are called *classificatory* questions.

The set $\mathcal{Q}(Q_t, Q_l, Q_m)$ of complete classificatory questions and the set $\mathcal{Q}(Q'_t, Q'_l, Q^{i/b})$ of complete delineatory questions are *compatible* if and only if they share the same yes/no questions for fix/control, i.e., $Q_t = Q'_t$ and $Q_l = Q'_l$. While in $\mathcal{Q}(Q_t, Q_l, Q_m)$ the fix/control questions are complemented by classificatory yes/no questions, in $\mathcal{Q}(Q'_t, Q'_l, Q^{i/b})$ the fix/control questions are complemented by delineatory yes/no questions of the form:

- $(Q_{i/b})^1_j$: "Is the cell associated with a yes answer to the control question Q^j_l an interior part of a region in which a classificatory quality determinate of Q_m inheres?"

- $(Q_{i/b})^2_j$: "Does the cell associated with a yes answer to the control question Q^j_l contain a boundary which separates regions in which distinct classificatory quality determinates of Q_m inhere?"

According to Fig. 3, the answer to the question $(Q^{i/b}_{\text{LSQ}})^1_{cell_1}$ is yes and the answer to question $(Q^{i/b}_{\text{LSQ}})^2_{cell_1}$ is no. By contrast, the answer to question $(Q^{i/b}_{\text{LSQ}})^1_{cell_3}$ is no and the answer to question $(Q^{i/b}_{\text{LSQ}})^2_{cell_3}$ is yes.

The family $\mathcal{Q}(Q_t, Q_l, Q^{i/b})$ of maximally complete delineatory questions and the family $\mathcal{Q}(Q_t, Q_l, Q_m)$ of maximally complete classificatory questions are complementary. This is because if a yes answer to $Q_b \in \mathcal{Q}(Q_t, Q_l, Q^{i/b})$ following a yes answer to $Q_c \in \mathcal{Q}(Q_t, Q_l, Q_m)$ could yield more than L bits of information then so could a complementary classificatory question $Q'_c \in \mathcal{Q}'(Q_t, Q'_l, Q_m)$. But this would contradict Propositions 12 and 13.

8   Conclusion

The arguments of the previous sections about the quantum nature of maximally complete classificatory (and delineatory) questions depended critically on the assumption that the underlying control questions refer to cells at a maximal level of resolution. Linking a maximal amount of information to a minimal unit of space, as it is evident in Postulate 6 as well as in Propositions 11 and 12, makes explicit that, in the context of the processing of classificatory and delineatory information about geographic regions, there is a maximal *information density* associated with every classificatory quality determinable Q_m. The notion of maximal information density then opens the possibility that larger amounts of information can be had at coarser levels of granularity. At those coarser levels the density of information would be lower and quantum information would behave very much like classical information.

References

1 R. G. Bailey. *Ecosystem Geography*. Springer Verlag, Berlin, 1996.

2 Enrico G. Beltrametti, Gianni Cassinelli, and Peter A. Carruthers. *The Logic of Quantum Mechanics*, volume 15 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1984. `doi:10.1017/CBO9781107340725`.

3 T. Bittner. Vagueness and the tradeoff between the classification and delineation of geographic regions. *Int. J. Geographical Information Systems (IJGIS)*, 25(5):825–850, 2011.

4 Thomas Bittner. Is there a quantum geography? In Timothy Tambassi, editor, *Philosophy of GIS*. Springer Verlag, 2020.

5 P. Bolstad. *GIS Fundamentals*. Eider Press, Minnesota, 2005.

6 P. Grenon and B. Smith. SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation*, 4(1):69–103, 2004.

7 Alexei Grinbaum. Information-Theoretic Princple Entails Orthomodularity of a Lattice. *Foundations of Physics Letters*, 18(6):563–572, November 2005. `doi:10.1007/s10702-005-1129-0`.

8 R. I. G. Hughes. Quantum Logic. *Scientific American*, 245(4):202–213, 1981. URL: `http://www.jstor.org/stable/24964585`.

9 W. Koeppen. *Grundriss der Klimakunde*. W. de Gruyter, Berlin, 1931.

10 James M. Omernik. Map Supplement: Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers*, 77(1):118–125, 1987. URL: `http://links.jstor.org/sici?sici=0004-5608%28198703%2977%3A1%3C118%3AMSEOTC%3E2.0.CO%3B2-V`.

11 Carlo Rovelli. Relational quantum mechanics. *Int. J. Theor. Phys.*, 35:1637–1678, 1996. `doi:10.1007/BF02302261`.

12 David H. Sanford. Determinates vs. Determinables. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, spring 2019 edition, 2019.

13 C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. `doi:10.1002/j.1538-7305.1948.tb01338.x`.

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

15 D. F. Sinton. The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. In G. Dutton, editor, *Harvard Papers on Geographic Information Systems*, pages 1–17. Addison Wesley, 1978.

16 B. Smith. Fiat Objects. *Topoi*, 20(2):131–48, 2001.

17 B. Smith and A.C. Varzi. Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research*, 60(2):401–420, 2000.

18 John Archibald Wheeler. Information, Physics, Quantum: The Search for Links. In *Proceedings III International Symposium on Foundations of Quantum Mechanics*, pages 354–358. Physics Department, Princeton University, 1989.

The Language of Architectural Diagrams

Thora Tenbrink[1] (ORCID)
School of Languages, Literatures and Linguistics, Bangor University, Wales, UK
t.tenbrink@bangor.ac.uk

Ruth C. Dalton[1] (ORCID)
The University of Northumbria at Newcastle, UK
ruth.dalton@northumbria.ac.uk

Anwen Jago Williams
School of Languages, Literatures and Linguistics, Bangor University, Wales, UK
elu37c@bangor.ac.uk

─── **Abstract** ───

Complex buildings frequently present a challenge to users' understanding, which may affect wayfinding as well as appreciation of the building's structure. In this paper we focus on the building's *diagram*, a representation by the building's architect that captures its main 'idea'. Motivated by the intuition that a building may be easier to understand if its conceptual diagram can be clearly and easily described, we explored perceivers' descriptions of such diagrams' features. We asked students of Language and students of Architecture to write about the buildings represented in a variety of diagrams, and then repeated the task for photographs of the actual buildings. Using Cognitive Discourse Analysis, we aimed to create a first qualitative exploration of the linguistic and conceptual patterns that are associated with the perception of diagrams and images of complex buildings. Among other factors, results show how perception of the diagram's meaning is fundamentally affected by subject expertise. Linguistic patterns demonstrate the ways in which written descriptions reflect observers' understanding and concepts of building representations, providing a starting point for future studies which may address the possible relationship between the *verbalisability* of a diagram and the *legibility* of a building.

1 Introduction

Have you ever stood in front of a complex public building, marvelled at its strange and fascinating forms – and wondered how to make sense of it, locate the entrance or how to find your way around inside? The building shown in Figure 1, Museu Paula Rego in Cascais, might trigger such thoughts – impressive and perhaps a bit intimidating. How easy is it to understand such buildings?

To represent our understanding of a complex building, how would we describe it in language? Words are, after all, our most commonly used tool to represent the world and our understanding of it. We use language to communicate our thoughts to others, and to express and develop our thought processes [5]. To some extent, people's concepts of a building can therefore be accessed by a close look at how they talk (or write) about them,

---

[1] The first two authors contributed equally to this paper.

i.e. by analysing verbal descriptions. In the context of architectural concepts, we might expect that the complexity of buildings and the concomitant understanding of them should be represented in linguistic description. Ultimately, we would also expect that a building that can be clearly understood (and hence clearly verbalised) should also be easy to use, or *usable* – applying a previous definition by Krukar et al. in which they suggest that "A building is usable when it allows the user to execute his/ her tasks effectively, efficiently and with satisfaction in the specified context of use." [14]. To a high extent, this will in practice depend on *navigability*, i.e. the ability to navigate to a destination within the building. This presupposes a degree of understanding of the building's structure, and thus relates to Lynch's idea of *legibility*: namely, "the ease with which its parts can be recognized and organized into a coherent pattern" [15].

Unique buildings such as the one shown in Figure 1 are designed by architects on the basis of an initial idea or concept, frequently represented in an architectural diagram (a notion we'll examine in some depth in Section 2), such as the one shown on the right. A diagram represents the architects', not the users', conceptualization – and it does so in visual form, not in language. Intuitively, there should be a connection: If a diagram and its associated building are easy to understand, they should also be relatively easy to put into words. However, the literature so far offers few insights as to how buildings, or their diagrams, are described intuitively by speakers with different degrees of architectural expertise. In this paper, we therefore start by exploring the ways in which various types of diagrams are verbalised by students of architecture and (for comparison) of language-related subject areas.

2 Architectural diagrams

A diagram can be thought of as a particularly specialist sub-class or type of drawing, in which a number of simplified or often symbolic depictions of real world objects are used to represent complex relations between those objects. The use of diagrams, who produces them, for what reasons, the features that they contain, and how they are used to communicate ideas have been the focus of study in various ways: for example the classic studies on the nature

Table 1 Characteristics of architectural diagrams and their production ("diagramming").

Who?	Primarily architects–sometimes working alone, sometimes working in a team. Less used by non-architect members of the team although they may still be active participants in the process
What?	Feature set consisting of (from Do, 2001): 1. configurations 2. connections 3. shapes 4. orientations 5. physical form/s
Where?	Typically, but not exclusively, produced in the architecture studio, office or atelier (yet includes the 'back of napkin' sketch)
When?	During the design process, particularly at an early, more explorative stage
Why?	To explore and resolve ideas as well as to produce insights and inferences (Do, 2001). To negotiate between 'determinacy and indeterminacy' (Buchanan, 1992)
How?	Typically a freehand sketch on paper (or paper equivalent) using pen/pencil/charcoal etc.

of representation by Peirce [17] and later Goodman [7] or on the use of diagrams in specific contexts such as in scientific texts [8]. In this paper we focus on the very particular type of diagrams, as used by architects. Architectural diagrams are typically produced during the design process [16]; they may be the "key form of visual thinking within architecture" [2].

In Table 1 we outline a range of characteristics of architectural diagrams and their production, which set them apart from the diagrams produced by other disciplines. Do [4] argued that the main distinctive feature of architectural design diagrams, as compared to diagrams in other domains, is "that the elements and spatial relations correspond to physical elements and spatial relations in the architectural problem". While diagrams often depict real world objects, it is only in architectural diagrams that the spatial relationships between those real world objects are elevated to a level that gives them equal to, if not greater importance than, the real world objects being depicted. This is because architecture is essentially a spatial (and specifically a spatial configurational) art: as Hillier says, "The designer is in effect a configurational thinker" [10].

Herbert [9] defined the architectural diagram as an analytic statement used to help the architect solve a problem. Design problems, in general, and architectural design problems specifically, are well-known examples of *wicked* or ill-formulated problems [18, 1]: problems that have no definitive formulation, no stopping rules, can always have more than one solution, and are unique in each case. This is aggravated by the constant need to negotiate between determinacy and indeterminacy [1] - a process that may be supported substantially by the use of diagrams. Somol [19] even suggested that, beyond a means of thinking, diagrams may actually be "the matter of architecture itself"; and furthermore that the architectural diagram has "seemingly emerged as the final tool. . . for architectural production and discourse".

This view of architectural diagrams acting as *more than* a mere representation aligns very well with Hillier's view in his book, Space is the Machine, where he suggests "the idea of architecture is at once a thing and an activity, certain attributes of buildings and a certain way of arriving at them. Product and process are not, it seems, independent. In judging architecture we note both the attributes of the thing and the intellectual process by which the thing is arrived at." [10] If, in this sense, architecture is both a thing and an activity, then surely, by extension, the architectural diagram can be both an activity (tellingly often

denoted by the verb 'diagramming' in architectural practice) and a thing, not only in of itself, but also as an interchangeable artifact standing for the, as yet, unrealised building. Even though some diagrams may never be realised as a building (which may make their significance debatable to some), they still represent their designer's intent.

3   From diagrams to language

If architectural diagrams are not merely a means for thinking about architectural design but have the potential to become the matter of architecture itself (c.f. Hillier's combining of product and process), the question arises whether there is a direct relationship between the qualities of an architectural diagram and the qualities of the resultant building. Does a 'clear' (however defined) diagram produce a more 'legible' building in Lynch's sense [15]? Does a diagram that is easy to understand result in a building that is also somehow clearer and hence more *usable* by the building's inhabitant? Could there be a translation from a building's diagram to its use that can be identified? To what extent would this depend on the observers' expertise? Do architects understand a building's diagram differently, or better, than non-trained observers? How does this relate to perceptions of the real building?

How could we begin to assess the comprehensibility or *clarity* of a diagram?[2] Hölscher and Dalton [11] asked architects and non-architects to gauge the complexity and perceived navigability of a set of buildings based on schematic floor plans. One interesting result was that building layouts that resembled commonly named-shapes (in this case a cross-shaped and a square-shaped layout) were judged very differently from the rest of the sample. These were prototypical examples of shapes with high "Prägnanz" (conciseness) as described in the literature on Gestalt psychology [13, 12]. Apart from representing highly familiar shapes for laypeople and architects alike, these layouts were also easily *describable*, since there existed common words to describe them. Thus, one measure of the *clarity* of a diagram might be how easily it could be described or 'put into words': how speakers *describe* diagrams may reflect what they *understand* about them. In this light, verbal descriptions of architectural diagrams might be key to the understanding of building complexity and, ultimately, usability.

So far, little is known about how diagrams are verbalised, and even less in the architectural domain. It is perhaps fair to say that the most relevant insights about the relation between diagrams and verbal description can be found in Barbara Tversky's work [22, 23, 24]. Tversky consistently takes verbal descriptions as a representation of thought, and finds that linguistic expression and other representation media, such as sketches and gestures, correspond to each other systematically in terms of structure and essential elements or features representing crucial aspects of conceptualisation. However, clearly there are also limits to the kinds of aspects that can or will be verbalised with respect to a diagram or any other pictorial representation. Linguistic representations generally focus on relevance [20] in a discourse context, rather than aiming to be fully exhaustive.

We will now present our study, which addresses the verbalisation of architectural diagrams directly, by investigating linguistic patterns in descriptions of diverse diagrams and building photographs, written by students with varying degrees of relevant subject knowledge.

---

[2] It is important not to confuse what we mean by clarity with Buchanan's assertion that (architectural) problem solving is about relationship between 'determinacy and indeterminacy' [1]. It could be argued that many 'clear' architectural diagrams can simultaneously exhibit both aspects of determinacy and indeterminacy but that this remains quite independent from the clarity of a diagram, since, in our view, clarity reflects the explicitness of the architectural intent.

3.1 Diagram selection and questionnaire design

In order to investigate the verbalisability of a diagram, we selected a consistent set of architectural diagrams, representative of a wide range of styles and from a diverse group of practicing architects, from a recent book on architectural sketches and diagrams[3] by Chris van Uffelen (2014) [25], as follows. We first identified a set of 37 diagrams that corresponded to our notions of an architectural diagram (namely exploratory, 'early-stage' diagrams that seemed to capture aspects of both determinacy and indeterminacy) but did not contain any words. We analysed this initial set in terms of their attributes, noting if they appeared to be drawn as a two-dimensional plan, section or elevation or as a three-dimensional view (or if the viewpoint was unclear). We then recorded the number of occurrences of current drawing elements, i.e. arrows, triangles, squares, rectangles, circles, ovals, spirals, curves/waves as well as 90° and non-90° angles. We considered whether the diagrams included graphic techniques such as hatching and shading and whether they included non-building symbols such as people/figures and foliage/trees. Finally we noted if and when the building corresponding to a diagram had been built in the real world.

From those buildings that had been constructed, we selected two that had a diagram drawn in plan view, two with a diagram representing a section (or elevation) view, and two with a diagram drawn in 3D. For each of these different iconic viewpoints we selected one relatively simple diagram (i.e., the frequency of graphical elements in the feature set was low, compared to the sample as a whole) and one more complex diagram (a relatively high number of graphical elements in the feature set). The final set therefore consisted of 1 x simple+plan; 1 x complex+plan; 1 x simple+section; 1 x complex+section; 1 x simple+3D and 1 x complex+3D yielding 6 diagrams in total (see Figure 3 for all diagrams used in this study[4], and Figure 2 for photographs of the actual buildings).

A questionnaire (approved, separately, by Northumbria University's Research Ethics Committee and by the College of Arts, Humanities, and Business Research Ethics Committee of Bangor University) was designed as follows. Prior to the main data collection, the questionnaire's purpose was explained and participants were asked to give their informed consent. Following the main data collection, anonymized demographic information was collected along with a set of questions designed to identify 'visual thinkers'.

Section 1 of the questionnaire presented participants with each of the six buildings' diagrams, in a sequential but randomized order, along with the instruction (repeated six times): "Please look at this image below. Describe the building as it is depicted in the image, in about three sentences." In Section 2, the same diagrams were shown again (re-randomized), along with the following instruction (again repeated for each diagram): "Please look at this image below. Identify and describe which 'elements' (i.e. lines, shapes, forms, patterns etc.) you can find in this image". In Section 3, a photograph, randomly ordered, of each of the real buildings for which the diagrams had originally been drawn was presented,[5] along with

---

[3] *Sketch* is a more general term; in architecture it typically means a freehand representation of what is seen, or what might be seen: i.e., a translation from vision to paper. The architectural *diagram* is more specific: this is about relations between building spaces, forms and functions, and about exploring these via the medium of the drawing.

[4] Copyrights for Figure 3: Top left: © Christian de Portzamparc Architect; Top middle: Fernando Romero, Mexico City, 2006; Top right: UNStudio Architects; Bottom left: copyright Daniel Libeskind; Bottom middle: Ana Rocha Architecture; Bottom right: Mr. Eduardo Souto Moura

[5] For copyright reasons, the pictures shown in Figure 2 differ slightly from the ones used in the study. Copyrights for Figure 2: Top left: Diego Baravelli; Top middle: Carlos Valenzuela; Top right: UNStudio Architects; Bottom left: Bernd Gross; Bottom middle: Ana Rocha Architecture; Bottom right: Chia Hsien Liao - LeonL. Top left and middle, bottom left and right are licensed under Creative Commons.

Figure 2 Building photographs. Top left: Cidade das Artes, Brasil: 2013 by architect Christian de Portzamparc; top middle: The Soumaya Museum, Mexico City: 2011 by Fernando Romero; top right: House Bierings, Utrecht: 2009 by Ana Rocha/Christian Richters; bottom left: Military History Museum, Dresden: 2011 by Studio Daniel Libeskind; bottom middle: Villa NM, New York: 2007 by Ben van Berkel; bottom right: Museu Paula Rego, Cascais: 2008 by Eduardo Souto Moura.

the following instruction (again repeated 6 times): "Here is one of the buildings that was shown as a drawing earlier. Please describe the real-world building in a few sentences." Thus, for each of the six buildings, three written descriptions were elicited: two for its diagram and one for its photograph.

3.2 Participant features

Two sets of students were invited to participate in the study: students of language-related subject areas (such as Linguistics or English Literature, henceforth "Language student" for short) at Bangor University in Wales, and students of Architecture at Northumbria University in England. These two subject areas were chosen because the study addresses the *language* used in the context of *architectural design*, produced by participants whose background is relevant in distinct ways. The questionnaire was available for two weeks. The only incentive was a prize draw for an Amazon voucher; no other payments were made.

Of the 37 respondents, 22 were female, 14 male and 1 preferred not to say. For consistency of analysis, we eliminated one age outlier (64 years) from the final data set, as well as 5 participants who were not native speakers of English, and 1 participant who failed to complete the questionnaire as asked. The final data set has 12 female, 5 male, and one gender-unidentified language students (mean age: 22.1; age range: 18-33), and 8 female and 4 male architecture students (mean age: 22.7; age range: 20-25).

6 female and 2 male architecture students and 7 female, 2 male, and 1 gender-unidentified language student self-identified as visual thinkers. 7 female and 1 gender-unidentified language students (but no male) considered themselves to be artistic, and 7 female and 3 male architecture students did so. Thus, while the data sets seemed fairly balanced in these respects, architecture students were (as might be expected) somewhat more likely to view themselves as visual thinkers and artistic. The same subject-related tendency was also reflected in the fact that all architecture students said they drew at least once a week or every day, whereas only four of the language students (3 female, 1 gender-unidentified) did so; 6 (3 males) said they could not remember when they last drew something, and 8 (2 males) drew once a month.
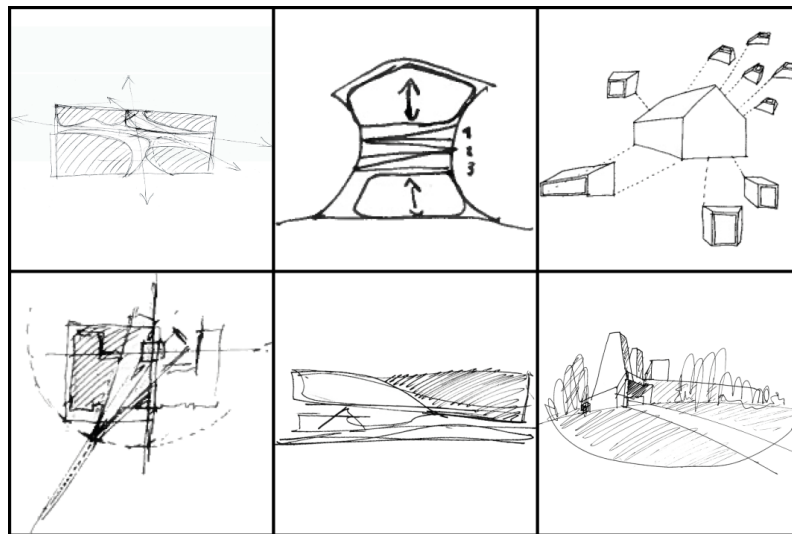
■ **Figure 3** Architectural diagrams used in the study, shown in the same order and configuration as the photographs of the corresponding buildings in Figure 2. Top row: simple; bottom row: complex. Diagrams on the left: plan views; middle: section views; right: 3D views.
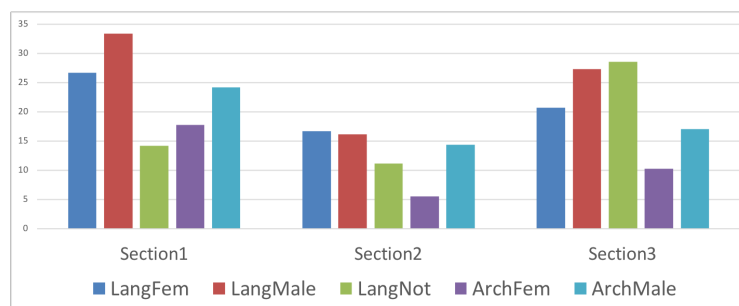
■ **Figure 4** Mean word count in each questionnaire section. LangNot = gender-unidentified.

Subject-related and gender differences were also reflected in word count (see Figure 4).[6] Language students produced more words on average (female: 384.6; male: 460.4) than architecture students (female: 201.6; male: 322); variability of word count was reduced in female architecture students (range: 119-261) as compared to other groups (male architecture students: 151–748; female language students: 161-643; male language students: 137-805).

4 Cognitive Discourse Analysis

Our aim was to gain insights into the concepts represented in the language in an explorative way, in light of various aspects of the study design: the two groups of students with their different subject expertise, how students described diagrams as opposed to photographs, the different building views and varying diagram complexity, and possible differences according to gender. None of these factors can be ignored, but they are too diverse to aim for any

---

[6] Words were counted semi-automatically in Microsoft Excel. Likewise, further analyses were supported by Excel's features as far as feasible.

specific hypotheses based on inferential statistics, especially with unconstrained language production in an only loosely controlled data elicitation exercise, using a small sample size. Instead, we will present a qualitative analysis of linguistic patterns in our data following the methodology of Cognitive Discourse Analysis [21], which aims to identify conceptual aspects in discourse on the basis of systematic linguistic choices (beyond content).[7]

In the absence of specific hypotheses based on earlier literature we inspected the data to gain first insights about patterns in the responses that highlight the participants' underlying concepts systematically, in light of our motivating research question as to how architectural diagrams may be put into words and how verbalisability might relate to complexity and other features of the diagrams. To achieve systematic and objective annotation of our fairly large data set across various aspects, we then semi-automatically counted the occurrence of keywords in the following categories (emerging from the data, rather than predefined).

- **Peculiar**: Words indicating a sense of peculiarity (signalling challenges for legibility), namely *difficult, strange, unusual, unclear, peculiar, odd*
- **Possibility**: Indicators of tentative (i.e., possible rather than certain) interpretations: *seem, could, perhaps, appear, maybe, possible/possibility, ?*
- **Familiarity**: Words marking the respondent's familiarity with something (signalling legibility): *standard, traditional, exactly what, normal*
- **Structure**: Descriptions of structural elements (highlighting what kinds of structures were legible and verbalisable): *rectangle/rectangular, prism, triangle/triangular, square, boomerang, pyramid, box, hourglass, L-shaped, circle*
- **Function**: Mention of possible building functions (suggesting comprehension of the building in this respect): *public, school, theatre, museum, office*
- **Building/house**: Using the words *building* and *house* (reflecting conceptualisations of the building as a whole 3D entity)
- **Building parts**: Mention of any building parts in the house, such as *wall* or *window* (reflecting more fine-grained real-world concepts of the actual building functions)
- **Aesthetics**: Using terms that, in this context, indicate a sense of aesthetics (signalling usability and legibility in a different way): *organic, contrast, clean, feature*
- **Architectural**: Terms we identified in this context as architectural jargon, namely *circulation, intervention, extrude/extruding/intrude/intruding/extrusion/intrusion, orient/orientation, void, mass, material, symmetry*, plus Adrian Forty's "key words in architecture" [6]: *character, context, design, flexibility, form, formal, function, history, memory, nature, order, simple, space, structure, transparency, truth, type, user.*

Note that there is no intention for these different categories to be equal in size, nor comparable in any sense to each other. They are, however, mutually exclusive.

5   Results

To gain a first intuitive impression of the language data, consider some exemplary descriptions:

Female language student, Section 1, complex, plan view (bottom left in Figure 3): "This looks like a birds-eye plan of a building. It is comprised of different shapes and would appear to have a triangular shaped building (or something) at the very top."

---

[7] To clarify, this approach does not exclude inferential statistics but only advocates it where feasible, which it is not in this explorative study.
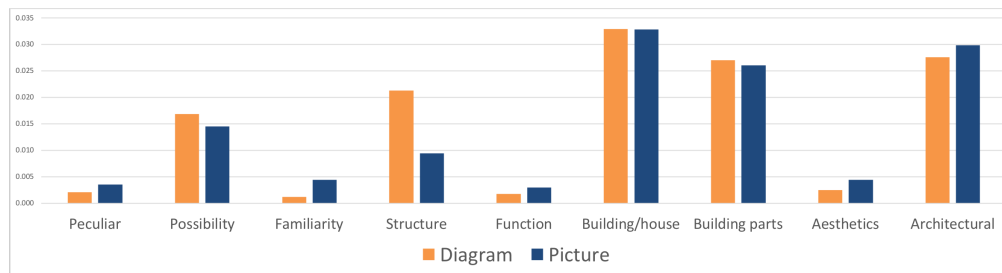
Figure 5 Comparing descriptions of architectural diagrams with pictures.

Male architecture student, Section 1, simple, section view (top middle in Figure 3): "Tall building, with dynamic shape. Two key large spaces with smaller intermediate floors. Long staircase covering entire width."

Male language student, Section 2, simple, 3D (top right in Figure 3): "The second drawing or outline in each dimension makes me think it has elevated surfaces."

Female architecture student, Section 3, complex, section (bottom middle in Figure 2): "Bold design which uses cladding which appears cold. The glazed aspect of the design allows views to a feature of the designs context."

All of these answers are relevant to the question asked; the students are evidently making an effort to describe the features they see and recognize. It is also remarkable that all descriptions, regardless of whether they pertain to a diagram (Section 1 and 2 of the questionnaire) or a photograph (Section 3) take the building seriously in its final design, although some uncertainty can be detected, particularly in the language students' examples. Indeed, intuitively the architecture students' descriptions seem different in some way; this is an effect of subject knowledge that we aimed to capture in more depth.

We found that female architecture students in Section 2, in particular, refrained from any kind of elaboration or speculation when describing the diagrams' elements, as reflected in a visible drop in the word count shown in Figure 4. Among those, a typical answer could be as short as "staircase" or "geometric cubes". Female language students' answers in Section 2 contrast sharply with this by being far more wordy and descriptive, and by frequently attributing function to form, as in "I see the circle which shows the area the building covers. Inside appears to be a shaded building, which could be the main focus, There appears to be a light sketch on the other side which could be the current building there." Again, the description seems tentative, trying to make sense of the diagram's features related to a *possible* building; in contrast, a female architecture student's crisp "Curved form floating above a void", like the other examples, seems to already describe *the building itself*.

Such observations motivated us to identify how linguistic indicators are spread throughout the data more systematically. We approached this by identifying patterns according to the various distinctions introduced by our design. Due to the high variability and diversity of patterns concerning word count reported above, our graphs show results in terms of percentages relative to total number of words in the relevant categories. Textual explanations provide raw numbers to demonstrate how often expressions actually occurred in the data. Patterns are described as they appear, with appropriate caution as to their significance; they may be suggestive but any stronger conclusions would require more controlled studies.

We start by noting that, in line with our initial intuition, descriptions of *diagrams* parallel those of building *photographs* concerning our chosen categories. As visualized in Figure 5, most categories are fairly equally represented in both types of representation. This is
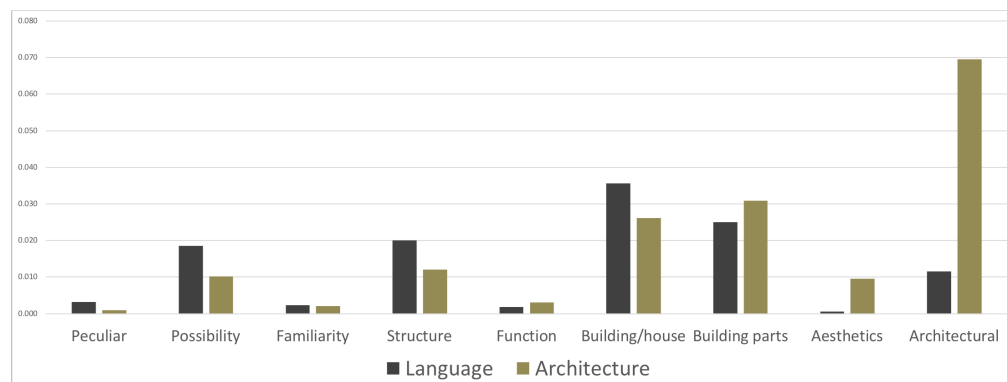
■ **Figure 6** Results according to students' background.

remarkable in light of the fact that participants were not asked to make a direct comparison and never saw a building's photograph side by side with its diagram. The differences that we do find in the graph intuitively make sense. Expressions of familiarity appear to be more frequent with pictures (N=15 out of 3380 words) than with diagrams (N=8 out of 6806 words); most of these (N=14 in total) point to the fairly traditional or standard form of House Bierings, Utrecht. In contrast, descriptions of structure are more frequent with diagrams (N=145; buildings: N=32). Thus, it appears that diagrams represent structures more clearly than the actual buildings do. Functions and aesthetic aspects, in contrast, appear to be somewhat more prominent with photographs.

Concerning our two participant groups, subject knowledge is most clearly evident through the architecture students' enhanced use of architectural terms (*circulation, intervention, mass*, etc.; N=174; language students: N=82). Beyond this, the graph shown in Figure 6 suggests a range of further differences that together account, to some extent, for our previously observed impression that descriptions by architecture students seem different on the whole. In particular, language students tended to use more expressions of peculiarity (*difficult, strange, unusual* etc.; N=23) than architecture students (N=3), as well as expressions of possibility (*seem, perhaps* etc.; N=134; architects: N=30). Also, they referred to structures they recognised within the diagrams using non-architectural terms (*square, pyramid, hourglass* etc.; N=143) more than architecture students (N=34), and used the words *building* and *house* more frequently (N=258; architects: N=70). Relative to the overall number of words written, which was far lower in architecture students (2946 as opposed to 7240 written by language students), architecture students provided suggestions of possible functions (*public, school, museum*, etc.) more often (N=9) than language students (N=13), and they referred more often to aesthetic aspects (*organic, feature* etc.), (N=28; language students: N=4).

Next, we consider possible differences based on diagram complexity (see Figure 7). Against predictions, references to peculiarity seemed more frequent with simple diagrams (N=19) than with complex ones (N=7). However, expressions of possibility were somewhat more frequent in complex diagrams (N=90) than in simple ones (N=74), and expressions of familiarity (*standard, traditional*, etc.) appeared more often in simple diagrams (N=17) than in complex ones (N=6). Suggestions of structure were less frequent with simple diagrams (N=76) than with complex ones (N=101), but the use of architectural terms appeared to decline with complexity (N=141 simple; N=116 complex).

The type of building view in the diagram (see Figure 8) appeared to affect language use only with respect to a few of our categories. References to the words *building* and *house* seemed more frequent in the case of a 3D diagram (N=136 as opposed to N=110 with a plan
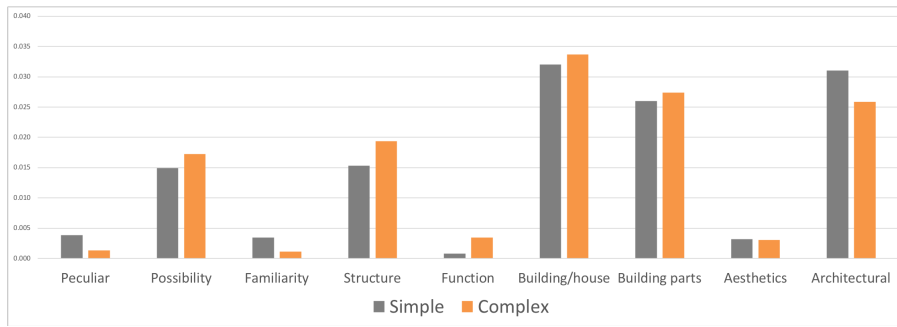
Figure 7 Results according to complexity of the diagram.
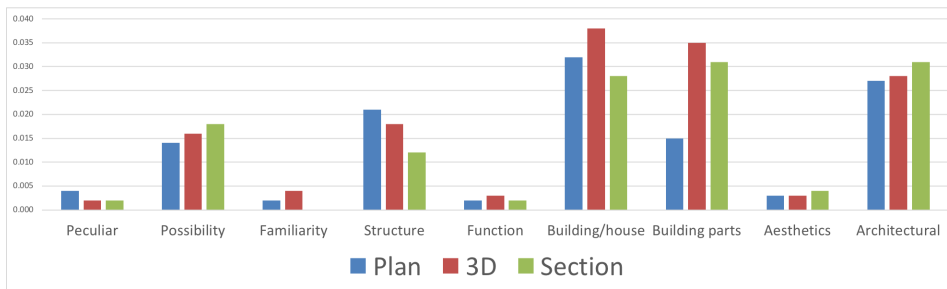
Figure 8 Results according to the type of diagram.

view and N=89 in the case of a section), suggesting that the idea of a building or house was more clearly visible. Structure seemed clearer with plan views (N=74) and 3D (N=66) than with section views (N=37). Plan views did not encourage recognition of specific building parts (N=51; 3D: N=125; section: N=96).

Finally, we can observe some tentative patterns concerning gender (Figure 9), beyond the word count differences noted in Section 3.2. Female participants tended to refer more to structures and used the words *building* and *house* more, and male participants used more expressions of possibility. (The gender-unidentified participant was excluded from this analysis.) We further noted that female architecture students were the only ones who used the words *North, East, South* or *West* in their responses (N=5, all in Section 1); the plan view diagrams may have invited this interpretation in the absence of actual compass information.
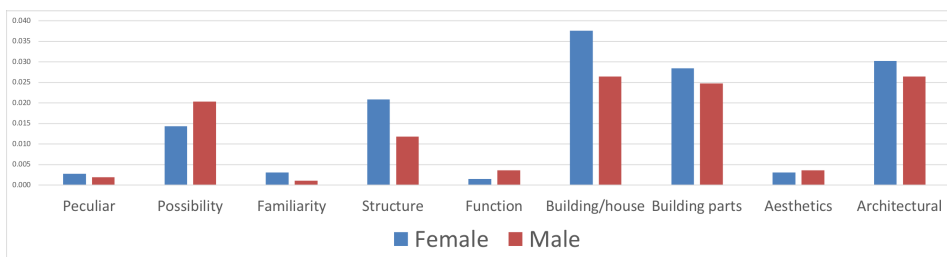
Figure 9 Results according to gender.

6 Discussion

We explored how architectural diagrams and their associated buildings are represented in language by people with different subject expertise, in light of various features of the diagrams. Results highlight a range of relevant observations that we hope will inspire future research. To start with, we note that this appeared a relatively easy task for both sets of students; in fact, the descriptions suggest that it may actually have been quite enjoyable[8] - even though it was an atypical and unusual task for all participants, for different reasons. It is not surprising that language students produced a far higher mean number of words, despite their lack of subject knowledge: architects generally prefer the media of drawing and talking more than the act of writing. The fact that it nevertheless proved quite easy to elicit text descriptions of the diagrams is encouraging, as it suggests that this might be a fruitful way to investigate the purpose, intent and range of possible interpretations of architectural diagrams.

The high degree of correspondence between descriptions of diagrams and building pictures, in terms of our linguistic categories, aligns with the idea that in architecture, diagrams have a particular significance in that they are as much the subject of architectural endeavour as the buildings themselves (see Section 2). Indeed, the designer does not actually 'make' the building – that is built by others – and so the only artifact that is touched by the hand of the 'creator' are the drawings (which includes diagrams). It is therefore reasonable that the drawings should *stand for* the creative intent of the architect and should have a status that is equivalent (in this respect) to the building. To what extent descriptions of photographs correspond to descriptions of actual building views and, ultimately, building experience and usage, is a matter of future investigation.

Our diverse analyses suggest various effects caused by the nature of the diagram (such as its level of complexity or whether it shows the building in plan or section view), as well as the observers' gender. There are, for instance, some indications that complexity matters for the degree of understanding of a building and its diagram, as shown by differences such as references to *possibility* or architectural terminology. These avenues could be pursued further in the future, to assess the relevance of any such systematic effects for building legibility.

Beyond these patterns, it is interesting to consider the ways in which subject knowledge affects diagram (and photograph) interpretation. Unsurprisingly, architecture students made heavier use of what we classed as technical terminology. Note however that only a few of the 'architectural' terms in our list (see Section 4 above) are exclusively used in the realm of architecture. Others, such as *structure*, are everyday words whose meanings may change with expertise [6]. Compare one language student's formulation *flat looking rectangular structure* with an architecture student's *simple pitched roof structure with extruded mirrored window boxes*. For language students, the term *structure* serves as a generic label used to describe whatever elements they can find; for architecture students, the term itself is meaningful, specifying, within their context, the nature of architectural design with respect to the elements described. Similar effects have been found in previous studies; for instance, professional background can affect how simple words such as *back* and *end* are used when describing pictures [3]. Ultimately, such differences in language use subtly convey diversity in how speakers *think* about what they perceive in a picture or in a diagram.

The idea that expertise affects how people conceive of diagrams and associated buildings is further corroborated by various other differences we found between our participant groups.

---

[8] To illustrate, here's a quote from the gender-unidentified language student, while describing the diagram of Villa NM in section 2: "An A-shape holding up a piece of pasta."

Altogether, the impression emerges that language students faithfully *describe what they see*, and use elaborate, cautious descriptions to speculate on possible meanings. For architects, in contrast, the simultaneous presence of both 'determinacy and indeterminacy' [1] is predominant; whilst the determinacy of the diagram shows the design-intent of the architect, it is the very indeterminent aspect of the drawing which permits diagrams to be 'read' on many different levels at the same time. Architecture students are being trained to produce such multi-level, multi-interpretable diagrams *for themselves*; this study indicates that this also contributes to the skill of interpreting the work *of others* in this way.

7   Conclusions and Outlook

This study brought together two hitherto fairly disconnected perspectives: the significance of the architectural diagram as a representation of design ideas [19], and the significance of linguistic choices in representing a speaker's perceptions [21]. The aim was to better understand how observers perceive architectural ideas as represented in diagrams.

Although explorative and qualitative in nature, a range of insights can be gained from this study. First, describing diagrams is feasible and yields meaningful linguistic data. Second, subject expertise (along with various further factors) affects descriptions in various ways; this highlights the different conceptualizations triggered by the visual information. It appears that the clarity of a diagram, or the ways in which it is understood, depends on who is interpreting it. It remains to be seen how these systematic differences in interpretation transfer to the real world building.

Supporting the idea of such a transfer, our third insight is that diagrams and photographs of buildings appear equally interpreted as representations of something real. The next step, accordingly, is to connect these representations to the actual buildings. Our ultimate aim (motivating this initial study) is to see whether a clear diagram (however defined) makes a more usable building (however defined), due to the legibility of its structures. This creates a clear need to take this work to the next stage: to relate diagram descriptions to navigation performance and further measures of the usability of the associated building.

References

1   Richard Buchanan. Wicked problems in design thinking. *Design issues*, 8(2):5–21, 1992.

2   Anthony Burke. Considering the diagram and design research. *Revista Lusófona de Arquitectura e Educação*, 11:345–355, 2014.

3   Claudia Cialone, Thora Tenbrink, and Hugo J Spiers. Sculptors, architects, and painters conceive of depicted spaces differently. *Cognitive Science*, 42(2):524–553, 2018.

4   Ellen Yi-Luen Do and Mark D Gross. Thinking with diagrams in architectural design. *Artificial Intelligence Review*, 15(1-2):135–149, 2001.

5   K Anders Ericsson and Herbert A Simon. *Protocol analysis: verbal reports as data.* MIT Press Cambridge, Mass, 1984.

6   Adrian Forty. *Words and buildings: A vocabulary of modern architecture.* Thames & Hudson London, 2000.

7   Nelson Goodman. *Languages of art: An approach to a theory of symbols.* Hackett, 1968.

8   Mary Hegarty, Patricia A Carpenter, and Marcel Adam Just. *Diagrams in the comprehension of scientific texts*, pages 641–668. Lawrence Erlbaum, Hillsdale, NJ, US, 1991.

9   Daniel M Herbert. *Architectural study drawings.* John Wiley & Sons, 1993.

10 Bill Hillier. *Space is the machine: a configurational theory of architecture.* Space Syntax, 2007.

11   Christoph Hölscher and Ruth Conroy Dalton. Comprehension of layout complexity: Effects of architectural expertise and mode of presentation. In *Design Computing and Cognition'08*, pages 159–178. Springer, 2008.

12   Kurt Koffka. *Principles of Gestalt Psychology*. Harcourt Brace Javanovich, 1935.

13   Wolfgang Köhler. *Gestalt Psychology*. New York: Liveright, 1929.

14   Jakub Krukar, Ruth Conroy Dalton, and Christoph Hölscher. Applying HCI Methods and Concepts to Architectural Design. In *Architecture and Interaction*, pages 17–35. Springer, 2016.

15   Kevin Lynch. *The Image of the City*. Harvard University Press, 1960.

16   Alexandre Menezes and Bryan Lawson. How designers perceive sketches. *Design Studies*, 27(5):571–585, 2006.

17   Charles S Peirce. Logic as semiotic: The theory of signs. *Philosophical writings of Peirce*, page 100, 1902.

18   Horst WJ Rittel and Melvin M Webber. Wicked problems. *Man-made Futures*, 26(1):272–280, 1974.

19   Robert E Somol. Dummy text, or the diagrammatic basis of contemporary architecture. *Diagram Diaries*, pages 6–25, 1999.

20   Dan Sperber and Deirdre Wilson. *Relevance: Communication and cognition*. Harvard University Press Cambridge, MA, 1986.

21   Thora Tenbrink. Cognitive discourse analysis: Accessing cognitive representations and processes through language data. *Language and Cognition*, 7(1):98–137, 2015.

22   Barbara Tversky. Visualizing Thought. *TopiCS*, 3(3):499–535, 2011. `doi:10.1111/j.1756-8765.2010.01113.x`.

23   Barbara Tversky, Julie Heiser, Paul Lee, and Marie-Paule Daniel. Explanations in gesture, diagram, and word. In Kenny Coventry, Thora Tenbrink, and John Bateman, editors, *Spatial language and dialogue*, pages 119–131. Oxford University Press, 2009.

24   Barbara Tversky, Masaki Suwa, Maneesh Agrawala, et al. Sketches for design and design of sketches. In *Human behaviour in design*, pages 79–86. Springer, 2003.

25   Chris Van Uffelen. *The book of drawings + sketches: architecture*. Salenstein: Braun, 2014.

Enabling the Discovery of Thematically Related Research Objects with Systematic Spatializations

Sara Lafia 🆔
Department of Geography, University of California, Santa Barbara, USA
slafia@ucsb.edu

Christina Last
School of Geographical Sciences, University of Bristol, UK
cl15540@bristol.ac.uk

Werner Kuhn 🆔
Department of Geography, University of California, Santa Barbara, USA
werner@ucsb.edu

—— Abstract ——

It is challenging for scholars to discover thematically related research in a multidisciplinary setting, such as that of a university library. In this work, we use spatialization techniques to convey the relatedness of research themes without requiring scholars to have specific knowledge of disciplinary search terminology. We approach this task conceptually by revisiting existing spatialization techniques and reframing them in terms of core concepts of spatial information, highlighting their different capacities. To apply our design, we spatialize masters and doctoral theses (two kinds of research objects available through a university library repository) using topic modeling to assign a relatively small number of research topics to the objects. We discuss and implement two distinct spaces for exploration: a field view of research *topics* and a network view of research *objects*. We find that each space enables distinct visual perceptions and questions about the relatedness of research themes. A field view enables questions about the distribution of research objects in the topic space, while a network view enables questions about connections between research objects or about their centrality. Our work contributes to spatialization theory a systematic choice of spaces informed by core concepts of spatial information. Its application to the design of library discovery tools offers two distinct and intuitive ways to gain insights into the thematic relatedness of research objects, regardless of the disciplinary terms used to describe them.

1 Introduction

In recent decades, the curation of scholarship and its access mechanisms have shifted from physical to virtual spaces. In the 1990s, physical card catalogs were migrated to online databases, trading collocation for scalability [4]. Similarly, library shelves with thematically collocated material are today largely accessed through virtual spaces, such as digital repositories organized by faceted categories [15]. This shift has increased the potential for exchange of scholarly information on the Web through semantically rich *research*

objects [5]. While online library services may provide scholars with access to millions of research objects, they do not necessarily improve the ability of scholars to serendipitously discover related objects. Such a capacity was naturally built into the physical spaces of book shelves, albeit in a limited form. Spatialization can recreate specially designed two-dimensional thematic spaces, such as neighborhoods and networks of themes. These spaces support exploration, browsing, and navigating and can be exploited in future search and discovery services, complementing standard known-item searches [10].

Exploratory search is already supported by library services, like GeoBlacklight[1] and DASH[2], which index research objects *geographically* and enable discovery and access through map interfaces. Such services curate and expose research objects based on their geographic footprints, derived from the named places that they are about (if any). They enable the integration of research perspectives by geographic locations, revealing spatial patterns, such as clusters or gaps [22]. They are especially useful in a university setting where research objects from different disciplines may refer to the same places [20]. However, geographic space only captures geographic notions of location and relatedness. Location, time, and theme are dimensions that can be used to organize observations [28] including research objects. Since in many cases, the temporal organization of research objects is comparatively straightforward (e.g. indexing research objects by their date of publication or the period they are about and displaying them using a time slider), we take on the bigger challenge of representing the relatedness of research themes.

We address this challenge by literally mapping it to the existing solution for discovery by geographic location. In other words, we ask how exploratory search for research objects can be improved by maps of thematic spaces in which related research themes are placed closer together. Conceptually and technically, we adapt our previous work to expose research objects by their geographic footprints [20] to enable discovery in specially designed two-dimensional thematic spaces, which we implement using spatialization techniques. Spatializations exploit people's familiarity with spaces in everyday life to produce intuitive visual information spaces that convey similarity through distance [18]. Spatializations, like self-organizing maps informed by cartographic principles, have been applied to efficiently visualize knowledge domains, such as the subdisciplines of geography [29]. Various types of spatializations, including point maps [24], network maps [12], and regions [11] have been proposed and empirically evaluated, demonstrating that viewers correctly interpret nearby items in abstract space as similar. Analogous to the "first law of geography" [34], this "first law of cognitive geography" states that viewers believe that closer things tend to be more similar [24].

Yet, spatialization remains underexploited, particularly in libraries, which have to deal with vast and context-dependent thematic search spaces. We see this as an opportunity to experiment with spatialization in a multidisciplinary university library repository of research objects. What further distinguishes our approach is that the spatial views we develop are designed based on core concepts of spatial information[3]; in this theory, a base concept (location), a set of content concepts (field, object, network, event), and a set of quality concepts (granularity, accuracy, provenance) capture what spatial information is about. This theory positions spatial information "at a level above data models, independent of particular application domains" [19]. We use these concepts to design two kinds of spatializations: fields of research topics and networks of research objects. A field of research topics reveals their

---

distribution, while a network of research objects reveals their connectivity and centrality. We implement these two spatial views by selecting the spatialization techniques of a self-organizing map [17] and of a planar network. To obtain the necessary visual interfaces for these abstract spaces, we extend the capabilities of the same web GIS platform (ArcGIS Online) that we previously used to display and discover research objects geographically. We show how the spaces that we design are configurable and enable intuitive exploration and discovery of related research objects across disciplines.

The remainder of this paper is organized as follows. In Section 2, we present a motivating scenario to illustrate the challenge of discovering related research. Section 3 explains our conceptual approach to systematize the design of search spaces through the core concepts of spatial information. In Section 4, we implement spatializations of research objects from a university repository of masters and doctoral theses. In Section 5, we apply the spatializations, demonstrating the types of questions that they enable with examples from the previous search scenario. Finally, in Section 6, we envision discovery in spatializations informed by other core concepts of spatial information.

2 Enabling Research Discovery Across Disciplines

Discovering thematically related research in a multidisciplinary setting is both important and challenging. This is a consequence of the siloing of scientific perspectives on the world into different disciplines and the heterogeneous terminologies used within them [33]. Specifically, scholars may find it challenging to identify collaborators and methods outside of their discipline. This is problematic, given that scientific studies and applications of geographic information are increasingly transdisciplinary [19]; they may, for example, combine knowledge from sociology and psychology, or borrow methods from computer science and engineering.

As a motivating scenario, consider two published Geography theses: "Representations of an Urban Neighborhood: Residents' Cognitive Boundaries of Koreatown, Los Angeles" [2]; and "A Temporal Approach to Defining Place Types based on User-Contributed Geosocial Content" [23]. How could the authors of these theses have gone about finding collaborators studying related topics or using related methods? Even for trained interdisciplinary researchers, disciplinary terminologies make it hard to discover related research, resulting in missed sources, insights, and opportunities for collaboration. How can researchers be made aware of thematically related research without needing to know its disciplinary terms?

A common approach to reduce mismatches in keyword-based search is to use ontologies to expand the set of search terms [3]. Such network-based approaches are often based on Linked Open Data and in the case of web journals, enable the discovery of networked data about authors, reviewers, and editors [16]. However, this approach loses the more intuitive similarity relations in the construction of terminological hierarchies [13], whose relations (e.g. broader, narrower) may not always be meaningful to the user. Thus, we propose to complement the terminological approach with an innovative spatial approach affording similarity judgments on research themes. Just as designs for successful everyday spaces, like neighborhoods and street networks, follow spatial patterns [1] and support important cognitive strategies, so can the designs for visual spaces that enable serendipitous discovery. These spatial patterns and strategies are well-understood in the geographic case (consider navigation or perspective-taking) and spatialization carries them over to abstract thematic spaces. The organizational affordances of space, well-known from geographic as well as desktop spaces, can be built into artificial spaces, creating useful and intuitive spatial structures for research themes.

3   Conceptual Approach: Making the Choices of Spaces Systematic

The core concepts of spatial information [19] offer a systematic approach to defining spatial structures by providing a typology of geographic (and other) spaces to guide the organization and interpretation of spatially referenced data. Thus, we recast spatialization as a conceptual choice of a lens through which to view data (i.e. viewing research objects as a field or network). The core concepts of spatial information provide lenses that enable distinct views on spatialized relationships, such as similarity. To go beyond purely cartographic design [22], we make our choices of spaces more systematic by basing our spatializations on those two core concepts that have a solid mathematical formalization: *fields*, formalized by continuous functions from location to theme, and *networks*, formalized by graph theory.

3.1   Choices of Spaces and their Entailments

We first review previous work to create *field* and *network* spatializations, highlighting their underlying spatial theories that inform and evolve our approach. Our thesis is that, if treated systematically and formally, there are distinct choices of spatial concepts that carry perceptual powers; these enable specific types of questions and associated insights.

Landscapes and Fields.   We begin with an example from Wise's [35] pioneering intelligence work, where a spatialized display of news documents shows viewers intuitive similarity relationships based on their proximity in the display. Documents are treated as objects, with *k–means* and *complete linkage hierarchical clustering* used to project documents to a two-dimensional plane. This results in a spatialization, where the position of every news document is surrounded by a neighborhood of topics. A surface is then fit over the display, representing a terrain with peaks of high frequency terms drawn from the corpus.

While this work introduces the metaphor of a landscape or terrain to information visualization, it conflates the field of topic vectors with one of topic frequencies, essentially performing a local map algebra operation. The two field views (topic neighborhoods and topic frequencies) can be separated and an additional view of documents as objects can be added; each affords different types of reasoning (on similarity, frequency, and clustering). In our work, we will show this idea for the case of research objects. While we omit frequencies, which are not supported by adequate amounts of data, we further develop the object view into a network view that shows specific connections between documents.

Another example of an information landscape is Fabrikant's [10] spatialization of a digital library's holdings. Like Wise's approach, multidimensional scaling is used as a projection method to create a surface of keywords. However, Fabrikant's work extends the landscape metaphor by explicitly referencing three spatial concepts: 1) distance (similarity), 2) scale (level of detail), and 3) arrangement (dispersion and concentration), based on Golledge's primitives of spatial knowledge [14]. These concepts are used to systematically inform what users can do in the landscape: looking (overview), navigating (to discover items of interest), changing level of detail, selecting individual documents, and discovering relationships between documents (detail on demand). While this example moves toward conceptual formalization, it does not yet support multiple views based on different spatial concepts.

Networks and Graphs.   "Maps of science" visualize research networks, ranging from co-citation networks to expertise profiles [7]. Börner et al. visualize a network of millions of university research articles embedded in an abstract spherical space. The network is rendered in a pseudo-Mercator projection, based on the idea that a Riemannian perspective, which

uses a sphere as the layout surface, offers continuous linkages. However, it is unclear what additional costs or benefits this choice imparts, as some network properties (like centrality) may be more challenging for viewers to ascertain in such a view.

The extraction of spatial and temporal information from digital text archives can inform more systematic spatializations [8]. Bruggmann and Fabrikant embed a network of toponyms (nodes) and their relationships (edges) in a geographic map to illustrate their connectivity and hierarchy. The inclusion of time in their analysis enables interesting questions about how certain places have risen or fallen in prominence over some period; this is encoded by node size (frequency of mention) along with edges (co-references with another place). The resulting networks are clear and effective, highlighting important relationships, like centrality, through systematic choices of node roles, edge roles, weighting, and embedding.

3.2   Locating Research Objects in Topic Space

Our conceptual design addresses university theses, which do not have any inherent way of locating them. We therefore model them as research objects in an n-dimensional vector space of topics. To locate them, we perform topic modeling on their titles and abstracts. Although the full text is available for most theses, we consider them to be adequately described at the metadata level; our approach gains efficiency and practicality, as only commonly available metadata are required for spatialization. Topic modeling assigns each thesis a vector of keywords (standing in for their topics) locatable in a two-dimensional topic map. We chose to assign topics to research objects, as this supports useful exploratory data analyses [6].

Field-based model.   Rather than using the topic model to compute on the similarities of theses, we spatialize it into a topic map that supports visual pattern detection and similarity inferences. Our first spatialization is based on the field concept, with topics as the field attribute. Fields enable questions about the value of an attribute at any position in a given spatial and temporal domain. Field-based models underlie, but do not imply the use of, a landscape metaphor. They involve explicit choices of a spatio-temporal framework and a type of attribute (scalar, vector, spinor, or tensor).

We create a self-organizing map (SOM) using the vectors of words that result from the topic model. The SOM creates a field with a two-dimensional abstract spatial framework and a vector attribute. It represents topic locations as hexagonal cells into which point objects (representing the theses) fall. This can be seen as an example of a relative Leibniz space, generated based on objects, rather than a pre-established absolute Newtonian space [26]. The SOM satisfies the criteria for field-based models as follows:

- In its **spatio-temporal framework**, time is held constant (covering the entire period of available theses), location is controlled by the topic map, and theme is measured.
- The measured attribute **value** is an n-dimensional topic vector of words associated with the topic, ordered by their probability of occurring in theses on the topic.
- Furthermore, the topic field is **continuous**, in that a small move in position in any of six directions results in a small change in attribute value.

Network-based model.   Our second choice of spatialization is based on the network concept. Networks provide views of objects that are not supported by a field view, such as questions about direct connections between objects and their centrality in the network [19]. Graphs formalize network models and give them inferential power and versatility.

Network models in general require the following explicit choices [25]: what plays the role of a node?, what plays the role of an edge?, how are edges labeled or weighted?, do they

have direction?, and is there an embedding of the nodes, edges, or both in another space? Like the field-based model, the planar network that we produce also exemplifies a relative Leibniz space. Our network spatialization of theses rests on the following choices:

- The theses (research objects) are conceptualized as **nodes**.
- The **edges** are defined based on a **binary** topical relation between theses; if two research objects have at least one of five "top topics" in common, they share an edge.
- The edges are **weighted** by the value of the topic attribute (0–1).
- The edges are **non-directed**, as topic sharing is symmetrical.
- The nodes are **embedded** in a planar space, also based on value of the topic attribute.

4   Technical Approach: Implementing Field and Network Spatializations

We spatialized masters and doctoral theses accessible through the Alexandria Digital Research Library (ADRL), a repository[4] curated by the UC Santa Barbara Library. It is named for the original Alexandria Digital Library (ADL), a project in which users could access multimedia library objects through a map interface [31]. Experimental work on ADL also resulted in a prototype "information landscape" of library objects based on frequent keywords [10]. Despite the lineage that ADRL shares with the original ADL geo-library project, it does not offer any spatial search capabilities, neither in geographic nor in thematic space; this design limitation presents an opportunity to develop spatial views that enable the discovery of research objects. We use established topic mapping and spatialization techniques [30] to:

- harvest the metadata of research theses from the ADRL repository,
- compute and assign topics to the theses using topic modeling, and
- spatialize the topics, producing a self-organizing map (SOM) and a network.

4.1   Metadata Harvesting

For our experiment, we chose research theses published by graduates of UC Santa Barbara between 2011 and 2016 that represent all 53 academic departments granting graduate degrees. The theses are accessible through a public-facing search interface, which provides keyword-based search and facets. The metadata are not accessible through an API, so we obtained permission from the UCSB Library to harvest them for the 1,731 research theses using a combination of *WGET*[5] and the Python libraries *Crummy* and *Beautiful Soup 4*[6]. The metadata follow the Portland Common Data Model[7] and are comprised of: a unique identifier; a title (of 50 words or less); a year of publication; an author; a degree grantor; a degree supervisor; a language; and a detailed abstract (no word limit) containing a problem statement, a description of methods and procedures used to gather data, and a summary of findings. Researcher contributed (uncontrolled) keywords were only available for research theses added after 2017, so we did not include keywords in our topic model.

---

[4] `https://alexandria.ucsb.edu/collections/f3348hkz`
[5] `https://www.gnu.org/software/wget/`
[6] `https://www.crummy.com/software/BeautifulSoup/bs4/doc/`
[7] `https://pcdm.org/2016/04/18/models`

4.2   Topic Modelling

We produced a topic model using *MALLET*[8], an open-source package developed for text-based machine learning applications. We applied Latent Dirichlet Allocation (LDA) to model the topics present in the combined text of the title and abstract of each thesis [6]. LDA is used to determine the thematic relatedness of theses by attributing the presence of each word in the combined title and abstract text to a topic characterized by a word vector. LDA is suitable, as it has been applied to many similar dimensionality reduction and classification problems [6]. LDA largely succeeds in capturing the notion of relatedness (relative to the set of inputs) despite the fact that different terms are used within those inputs (e.g., "variability" and "change" are likely to be grouped into a single topic). Thus, it is a pragmatic solution for dealing with complex notions of topics and their relatedness.

We removed the standard English stop words using a list from the MALLET package. We then experimented with between 30 to 100 topics, roughly corresponding to the number of academic departments at UCSB, which indicates a rather coarse topic granularity and targets the cross-disciplinary scope of our inquiry. We found that 71 topics provided the lowest log-likelihood value, a criterion that optimizes for the tightest possible lower bound [6]. We then assigned topic probabilities to the research objects, coded from 0 to 70. We chose to leave the topics unlabelled; they are characterized only by their word vectors. The assignment of topics provides the basis for relatedness in the following steps.

4.3   Field Spatialization

We adapted a method developed by Bruggmann to spatialize the output of a topic model [8] by using a self-organizing map (SOM) toolbox[9] for ArcGIS 9 written by Lacayo-Emery. This toolbox implements the SOM algorithm [17] in existing cartographic software, leveraging its clustering and dimensionality reduction to produce a 2-dimensional map that is readily visualized. We set the following parameters: the x / y dimension of the SOM was 42 x 42 (1,764 hexagons); the length of training was 50,000 / 500,000 runs; and the initial neighborhood radius was 42 / 6. We used the probability distribution matrix that resulted from topic modeling to produce our SOM template in ArcGIS Desktop. For cartographic readability, we only display theses from the most productive departments (those with over 50 theses). This resulted in a SOM showing 775/1,731 theses from 10 departments. Figure 1 shows the SOM, which is also published to ArcGIS Online as an interactive web application[10].

4.4   Network Spatialization

We applied a hierarchical clustering method adapted from Leicht et al. [21], which is a compromise between the single-linkage clustering method (in which a single edge is defined based on the most related pairs of nodes) and average-linkage clustering (in which an edge is defined based on the average relatedness of all pairs of nodes). We used the *tidyverse*[11] package in R to construct the edge list, assigning theses the role of nodes and shared topics the role of edges; for cartographic readability, we restrict shared topics to 5. Specifically, each thesis is characterized by the same topics and associated word vectors used to produce the SOM. For example, if *Thesis A* is characterized by Topics 2, 11, 22, 34, and 60 and *Thesis B* is characterized by Topics 4, 11, 27, 33, and 51 they share one edge based on shared Topic 11.

---

[8] http://mallet.cs.umass.edu/
[9] http://code.google.com/p/somanalyst
[10] http://arcg.is/0vyezH
[11] https://www.tidyverse.org/
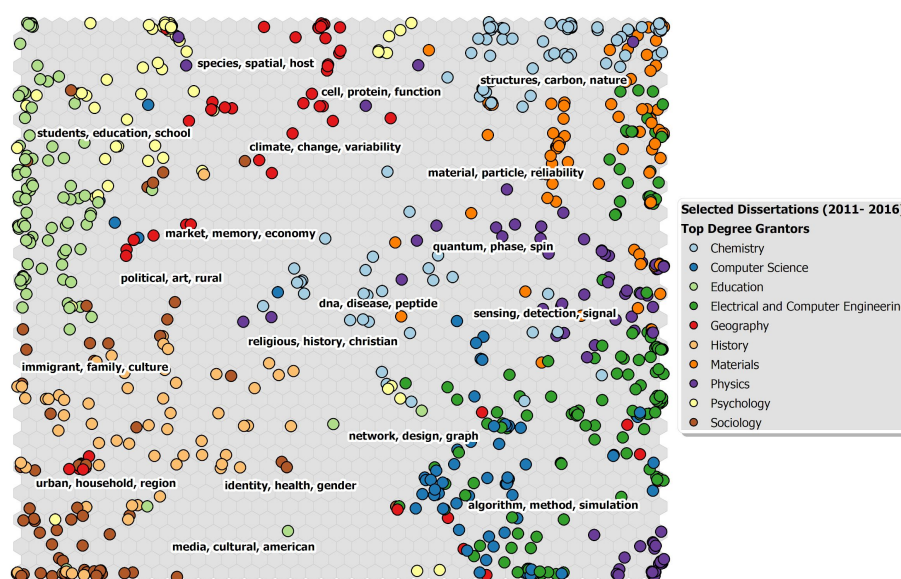
Figure 1 Theses (color coded by academic department) located in the SOM (gray tessellated topic field of themes, shown with representative terms from word vectors).

We scale node size relative to the amount that two nodes share a corresponding topic; thus, a larger node corresponds strongly with its shared topic and a smaller node does not. For example, if *Thesis A* is characterized 70% by Topic 11, its node size will be 0.7 (out of a maximum size of 1). We also embed nodes in a planar space (distinct from that of the SOM) that reflects how strongly each node corresponds to its "top-topic"; the position of each node reflects the value (0–1) of the top topic vector. To enable comparisons between the SOM and the network, we randomly sampled without replacement 775 nodes, embedded in a planar space, and connected them with edges standing in for a "top-five" topic. Figure 2 shows the network constructed with the *networkx*[12] library, which is also published in a reproducible Jupyter Notebook[13] and deployed using Binder[14].

5   Application: Discovering Thematically Related Research

The spatializations that we produce enable scholars to discover thematically related research objects, unlike the current ADRL, which does not offer any such capabilities. We apply the field and network concepts of spatial information to the motivating scenario offered in Section 2, referencing specific research objects related to the theses from the scenario. Patterns of relatedness are interpreted using Golledge's spatial primitives of *distance*, *arrangement*, and *scale* [14], which have informed previous conceptual formalizations [10].

---

[12] https://networkx.github.io/

[13] https://github.com/saralafia/adrl/tree/master/3_network

[14] https://mybinder.org/v2/gh/saralafia/adrl/master
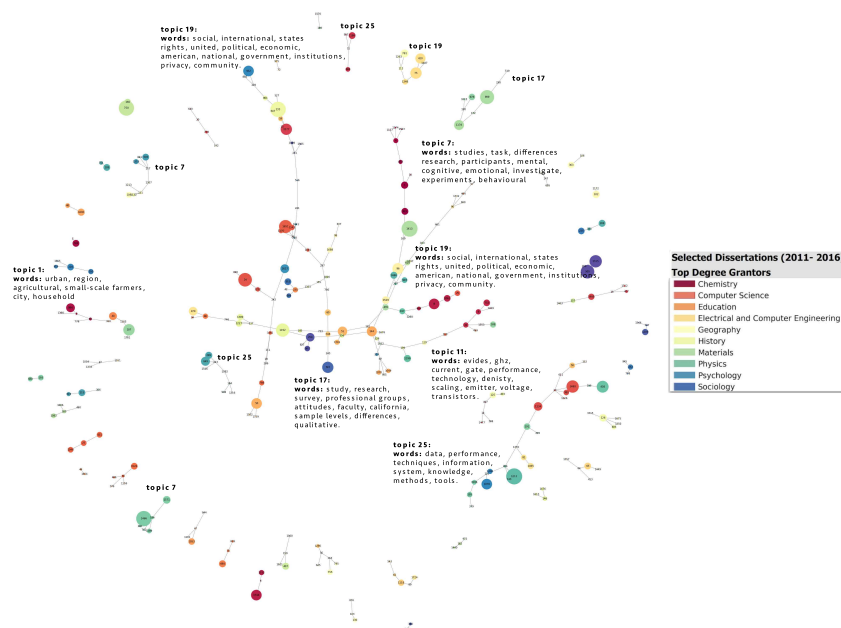
■ **Figure 2** Theses (color coded by academic department) connected by their shared five top topics (shown with representative terms from word vectors).

5.1 Questions Enabled by a Field of Research Topics

Both the field, in the form of a self-organizing map (SOM), and the research objects used to produce it enable the discovery of related research topics. Fields enable questions about value (i.e., research topic) at a given location. A continuous field function satisfies Tobler's First Law of Geography [34], so that nearby topics in the SOM are similar. For pairs of objects, similarity can therefore be assessed by *distance*. Researchers interested in a particular area of research can see related theses by examining those closest to that area of interest in the SOM. Closely related research objects tend to fall within the area's neighborhood (i.e. a single hexagonal topic location or an aggregate of several such cells).

In the case of Bae's research from our scenario, the SOM displays six research objects from Geography, History, and Sociology within a neighborhood. Neighborhoods can be defined based on various distance thresholds. In addition to shared topics, relatedness may also reflect shared methods and techniques, as these are typically captured in abstracts as well. McKenzie's research, for example, is in a neighborhood of research objects from Computer Science and Electrical and Computer Engineering. While the subject matter of some research is different (e.g., photography or drugs), the theses share methods (e.g. "spatial, data, search..." and "learning, place, knowledge..."). Figure 3 illustrates these related research objects from the scenario, located in the SOM.

Beyond similarity of themes or methods, *arrangement*, such as the dispersion or concentration of research themes in a topic space, are also demonstrated in the field view. Theses that address the "urban, region, local..." topic are clustered and centered in the SOM, indicating that this topic pertains to many theses; conversely, topics (and their associated research objects) at the periphery of the SOM are less related to other research topics (e.g. "dna, disease, peptide...") and pertain to fewer theses. Compared with concentrated theses from other departments (like Materials, shown previously in Figure 1) the Geography department theses are dispersed; although Bae and McKenzies' theses share topics ("urban, region, local..." and "models, based, system..."), they are still distant from each other.
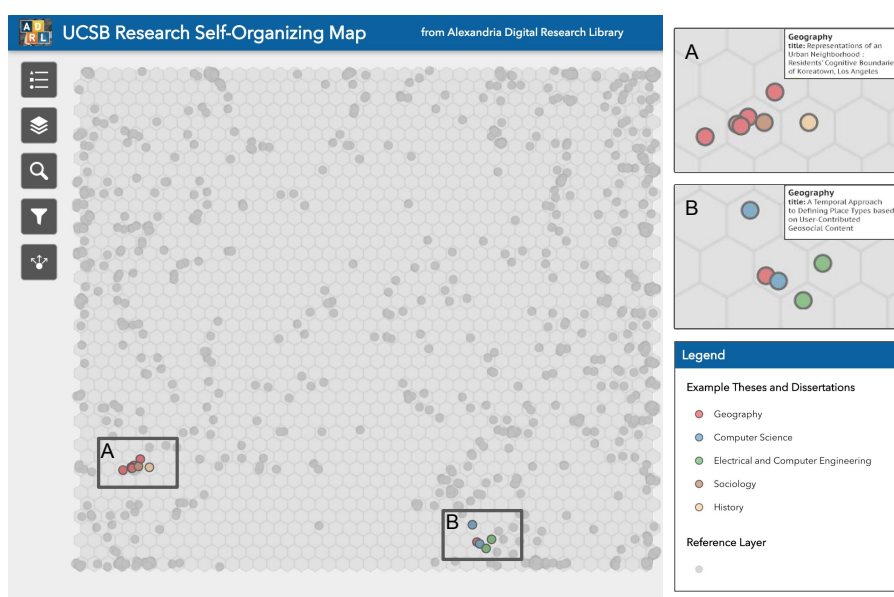
■ **Figure 3** Selected theses (color coded by academic department) located in the SOM and surrounding: (A) Bae's geography thesis; and (B) McKenzie's geography thesis.

The field view with the thesis objects placed in it also reveals the presence and absence of research areas among existing theses. Portions of the field that do not contain any theses show research areas that may not be addressed in the repository, possibly suggesting interesting themes not yet studied and signaling opportunities for research at the boundary between disciplines. It should be noted that such gaps can also result from distortions in distance; cartogram techniques, which we have not yet applied to our field view, can be used to account for this by warping the SOM basemap [9]. Nonetheless, gaps between History and Geography surrounding Bae's research for example might suggest opportunities for integration of subject matter and techniques in this area (e.g., in the spatial humanities).

Scale in the field view is determined by topic modeling (number of input topics) and the parameters of the SOM (spatial resolution of the cells that locate topics). The size of the cells in relation to the whole field, and the dimensions of the field influence the position of topics and research objects. In our SOM, only one other thesis shares a top topic with McKenzie's research; this would likely change if the resolution of the cells changed, resulting in different topic groupings. Prevalent themes of research objects are visible at multiple levels. At the repository level shown in Figure 1, a prevalent topic appears to be about "spatial, visual, search. . ." and relates to research across many departments, including Psychology, Geography, and Computer Science. Prevalent topics of departments can also be seen from the color coding of theses by academic department (rather than by academic advisor or year of publication, which would be other possible choices).

5.2   Questions Enabled by a Network of Research Objects

Questions about the similarity, distribution, and prevalence of research topics in a repository are handled in the SOM view; however, questions about explicitly modeled relationships between the objects are not. Networks deal with these questions by encoding the relationships in their edges: for instance, are the theses of Bae and McKenzie topically related, and if so how? Figure 4 illustrates how networks convey connectivity, showing topical correspondence between departments and topical diversity within departments.
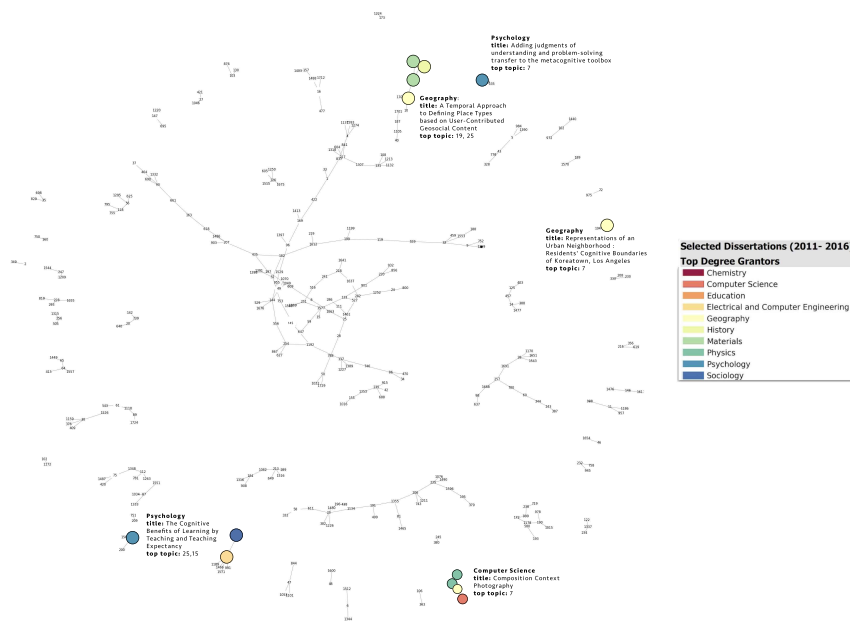
Figure 4 Selected theses (color coded by academic department and labelled by title) connected to other theses if the pair shares any five "top topics".

A network view answers questions about the specific relation encoded by network *distance*. Bae's research is most thematically related to other research objects one edge apart. A comparison between the network and the SOM shows additional similar theses, such as one from Marine Sciences, which is also characterized by the "urban, region, local..." topic.

In comparison to the SOM, where McKenzie's thesis is located next to a Computer Science thesis, there is a larger distance between them. In the network, McKenzie's thesis is closer to History and Materials theses, characterized by Topics 19 ("international, social, political...") and 25 ("data, performance, techniques..."). The Computer Science thesis shares a stronger topical relation with Geography and Physics theses, which are characterised by Topic 7 ("image, multiple, technology...").

Arrangement is related to node embedding; the most central topics in the network visualization are shared by the most research objects. More specifically, topics that intersect the central region of the network are less specific than topics that describe multiple research nodes. Niche topics are pushed toward the edge of the network; thus, theses that are heavily characterized by these topics cohere to them strongly. As shown in Figure 2, research nodes occupying a central location in the network are characterised by the generic terms "study, research, survey..." and by "studies, tasks, differences...". Conversely, theses such as those represented by the specific terms "work, material, particle..." share the fewest edges and therefore, are least central. Such theses deal with technical themes shared only by a few departments (in this case, those of Materials, Chemistry, and Electronic Engineering).

The *scale* of the network view shows a hierarchy with three levels: individual research objects, academic departments, and the repository as a whole. The nodes and the edge relations in a network can be defined in many ways. A node could represent a particular researcher and its attributes could be a list of theses published or supervised by the academic. Instead of representing a shared topic, edges could stand for a shared advisor, creating a network of "academic families or schools". While the choice to restrict edges to five top topics was pragmatic, it also illustrates the flexibility of the design approach; any kind of binary relations between research objects can be visualized.

6   Conclusion

6.1   Summary

In order to enable discovery in a multidisciplinary setting, we develop two systematic spatializations that allow users to identify thematically similar research objects. These spatializations provide a helpful alternative to known-item search by facilitating exploration; they do not require users to have prior disciplinary knowledge. To produce them, we conceptually reframe existing spatialization techniques using core concepts of spatial information. From this reframing, we produce two applications: a self-organizing map of research topics (a *field* view) and a network of connected research objects (a *network* view). In both spatializations, the relatedness of research objects can be ascertained by their *distance*; nearby topics (in neighborhoods) or objects (separated by an edge) are more related. The *arrangement* of topics and objects in each spatialization also indicates their overall relatedness; central research topics or objects tend to be more shared, while those on the periphery are niche. Finally, *scale* in both spatializations is determined during pre-processing (e.g. number of topics in the model) and spatialization (e.g. cell size; node or edge assignment). While made systematic, these choices are parameters that can be reconfigured during subsequent analysis.

6.2   Outlook

Spatializations in library services enable thematic search for research objects and complement our previous implementation of geographic search for them. Spatializing research themes extends the power of spatial search from geographically-referenced information into topic spaces, formalized in this work by core concepts of spatial information: *fields* and *networks*.

Information displays that index research by theme, location, and time [28] enable scholars to ask novel questions. The relatedness of research, indicated by proximity either in geographic location (e.g. Central American archaeology and entomology research) or thematic location (e.g. archaeological excavations of diverse ancient cultures) shows the potential for interplay between thematic and geographic views that our work enables. Furthermore, we envision allowing users to explore the spatializations in combination, gaining distinct yet complementary views of the same repository. While exploring a self-organizing map (SOM), a user can gain an overview of topics in the repository and from this, identify a specific area of interest. The subset of research objects falling into that area of the SOM can then be explored in the network, enabling further interrogation of connections. We are interested in assessing the design of our spatializations using standard usability tests, where test subjects are given questions to answer with each spatialized theme.

Temporal visualization beyond time sliders may also play a role in enabling research discovery. The meaningful representation of time-varying information [32] in disciplines like the digital humanities is notoriously fuzzy, inconsistent, and spatially variable [27]. We envision using temporal information inhering in research theses (e.g. publication date; events or periods studied) to be modeled by *events* and support reasoning on periods (time intervals). Time, made explicit and linked to spatializations, could show how research topics evolve in geographic and thematic spaces. However, events do not yet seem to provide a useful metaphor for spatializations, as they are notoriously difficult to visualize in static maps.

Visualizing the quality (as opposed to the content) core concepts of spatial information, which include *granularity*, *accuracy*, and *provenance* [19], suggests many directions for future spatialization work. *Granularity*, or level of detail, relates both to geographic scale and to the coarsened or refined topics shown in spatializations. At present, granularity

provides a clear and important intuition, as it relates directly to visualization (e.g. detail on demand). *Accuracy* relates to validity, possibly determined through comparison of multiple spatializations against domain ontologies. Finally, *provenance* may provide a way to explore the lineage of ideas (e.g. discovering related research through co-citation networks).

The long-term goals for this work are to increase awareness of relevant previous or ongoing research by applying spatial thinking to the discovery of thematically related work. Integrating research by spatialized topic, rather than siloing it by discipline, is likely to enable increased collaboration across academic disciplines. Much like browsing stacks of books in a physical library, exploring a spatialized library repository can transform a common research task into a learning opportunity or a serendipitous discovery.

References

1   Christopher Alexander. *A pattern language: towns, buildings, construction.* Oxford university Press, 1977.

2   Crystal Ji-Hye Bae. *Representations of an urban neighborhood: residents' cognitive boundaries of Koreatown, Los Angeles.* University of California, Santa Barbara, 2015.

3   Ricardo Baeza-Yates and Berthier de Araújo Neto Ribeiro. *Modern information retrieval.* New York: ACM Press; Harlow, England: Addison-Wesley, 2011.

4   Nicholson Baker. Discards. *The New Yorker*, page 64–86, April 1994.

5   Sean Bechhofer, David De Roure, Matthew Gamble, Carole Goble, and Iain Buchan. Research Objects: Towards Exchange and Reuse of Digital Knowledge. In *The Future of the Web for Collaborative Science (FWCS 2010)*. Nature Precedings, 2010.

6   David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022, 2003.

7   Katy Börner, Richard Klavans, Michael Patek, Angela M Zoss, Joseph R Biberstine, Robert P Light, Vincent Larivière, and Kevin W Boyack. Design and update of a classification system: The UCSD map of science. *PloS one*, 7(7):e39464, 2012.

8   André Bruggmann and Sara I Fabrikant. How to visualize the geography of Swiss history. In *Proceedings of the AGILE'2014 International Conference on Geographic Information Science,*. AGILE Digital Editions, 2014.

9   André Bruggmann, Marco M Salvini, and Sara Fabrikant. Cartograms of self-organizing maps to explore user-generated content. In *26th International Cartographic Conference*, pages 25–30, 2013.

10   Sara I Fabrikant. Spatialized browsing in large data archives. *Transactions in GIS*, 4(1):65–78, 2000.

11   Sara I Fabrikant, Daniel R Montello, and David M Mark. The distance-similarity metaphor in region-display spatializations. *IEEE Computer Graphics and Applications*, 26(4):34–44, 2006.

12   Sara I Fabrikant, Daniel R Montello, Marco Ruocco, and Richard S Middleton. The distance–similarity metaphor in network-display spatializations. *Cartography and Geographic Information Science*, 31(4):237–252, 2004.

13   Peter Gärdenfors. Semantics. In *Conceptual Spaces: The Geometry of Thought*. MIT Press, 2000.

14   Reginald G Golledge. Primitives of spatial knowledge. In *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, pages 29–44. Springer, 1995.

15   Marti Hearst. User interfaces and visualization. *Modern information retrieval*, pages 257–323, 1999.

16   Yingjie Hu, Krzysztof Janowicz, Grant McKenzie, Kunal Sengupta, and Pascal Hitzler. A linked-data-driven and semantically-enabled journal portal for scientometrics. In *International Semantic Web Conference*, pages 114–129. Springer, 2013.

17   Teuvo Kohonen. Learning vector quantization. In *Self-Organizing Maps*, pages 175–189. Springer, 1995.

18   Werner Kuhn. Handling data spatially: Spatializing user interfaces. In *Advances in GIS research II: Proceedings of the 7th International Symposium on Spatial Data Handling*, volume 2, page 13B, 1996.

19   Werner Kuhn. Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information Science*, 26(12):2267–2276, 2012.

20   Sara Lafia, Jon Jablonski, Werner Kuhn, Savannah Cooley, and Antonio Medrano. Spatial discovery and the research library. *Transactions in GIS*, 20(3):399–412, 2016.

21   Elizabeth A Leicht, Gavin Clarkson, Kerby Shedden, and Mark Newman. Large-scale structure of time evolving citation networks. *The European Physical Journal B*, 59(1):75–83, 2007.

22   Alan M MacEachren. *How maps work: representation, visualization, and design.* Guilford Press, 2004.

23   Grant D McKenzie. *A temporal approach to defining place types based on user-contributed geosocial content.* University of California, Santa Barbara, 2015.

24   Daniel R Montello, Sara I Fabrikant, Marco Ruocco, and Richard S Middleton. Testing the first law of cognitive geography on point-display spatializations. In *International Conference on Spatial Information Theory*, pages 316–331. Springer, 2003.

25   Mark Newman. *Networks.* Oxford University Press, 2018.

26   Joan Nunes. Geographic space as a set of concrete geographical entities. In *Cognitive and linguistic aspects of geographic space*, pages 9–33. Springer, 1991.

27   Adam Rabinowitz, Ryan Shaw, Sarah Buchanan, Patrick Golden, and Eric Kansa. Making sense of the ways we make sense of the past: The PeriodO project. *Bulletin of the Institute of Classical Studies*, 59(2):42–55, 2016.

28   David Sinton. The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. *Harvard papers on geographic information systems*, 1978.

29   André Skupin. The world of geography: Visualizing a knowledge domain with cartographic means. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5274–5278, 2004.

30   André Skupin and Sara I Fabrikant. Spatialization Methods: A Cartographic Research Agenda for Non-geographic Information Visualization. *Cartography and Geographic Information Science*, 30(2):99–119, 2003.

31   Terence R Smith and James Frew. Alexandria digital library. *Communications of the ACM*, 38(4):61–62, 1995.

32   Richard T Snodgrass. Temporal databases. In *Theories and methods of spatio-temporal reasoning in geographic space*, pages 22–64. Springer, 1992.

33   Dagobert Soergel. The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science*, 50(12):1119–1120, 1999.

34   Waldo R Tobler. A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1):234–240, 1970.

35   James A Wise. The ecological approach to text visualization. *Journal of the American Society for Information Science*, 50(13):1224–1233, 1999.

The Future of Geographic Information Displays from GIScience, Cartographic, and Cognitive Science Perspectives

Tyler Thrash [ORCID]
Department of Geography / Digital Society
Initiative, University of Zurich, Switzerland
tyler.thrash@geo.uzh.ch

Sara Lanini-Maggi [ORCID]
Department of Geography / Digital Society
Initiative, University of Zurich, Switzerland
sara.maggi@geo.uzh.ch

Sara I. Fabrikant [ORCID]
Department of Geography / Digital Society
Initiative, University of Zurich, Switzerland
sara.fabrikant@geo.uzh.ch

Sven Bertel
Center for Interaction, Visualization, and
Usability, Hochschule Flensburg, Germany
sven.bertel@hs-flensburg.de

Annina Brügger [ORCID]
Department of Geography,
University of Zurich, Switzerland
annina.bruegger@geo.uzh.ch

Sascha Credé [ORCID]
Department of Geography,
University of Zurich, Switzerland
sascha.crede@geo.uzh.ch

Cao Tri Do
Institute for Biomedical Engineering,
University of Zurich / ETH Zurich, Switzerland
cdo@biomed.ee.ethz.ch

Georg Gartner [ORCID]
Department of Geodesy and Geoinformation,
TU Wien, Austria
georg.gartner@tuwien.ac.at

Haosheng Huang [ORCID]
Department of Geography,
University of Zurich, Switzerland
haosheng.huang@geo.uzh.ch

Stefan Münzer
Department of Psychology,
University of Mannheim, Germany
stefan.muenzer@uni-mannheim.de

Kai-Florian Richter [ORCID]
Department of Computing Science,
Umea University, Sweden
kai-florian.richter@umu.se

---- **Abstract** ----

With the development of modern geovisual analytics tools, several researchers have emphasized the importance of understanding users' cognitive, perceptual, and affective tendencies for supporting spatial decisions with geographic information displays (GIDs). However, most recent technological developments have focused on support for navigation in terms of efficiency and effectiveness while neglecting the importance of spatial learning. In the present paper, we will envision the future of GIDs that also support spatial learning in the context of large-scale navigation. Specifically, we will illustrate the manner in which GIDs have been (in the past) and might be (in the future) designed to be context-responsive, personalized, and supportive for active spatial learning from three different perspectives (i.e., GIScience, cartography, and cognitive science). We will also explain why this approach is essential for preventing the technological infantilizing of society (i.e., the reduction of our capacity to make decisions without technological assistance). Although these issues are common to nearly all emerging digital technologies, we argue that these issues become especially relevant in consideration of a person's current and future locations.

1   Introduction

One serious consequence of global urbanization is the additional strain on cities' transportation networks. From the perspectives of engineers, economists, and planners, an apparent solution to this challenge is to allocate more resources towards public (and automated) modes of transportation (e.g., trains) and their optimization in terms of network efficiency [31]. However, people from different cultures vary with respect to their acceptance of public transportation and may consider public transportation as an affront to their own autonomy. Another possible solution to this transportation challenge is to indirectly improve network efficiency (in terms of user-centered metrics [31]) and user experience by improving individuals' spatial decision-making. This solution may be achieved with the design of geographic information displays (GIDs) that are context-responsive, personalized, and supportive for spatial learning. While a complete reliance on automated public transportation may require the development of artificial intelligence, our vision for the future of GIDs emphasizes the importance of intelligent assistance that provides relevant information to a person capable of reasoning.

For the present paper, we define GIDs as primarily visual displays that present spatial information and can facilitate navigation through a large-scale, real-world environment. Following Montello [37], we consider a large-scale environment as one that is larger than the human body and requires locomotion for apprehension. Our recommendations focus on the geographic information provided to individual pedestrians but may be extended to individuals within a multimodal transportation network. Previous research in this area tends to investigate spatial information, the visualization of spatial information, or users' knowledge of spatial information gained from different visualizations. We approach the problem from all three of these complementary perspectives: GIScience, cartography, and cognitive science (respectively).

2   GIDs from a GIScience perspective

From a GIScience perspective, some of the primary challenges associated with GIDs are context modeling, context inference, context management, and context adaptation. "Context" refers to the information used to characterize a person, place, or object that is relevant for human-system interaction [11]. Research on context modeling and GIDs attempts to derive a classification structure for sets of (spatial and non-spatial) information that are relevant for a particular task. For example, Sarjakoski and Nivala [53] classify contextual factors along seven dimensions, including user characteristics (e.g., demographics, goals, cognitive abilities), location/orientation, time, navigation history, technical properties, properties of the physical environment, and properties of the social situation. One open issue for context models and GIDs is the determination of appropriate methods for identifying and quantifying the relevance of particular sets of information.

Once relevant contextual factors are modeled, researchers can attempt to infer high-level context information from various sources of low-level raw data [24]. GIDs require context inference because raw sensor data would be uninformative for the user. Possible sources of 'raw sensor data' for GIDs include physical sensors (e.g., physiological sensors worn by the user), data from web applications and services (e.g., online social media), and users' implicit (e.g., past experience) and explicit (e.g., button press) outputs. Context inference also requires methods for integrating these different data sources because a contextual factor may not be reducible to one particular source. For example, the relative difficulty of navigating two possible routes from the user's current location to her destination may depend on both her past experience with that route and the current weather. Effective (and automated) methods and models for integrating different data sources and inferring context are still missing today [24]. Here, machine learning may prove to be an especially efficient and effective method for GIDs.

Context modeling, context inference, and context management often assume that human-system interactions will remain stable over time. Despite the growing number of raw data sources and the increasing complexity of context inference, GIDs must allow for the efficient retrieval and update of contextual information provided to the user. Human behavior changes as humans acquire more information regarding a system, and a particular contextual factor may not remain relevant forever. In order to address this challenge, technical systems must adapt to changes in context [21]. For example, during navigation, there are often frequent and rapid changes of spatial, attribute, and task contexts. Such context management is also critical for situations in which different contextual factors are interdependent. For example, users might realize that they are lost when they reach a particular boundary and then change their intended goal. Context adaptation requires an environmental awareness of changes in context and autonomous adjustments by the system in response to these changes. In addition, context adaptation can change the manner in which users interact with the system [9, 29].

3 GIDs from a cartographic perspective

From a cartographic perspective, the graphic elements represented by a GID (e.g., symbols indicating landmarks) should vary according to several "visual variables" in order to facilitate a user's understanding [6, 51]. The seven original visual variables include location, size, shape, orientation, color hue, color value, and texture. Some visual variables are more appropriate for encoding categorical, ordinal, or continuous values [51]. For example, an ordinal value on a geographic representation should be encoded with respect to the size of the graphic element rather than its shape. In contrast, a categorical value may be better represented by different shapes instead of different sizes. Additional visual variables (i.e., color saturation, arrangement, crispness, resolution, transparency) have been used to characterize existing cartographic designs and tested with real users [34]. For example, transparency effectively focuses users' attention on relevant geographic features compared to color saturation and color value [47, 62].

Visual variables may simplify geographic visualizations so that the encoded information is easy to comprehend, but new digital technologies allow for more realistic, 3D, high-resolution, animated, interactive visualizations than were previously possible. Animations have also been characterized in terms of dynamic visual variables, including moment, duration, frequency, order, rate of change, and synchronization [30, 12]. Users often prefer these relatively sophisticated visualizations over simpler visualizations of the same geographic

information [22]. For example, Hegarty and colleagues [22] found that undergraduate students and expert meteorologists rated realism, animation, detail, and 3D as desirable and effective characteristics for visual displays.

Sophisticated visualizations can also lead to improvement in the performance of navigation-related tasks. Researchers have developed and tested GIDs with advanced features such as the simultaneous representation of to-be-walked routes at multiple scales [10]. Delikostidis and colleagues [10] found that their "LandNavin" prototype led to more efficient and effective navigation behavior compared to an earlier version of Google Maps. The simultaneous representation of routes at different scales also reduced the need to frequently zoom in and out in order to orient, although the authors note that this function was somewhat confusing for some participants [10].

However, some research may indicate a performance advantage for simple visualizations because of fewer extraneous details (or "clutter") that are not task-relevant [50]. For example, Hegarty and colleagues [22] found that realistic weather maps negatively affected novices' (but not experts') performance on map inference tasks. Similarly, Wilkening and Fabrikant [68] found that realistically shaded relief maps led to less accurate performance in a slope detection task than a simple contour map. These realistic relief maps also led to worse performance than a slope map that contained more visual clutter but explicitly represented task-relevant information [68].

The extent to which one visualization leads to better performance on a particular task than another visualization also depends largely on expertise [35, 22, 54] and emotional context [14]. Expertise may even influence the definition of a particular visualization as simple or sophisticated. For example, a circuit diagram or architectural plan might appear simple and concise to an expert engineer or architect (respectively), but the same visualization might appear sophisticated and confusing to the uninitiated [54]. Similarly, different visualizations may be more or less effective in different emotional contexts during navigation. Emotionally laden landmarks may also enhance users' experience of location-based services [17] and improve recognition for the landmarks themselves [3]. In turn, memory for specific landmarks may facilitate the mental representation of the overall environment in a flexible manner.

4   GIDs from a cognitive science perspective

From a cognitive science perspective, GIDs should promote the user's mental representation of the variety of spatial relations that can be employed during navigation. Humans are extremely flexible with respect to the types of spatial relations they can mentally represent, but researchers often fail to distinguish between the corresponding types of mental spatial representations [56]. For example, people can remember the structure of cells viewed through a microscope or the arrangement of galaxies viewed through a telescope. Our discussion of spatial memory is limited to mental and external representations of large-scale environments. However, spatial memory may vary along several additional dimensions of representation, including level of abstraction, reference frame, and metric [56, 66].

Changes in the scale of an external representation of an environmental space (i.e., a map) are often accompanied by a change in the generalization of features represented by that map [61]. Specifically, less features tend to be visible at smaller map scales (i.e., for larger spaces). Similarly, mental representations tend to be more abstract (i.e., grouped into higher-level categories with fewer details) when they are acquired among a larger stimulus set with partially overlapping features (i.e., via interference) [23] or when there is a larger delay between learning and testing (i.e., via decay) [26]. In order to account for cognitive processes

that result in abstraction, modern map applications automatically generalize features with changes in map scale [5]. However, these schematized maps may not perfectly match the intentions of the human cartographers, and human cartographers may not always predict the most effective map design a priori. Thus, future research on GIDs may focus on both the implications of map design for abstraction in spatial memory and the incorporation of expert knowledge into schematization/generalization algorithms.

Spatial reference frames are the contextual information required to specify a location and orientation in space [57]. Researchers often investigate the extent to which different sources of contextual information result in a spatial memory that is oriented with respect to one source or another as evidenced by patterns of error or response time during recall [40, 20]. While people tend to prefer one reference frame over another [18], navigation often requires the translation of spatial information from one reference frame to another [63]. Mental translations between map and bodily reference frames are needed, for example, when navigators use GIDs to find a goal.

In the spatial cognition literature, researchers often manipulate reference frames using stimuli from either a bird's-eye or ground-level perspective. Cartographic maps from a bird's-eye perspective may vary with respect to their orientation (e.g., north-up versus track-up) [39]. The primary differences between north-up and track-up maps are the alignment of bodily and map reference frames and variability in the orientation of the map [39]. GIDs may simplify the mental translations required during navigation by providing a track-up map and thus improve navigation efficiency [41]. However, maps with a consistent orientation (e.g., north-up) may facilitate spatial memory for object-to-object relations (i.e., allocentric memory) [41]. In addition, images or prompts from a ground-level perspective may reflect past first-person experience of a space without aligning with the observer's current orientation [19].

Spatial memories may also vary in terms of metric (i.e., a distance function relating each location to each other location in a space). In the spatial cognition literature, a metric may be defined with respect to the underlying coordinate system of a mental representation or a property of the space being learned. According to the "dominant" theory of spatial knowledge acquisition [8, 27, 60], spatial memories become more metric and more Euclidean as the observer learns a space during navigation. In other words, with experience, the distances between mentally represented locations become more consistent, and the distance function relating different remembered locations begins to resemble the straight-line distance normally experienced in the physical environment (assuming no obstacles). However, people rarely develop an Euclidean spatial memory of familiar environments [64, 52], and spatial memories with different metrics may be acquired from the same environment simultaneously [27, 55].

The metrics of spatial memories also tend to vary because of properties of the space being learned. Euclidean memories may be rare because environments contain physical obstacles. Thus, the functional distance between locations (i.e., the amount of time required to move from one location to another) may be a more relevant metric for understanding mental spatial representations [49]. In addition, GIDs can provide spatial information with an underlying metric that is not Euclidean or functional. Indeed, many public transportation maps represent graph distances between locations. For example, signs in the London tube indicate the number of stops between locations. Some researchers suggest that people tend to mentally represent navigable spaces as graphs [28, 36], but the consequences of GIDs that represent spaces as graphs for spatial memory are largely unknown.

In general, a closer correspondence between the external representation of spaces provided by GIDs and the internal representation of spaces acquired in spatial memory is assumed to improve navigation efficiency (i.e., the speed with which one finds a goal location), although this correspondence is often under-specified in the literature [54]. In addition, there are

several reasons to consider whether this improvement in efficiency necessarily corresponds to an improvement in the accuracy of spatial memory. First, external representations of environmental features reduce the necessity of actively encoding these features in spatial memory [42, 44, 45]. Active encoding (e.g., rehearsing and elaborating familiar content) improves most types of memory, including survey knowledge [67]. Second, by providing navigation instructions along a route, GIDs reduce the need for users to make explicit navigation decisions [2, 9]. Explicit decision-making may be especially important for learning a route so that the route may be reproduced in the future without the GID. Third, GIDs draw visual attention away from the environment, so users experience the space less directly [15, 16]. Visual attention on the environment is especially important for incorporating landmark knowledge into spatial memory [14].

5   Vision for the future of GIDs

The future of GIDs requires a better understanding of users' cognitive and emotional processes. From this interdisciplinary perspective, we need new design guidelines for the development of effective and efficient GIDs that are adapted to different contexts, application domains, and presentation forms. These GIDs should also be personalized in terms of individual and group differences such as spatial learning abilities and familiarity with the environment [4]. To conclude, we propose several examples of design recommendations for future GIDs in the context of pedestrian wayfinding (see Table 1 for a summary).

Table 1 Summary of design recommendations for future GIDs.

GID element	Design recommendation
Landmarks	Emphasize emotionally relevant landmarks
Landmarks	Provide virtual landmarks via augmented reality
Landmarks	Emphasize landmarks at critical decision points
Routes	Provide multiple route options
Routes	Personalize route options to match individual preferences
Topography	Only provide sparse information under time pressure
Topography	Provide richer details without time pressure

To ameliorate the negative side effects of GIDs on different aspects of spatial memory, researchers should develop GIDs that support active encoding, explicit decision-making, and visual attention on the physical environment. For active encoding, users should be required to use the information provided by the GID in a more effortful way than current systems demand. For example, to promote explicit decision-making, GIDs could provide several route options to users at various decision points [65]. The provision of additional options (up to a point) may increase satisfaction with the option eventually chosen and improve memory for that particular option. GIDs could also employ this approach in order to provide more personalized route recommendations in the future. Finally, augmented reality applications could enhance the visualization of critical landmarks along a route in order to maintain users' attention on their immediate surroundings. Increased visual attention towards landmarks at critical decision points may improve recognition of those landmarks, and improved recognition for particular landmarks can facilitate route knowledge [55]. Together, the evidence suggests that such changes to the visualizations and instructions provided by GIDs could improve spatial memory without incurring a substantial cost or requiring significant advances in GID technologies [7].

Interaction with technology is an omnipresent and integrated part of our everyday lives. Importantly, the way we design technologies will change their everyday use, as well as the way in which we think and interact with the world in general [38]. The rise of mobile navigation technologies has a variety of benefits for users and for the efficiency of wayfinding and transportation in a complex urban society. However, scientific evidence for the negative influences of current GIDs on spatial memory and human cognition have led to discontent and worry regarding the autonomy of the navigator. More generally, technological systems with different "levels of automation" [43, 59, 7] result in different levels of user engagement while performing a particular task. For example, a passenger airplane does not require the pilot to constantly monitor and steer the vehicle over the course of a long flight, but the pilot should be sufficiently engaged with the task of flying to intervene in case of an emergency [13]. At the societal level, such GIDs are now widely accepted and intensively used, increasing the efficiency of transportation networks and sometimes preventing dangerous situations. However, given the visibility of accidents attributable to autonomous systems, they are also sometimes considered a threat to human safety [33]. Indeed the future technological progress of society may depend on the extent to which humans accept being part of an autonomous system.

The term "technological infantilizing" has been used to describe the process by which technology acquires the responsibility of humans for reasoning and leads to a gradual decrease in cognitive skills [38]. With the growing number of smartphones in the world, users may tend to extensively rely on mobile applications such as GIDs. The practical and ethical implications of a potential large-scale decrease in individuals' spatial abilities are far-reaching. A widespread dependency on mobile technology might weaken the individual and empower the corporations and institutions that provide these services, leading to oppression and control. It is therefore necessary for us to understand the extent to which the technological infantilization may surpass a users' ability to reason about space. When one seeks to develop a novel GID, he should consider the ratio between the potential benefits of such technologies and these associated risks [58].

The core functionality of future systems will still be the efficiency with which they guide us from one location to another by providing cognitively economic route instructions. With the capabilities of new GIDs to identify contextual states (e.g., traffic jams) and users' psychological states (e.g., positive or negative moods) in real time using advanced sensors, there is the potential for developers to extend beyond this core functionality. For example, a device may be able to combine movement data (e.g., velocity) with data from physiological sensors (e.g., arousal in terms of electrodermal activity) and assess users' level of stress. In a high stress state (i.e., high velocity and high arousal), the information display may emphasize sparse route information. In a low stress state (i.e., low velocity and low arousal), the user's attentional resources might allow for the processing of richer information, and devices could display additional details and/or landmarks.

Emerging technologies such as augmented reality may be useful for providing an additional layer of support for spatial navigation [48] and spatial learning [32]. One remaining challenge is the design of experimental tasks that can be used to evaluate the usability of augmented reality in a meaningful manner [25]. Here again, we must carefully consider responsive designs [46]. While such technologies provide additional data for the creation of context-responsive and personalized GIDs, we must also consider new social, ethical, and legal aspects of GID usage, including user privacy [24, 1].

References

1 Roba Abbas, Katina Michael, and MG Michael. The regulatory considerations and ethical dilemmas of location-based services (LBS). *Information Technology & People*, 27(1):2–20, February 2014.

2 Jonathan Z Bakdash, Sally A Linkenauger, and Dennis Proffitt. Comparing decision-making and control for learning a virtual environment: Backseat drivers learn where they are going. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 52, pages 2117–2121. SAGE Publications Sage CA: Los Angeles, CA, 2008.

3 Ceylan Z Balaban, Florian Roser, and Kai Hamburger. The effect of emotions and emotionally laden landmarks on wayfinding. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 36, 2014.

4 Andrea Ballatore and Michela Bertolotto. Personalizing Maps. *Communications of the ACM*, 58(12), 2015.

5 Thomas Barkowsky, Longin Jan Latecki, and Kai-Florian Richter. Schematizing maps: Simplification of geographic shape by discrete curve evolution. In *Spatial Cognition II*, pages 41–53. Springer, 2000.

6 Jacques Bertin, William J Berg, and Howard Wainer. *Semiology of graphics: diagrams, networks, maps*, volume 1. University of Wisconsin Press: Madison, 1983.

7 Annina Brügger, Kai-Florian Richter, and Sara Irina Fabrikant. How does navigation system behavior influence human behavior? *Cognitive research: principles and implications*, 4(1):5, 2019.

8 Elizabeth R Chrastil and William H Warren. From cognitive maps to cognitive graphs. *PloS ONE*, 9(11):e112544, 2014.

9 Jaewoo Chung, Francesco Pagnini, and Ellen Langer. Mindful navigation for pedestrians: Improving engagement with augmented reality. *Technology in Society*, 45:29–33, 2016.

10 Ioannis Delikostidis, Corné P.J.M. van Elzakker, and Menno-Jan Kraak. Overcoming challenges in developing more usable pedestrian navigation systems. *Cartography and Geographic Information Science*, 43(3):189–207, May 2016.

11 Anind K Dey. Understanding and Using Context. *Personal Ubiquitous Comput.*, 5(1):4–7, January 2001.

12 David DiBiase, Alan M MacEachren, John B Krygier, and Catherine Reeves. Animation and the role of map design in scientific visualization. *Cartography and geographic information systems*, 19(4):201–214, 1992.

13 Mica R Endsley. Toward a theory of situation awareness in dynamic systems. *Human factors*, 37(1):32–64, 1995.

14 Aaron Gardony, Tad T Brunyé, Caroline R Mahoney, and Holly A Taylor. Affective states influence spatial cue utilization during navigation. *Presence: Teleoperators and Virtual Environments*, 20(3):223–240, 2011.

15 Aaron L Gardony, Tad T Brunyé, Caroline R Mahoney, and Holly A Taylor. How navigational aids impair spatial memory: Evidence for divided attention. *Spatial Cognition & Computation*, 13(4):319–350, 2013.

16 Aaron L Gardony, Tad T Brunyé, and Holly A Taylor. Navigational aids and spatial memory impairment: the role of divided attention. *Spatial Cognition & Computation*, 15(4):246–284, 2015.

17 Georg Gartner. Putting Emotions in Maps – The Wayfinding Example, 2012. URL: http://http://www.mountaincartography.org/publications/papers/papers_taurewa_12/papers/mcw2012_sec3_ch08_p061-065_gartner.pdf.

18 Klaus Gramann, Hermann J Müller, Eva-Maria Eick, and Bernd Schönebeck. Evidence of separable spatial representations in a virtual navigation task. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6):1199, 2005.

19 Nathan Greenauer, Catherine Mello, Jonathan W Kelly, and Marios N Avraamides. Integrating spatial information across experiences. *Psychological research*, 77(5):540–554, 2013.

20   Nathan Greenauer and David Waller. Intrinsic array structure is neither necessary nor sufficient for nonegocentric coding of spatial layouts. *Psychonomic Bulletin & Review*, 15(5):1015–1021, 2008.

21   Amy L Griffin, Travis White, Carolyn Fish, Beate Tomio, Haosheng Huang, Claudia Robbi Sluter, João Vitor Meza Bravo, Sara I Fabrikant, Susanne Bleisch, Melissa Yamada, et al. Designing across map use contexts: a research agenda. *International Journal of Cartography*, 3(sup1):90–114, 2017.

22   Mary Hegarty, Harvey S Smallman, Andrew T Stull, and Matt S Canham. Naïve cartography: How intuitions about display configuration can hurt performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3):171–186, 2009.

23   Douglas L Hintzman. " Schema abstraction" in a multiple-trace memory model. *Psychological review*, 93(4):411, 1986.

24   Haosheng Huang, Georg Gartner, Jukka M. Krisp, Martin Raubal, and Nico Van de Weghe. Location based services: ongoing evolution and research agenda. *Journal of Location Based Services*, 2018.

25   Haosheng Huang, Manuela Schmidt, and Georg Gartner. Spatial Knowledge Acquisition with Mobile Maps, Augmented Reality and Voice in the Context of GPS-based Pedestrian Navigation: Results from a Field Test. *Cartography and Geographic Information Science*, 39(2):107–116, January 2012.

26   Janellen Huttenlocher, Larry V Hedges, and Susan Duncan. Categories and particulars: prototype effects in estimating spatial location. *Psychological review*, 98(3):352, 1991.

27   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(2):93–129, 2006.

28   Amir-Homayoun Javadi, Beatrix Emo, Lorelei R Howard, Fiona E Zisch, Yichao Yu, Rebecca Knight, Joao Pinelo Silva, and Hugo J Spiers. Hippocampal and prefrontal processing of network topology to simulate the future. *Nature communications*, 8:14652, 2017.

29   Peter Kiefer, Ioannis Giannopoulos, Vasileios Athanasios Anagnostopoulos, Johannes Schöning, and Martin Raubal. Controllability matters: The user experience of adaptive maps. *GeoInformatica*, 21(3):619–641, July 2017. URL: http://link.springer.com/10.1007/s10707-016-0282-x.

30   Barend Köbben and Mustafa Yaman. Evaluating dynamic visual variables. In *Proceedings of the seminar on teaching animated cartography, Madrid, Spain*, pages 45–51, 1995.

31   David Levinson. Perspectives on efficiency in transportation. *International Journal of Transport Management*, 1(3):145–155, 2003.

32   Hengshan Li and Nicholas Giudice. Assessment of multi-level structural and topological properties on cognitive map development in multi-level built environments. In *International Conference on Spatial Cognition 2016*, 2016.

33   Allen Yilun Lin, Kate Kuehl, Johannes Schöning, and Brent Hecht. Understanding " Death by GPS ": A Systematic Analysis of Catastrophic Incidents Associated with Personal Navigation Technologies. In *CHI 2017*, Denver, CO,USA, 2017.

34   Alan M MacEachren, Robert E Roth, James O'Brien, Bonan Li, Derek Swingley, and Mark Gahegan. Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2496–2505, 2012.

35   Sara Maggi, Sara Irina Fabrikant, Jean-Paul Imbert, and Christophe Hurter. How do display design and user characteristics matter in animations? An empirical study with air traffic control displays. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 51(1):25–37, 2016.

36   Tobias Meilinger, Gerald Franz, and Heinrich H Bülthoff. From isovists via mental representations to behaviour: first steps toward closing the causal chain. *Environment and Planning B: Planning and Design*, 39(1):48–62, 2012.

37   Daniel R Montello. Scale and multiple psychologies of space. In *European conference on spatial information theory*, pages 312–321. Springer, 1993.

38   Daniel R Montello. Cognitive research in GIScience: Recent achievements and future prospects. *Geography Compass*, 3(5):1824–1840, 2009.

39   Daniel R Montello. You are where? The function and frustration of you-are-here (YAH) maps. *Spatial Cognition & Computation*, 10(2-3):94–104, 2010.

40   Weimin Mou and Timothy P McNamara. Intrinsic frames of reference in spatial memory. *Journal of experimental psychology: learning, memory, and cognition*, 28(1):162, 2002.

41   Stefan Münzer, Hubert D Zimmer, and Jörg Baus. Navigation assistance: A trade-off between wayfinding support and configural learning support. *Journal of experimental psychology: applied*, 18(1):18, 2012.

42   Stefan Münzer, Hubert D Zimmer, Maximilian Schwalm, Jörg Baus, and Ilhan Aslan. Computer-assisted navigation and the acquisition of route and survey knowledge. *Journal of environmental psychology*, 26(4):300–308, 2006.

43   Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. A Model for Types and Levels of Human Interaction with Automation, 2000.

44   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.

45   Patrick Péruch and Paul N Wilson. Active versus passive learning and testing in a complex outside built environment. *Cognitive Processing*, 5(4):218–227, 2004.

46   Verena Radoczky. How to design a pedestrian navigation system for indoor and outdoor environments. In *Location based services and telecartography*, pages 301–316. Springer, 2007.

47   Tumasch Reichenbacher. The concept of relevance in mobile maps. In *Location based services and telecartography*, pages 231–246. Springer, 2007.

48   Gerhard Reitmayr and Dieter Schmalstieg. *Collaborative augmented reality for outdoor navigation and information browsing.* na, 2004.

49   John J Rieser, Jeffrey J Lockman, and Herbert L Pick. The role of visual experience in knowledge of spatial layout. *Perception & Psychophysics*, 28(3):185–190, 1980.

50   Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of vision*, 7(2):17–17, 2007.

51   Robert E Roth. Visual Variables. *International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology*, pages 1–11, 2016.

52   Edward K Sadalla, W Jeffrey Burroughs, and Lorin J Staplin. Reference points in spatial cognition. *Journal of experimental psychology: human learning and memory*, 6(5):516, 1980.

53   Annu-Maaria Sarjakoski L. Tiina and Nivala. Adaptation to Context — A Way to Improve the Usability of Mobile Maps. In T. Meng L. and Reichenbacher and Zipf A, editors, *Map-based Mobile Services: Theories, Methods and Implementations*, pages 107–123. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

54   Mike Scaife and Yvonne Rogers. External cognition: how do graphical representations work? *International journal of human-computer studies*, 45(2):185–213, 1996.

55   Victor R Schinazi and Russell A Epstein. Neural correlates of real-world route learning. *Neuroimage*, 53(2):725–735, 2010.

56   Victor R Schinazi, Tyler Thrash, and Daniel-Robert Chebat. Spatial navigation by congenitally blind individuals. *Wiley Interdisciplinary Reviews: Cognitive Science*, 7(1):37–58, 2016.

57   Amy L Shelton and Timothy P McNamara. Systems of spatial reference in human memory. *Cognitive psychology*, 43(4):274–310, 2001.

58   Stephen RJ Sheppard and Petr Cizek. The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualisation. *Journal of environmental management*, 90(6):2102–2117, 2009.

59 T. B. Sheridan. *Humans and automation: System design and research issues.* John Wiley & Sons, Inc.., 2002.

60 Alexander W Siegel and Sheldon H White. The development of spatial representations of large-scale environments. In *Advances in child development and behavior*, volume 10, pages 9–55. Elsevier, 1975.

61 Terry A Slocum, Robert M McMaster, Fritz C Kessler, Hugh H Howard, and Robert B Mc Master. *Thematic cartography and geographic visualization.* Prentice hall, 2008.

62 Olivier Swienty, Tumasch Reichenbacher, Simone Reppermund, and Joseph Zihl. The Role of Relevance and Cognition in Attention-guiding Geovisualisation. *The Cartographic Journal*, 45(3):227–238, August 2008.

63 Perry W Thorndyke and Barbara Hayes-Roth. Differences in spatial knowledge acquired from maps and navigation. *Cognitive psychology*, 14(4):560–589, 1982.

64 Barbara Tversky. Cognitive maps, cognitive collages, and spatial mental models. In *European conference on spatial information theory*, pages 14–24. Springer, 1993.

65 Hoang Tam Vo, Peng Ni, Chengbin Yan, Thirumurugan Eswaran, Jugal Shah, and Bob Massarczyk. Diversification of route planning results for improved user satisfaction. In *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, pages 2625–2630. IEEE, 2015.

66 Ranxiao Frances Wang. Theories of spatial representations and reference frames: What can configuration errors tell us? *Psychonomic bulletin & review*, 19(4):575–587, 2012.

67 Wen Wen, Toru Ishikawa, and Takao Sato. Individual differences in the encoding processes of egocentric and allocentric survey knowledge. *Cognitive science*, 37(1):176–192, 2013.

68 Jan Wilkening and Sara Irina Fabrikant. How do decision time and realism affect map-based decision making? In *International Conference on Spatial Information Theory*, pages 1–19. Springer, 2011.

Twenty-Five Years of COSIT: A Brief and Tasty History

Daniel R. Montello

Departments of Geography and of Psychological & Brain Sciences,
University of California Santa Barbara, USA
montello@geog.ucsb.edu

---- **Abstract** ----

In this talk, I offer a few thoughts in celebration of COSIT '14, said to be the 25th anniversary of the Conference on Spatial Information Theory.[1] I reminisce about some of the many interesting and entertaining people who have participated in COSIT over the years, and wax nostalgic over many of the incredible memories it has engendered, from Pisa to L'Aquila, and now Regensburg. Many fascinating and truly interdisciplinary explorations have occurred, and continue to occur, at this international meeting of the minds and bodies and symbol systems. I specifically touch on three of the 'curious concepts of COSIT' in my talk: *cognitive maps*, *qualitative reasoning*, and *ontologies*.

1 Cognitive maps

Cognitive map is a concept typically attributed to the rat psychologist Edward Tolman, and it appears he did coin the term [6], but the concept had been around for some time. In his 1913 paper, C.C. Trowbridge [7] called them 'imaginary maps'. But the concept certainly goes back long before that. Here is a quote from a Roman military commander [1], circa 300 CE:

> In the first place, a commander should have itineraries of all the war zones very fully written out, so that he may thoroughly acquaint himself with the intervening terrain, as regards not only distance but standard of roads, and may study reliable descriptions of shortcuts, deviations, mountains and rivers. In fact, we are assured that the more careful commanders had, for provinces in which there was an emergency, itineraries that were not merely annotated but even drawn out in colour (picta), so that the commander who was setting out could choose his route not only with a mental map but with a constructed map to examine (pp. 236–237; translated from Vegetius '*De re Militari*' [Military Institute of the Romans]).

In any case, the meaning of the concept as mostly used by environmental psychologists, geographers, and planners/landscape architects treats 'map' as a broad metaphor [5], being neutral as to the specifics of its form and geometric sophistication, but insisting that people have beliefs (knowledge) in their mind about the layouts and contents of environments in a variety of formats, including pictorial, verbal, numerical, etc. In contrast, nonhuman animal behaviorists and many perceptual/cognitive psychologists, and now neuroscientists, reserve the term for what I would call a 'survey' or 'configurational' map. In other words,

---

[1] The entire list of COSIT meetings appears below.

in the broad sense I was schooled in, a route map or even a list of landmark images is still a cognitive map. Of course, no one can be said to have a map in the cartographic sense of a comprehensive, fully covering, metrically accurate, and consistently scaled geospatial representation – in fact, not even cartographic maps really fit that description!

2 Qualitative Spatial Reasoning

This is something that behavioral/cognitive scientists such as myself first heard about at early COSIT meetings, and discussions I had with formal modelers/computer scientists at those early COSIT meetings provide for me a stark example of the challenges of interdisciplinary communication. Part of my struggle was that the concept as used by computational modelers (e.g., [3]) seemed to me to conflate two properties of spatial information: the degree or nature of quantification of spatiality (as opposed to nonquantitative expression), and the precision or resolution of the information. After all, if I am expressing distances to the nearest 100 km, then it is both quantitative and accurate to say that any distance from 450 to 550 km is '500 km'; but it is relatively vague in many contexts. Also, the term *metric* typically refers to spaces – geometries – that follow the axioms of metric geometry (which, by the way, it is trivially easy to demonstrate human spatial knowledge violates regularly). At the same time, *metric* is also used to refer to metric scales or levels of measurement, which are either interval or ratio. It is in the latter sense that I insist we should describe human spatial knowledge as metric rather than nonmetric, albeit of relatively vague resolution and often relatively inaccurate.

3 Ontology

Finally, we come to the 'O-Word'. Philosophically, this is the issue of describing the nature of that which is, that which exists. But in the 1970s, it came to mean the expression of 'reality' as instantiated by a computational model or information system [2]. Then that was quickly enough applied to models of reality as expressed by individual and group cognition [4], often in linguistic form; this usage gained some currency even though we already had perfectly fine terms for it like 'conceptual system'. In my talk, I clarify the use of *ontologies* (in the plural) as a spatial-information concept by recounting a conversation between Farmer Smith and Farmer Mark.

4 All Meetings of COSIT

COSIT 0	1992 *Pisa*, Italy	**COSIT 1**	1993 *Elba*, Italy
COSIT 2	1995 *Semmering*, Austria	**COSIT 3**	1997 *Laurel Highlands*, USA
COSIT 4	1999 *Stade*, Germany	**COSIT 5**	2001 *Morro Bay*, USA
COSIT 6	2003 *Kartause Ittingen*, Switzerland	**COSIT 7**	2005 *Ellicottville*, USA
COSIT 8	2007 *Melbourne*, Australia	**COSIT 9**	2009 *Aber W'rach*, France
COSIT 10	2011 *Belfast*, USA	**COSIT 11**	2013 *Scarborough*, UK
COSIT 12	2015 *Santa Fe*, USA	**COSIT 13**	2017 *L'Aquila*, Italy
COSIT 14	2019 *Regensburg*, Germany		

References

1 O. A. W. Dilke. Itineraries and geographical maps in the early and late Roman empires. In J. B. Harley and David Woodward, editors, *The History of Cartography, Volume 1: Cartography in Prehistoric, Ancient, and Medieval Europe and the Mediterranean*, pages 234–257. The University of Chicago Press, Chicago, 1987.

2 Andrew U. Frank. Spatial ontology: A geographical information point of view. In Oliviero Stock, editor, *Spatial and Temporal Reasoning*, pages 135–153, Dordrecht, The Netherlands, 1997. Springer.

3 Christian Freksa. Using orientation information for qualitative spatial reasoning. In Andrew U. Frank, Irene Campari, and Ubaldo Formentini, editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, volume 639 of *Lecture Notes in Computer Science*, pages 162–178, Berlin, Heidelberg, 1992. Springer.

4 Werner Kuhn. Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information*, 26(12):2267–2276, 2012.

5 Daniel R. Montello and Scott M. Freundschuh. Sources of spatial knowledge and their implications for GIS: An Introduction. *Geographical Systems*, 2:169–176, 1995.

6 Edward C. Tolman. Cognitive maps in rats and men. *The Psychological Review*, 55(4):189–208, 1948.

7 C. C. Trowbridge. On fundamental methods of orientation and "imaginary maps". *Science, New Series*, 38(990):888–897, 1913.

Perception of Space in Virtual and Augmented Reality

Sarah H. Creem-Regehr

University of Utah, Salt Lake City, USA
`https://psych.utah.edu/people/faculty/creem-regehr-sarah.php`

Abstract

Virtual and Augmented Reality (VR and AR) methods provide both opportunities and challenges for research and applications involving spatial cognition. The opportunities result from the ability to immerse a user in a realistic environment in which they can interact, while at the same time having the ability to control and manipulate environmental and body-based cues in ways that are difficult or impossible to do in the real world. The challenge comes from the notion that virtual environments will be most useful if they achieve high *perceptual fidelity* – that observers will perceive and act in the mediated environment as they would in the real world. Consider two approaches to the use of VR/AR for in cognitive science. The first is to serve applications. For this, I argue in many cases we need to achieve and measure perceptual fidelity. Specifically, perceiving sizes and distances similarly to the real world may be critical for applications in design or training where the accuracy in scale matters. The second approach is to use VR/AR to manipulate environment-body interactions in ways that test perception-action mechanisms. Our lab and collaborators take both of these approaches, as they often mutually inform each other.

I will present two examples of this dual approach to the use of VR that take advantage of the body-based feedback available in immersive virtual environments, in adults and children. The study of children's spatial cognition is an important new direction in VR research, now feasible with the emergence of head-mounted-display technologies that fit those with smaller heads. Immersive VR has great potential for education, specifically in advancing complex spatial thinking, but a foundational understanding of children's perception and action must first be established. This is particularly important because children's rapidly changing bodies likely lead to differences compared to adults in how they represent and use their bodies for perception, action, and spatial learning. Even with rapidly advancing VR technologies, one continuing challenge is how to accurately update one's spatial position in a large virtual environment when real walking is constrained by limited physical space or tracking capabilities. In my first example, I will present research that compares different modes of locomotion that vary the extent of visual or body-based information for self-motion, and tests the ability of users to keep track of their positions during self-movement. Differences in adults and children suggest reliance on different cues for spatial updating. Research in space perception in VR suggests that viewers underestimate egocentric distances in VR as compared to the real world, although the new commodity-level head-mounted-displays have somewhat reduced this effect. In a second example, I will present research that examines the role of bodies in scaling the *affordances* of environmental spaces. We use judgments of action capabilities both to evaluate the perceptual fidelity of virtual environments and to test the role of visual body representations on these judgments. Finally, I will present extensions of the use of affordances to evaluate perceptual fidelity in VR to new possibilities with AR, in which virtual objects are embedded in the real world. This work demonstrates that augmented reality environments can be acted upon as the real world, but some differences exist that may be due to current technology limitations.

Planning and Explanations with a Learned Spatial Model

Susan L. Epstein

Department of Computer Science, Hunter College and The Graduate Center of
The City University of New York, New York, New York 10065, USA
http://www.cs.hunter.cuny.edu/~epstein/
susan.epstein@hunter.cuny.edu

Raj Korpan

Department of Computer Science, The Graduate Center of The City University of New York,
New York, New York 10016, USA
rkorpan@gradcenter.cuny.edu

Abstract

This paper reports on a robot controller that learns and applies a cognitively-based spatial model as it travels in challenging, real-world indoor spaces. The model not only describes indoor space, but also supports robust, model-based planning. Together with the spatial model, the controller's reasoning framework allows it to explain and defend its decisions in accessible natural language. The novel contributions of this paper are an enhanced cognitive spatial model that facilitates successful reasoning and planning, and the ability to explain navigation choices for a complex environment. Empirical evidence is provided by simulation of a commercial robot in a large, complex, realistic world.

1 Introduction

As robots that navigate autonomously among people become increasingly prevalent, the software that controls them must address substantive issues for their control. This paper focuses on effective planning, natural communication, and adaptation to a new environment based on the robot's experience there. The thesis of our work is that a learned cognitive spatial model based on spatial affordances can support robot navigation in built environments (henceforward, *worlds*). In previous work, we demonstrated how a small, inexpensive robot could learn such a model to approximate simple worlds [13]. This paper tackles considerably more challenging worlds, for which it extends the model, plans from it, and uses it to formulate natural explanations of the robot's navigation behavior. The principal results of this paper are that the enhanced spatial model can mitigate the impact of discretization, and that planning with it supports effective navigation, transparent reasoning, and human-friendly communication.

A robot's *controller* is its decision-making software. The context of our work is *SemaFORR*, a robot controller that relies on only its travel experience and a laser range finder to learn a cognitive spatial model for a new world. The next section of this paper provides necessary background and related work. Subsequent sections describe SemaFORR's improved spatial model and how that model supports both planning and WHY, SemaFORR's question answerer for its behavior and intentions. The paper then describes and discusses empirical results that support our thesis in a realistic, large-scale world.

2 Background and Related Work

A *robot* is an embodied, artificial, mobile agent whose behavior is produced by a sense-decide-act loop. *ROS* is the state-of-the-art robot operating system [33]. We have written both SemaFORR and WHY as ROS modules. Although robots perceive continuous space and their hardware allows a broad range of possible actions, most robot controllers, including SemaFORR, discretize both space and their action set to make computation tractable.

To plan for and communicate about navigation, a robot must both represent and reason about space. Because such communication is simpler when the robot's spatial representation and reasoning are human-like, the ways people represent and reason about space are important. Rather than find common ground between the robot's egocentric perspective of its world and that of a person whose perspective is unknown, we assume an allocentric (i.e., with respect to some external fixed point) perspective.

2.1 Spatial Models and Reasoning for Humans

To represent space, a person learns a *cognitive spatial model*, a mental representation of her world. The earliest clinical evidence for this was from studies of the behavior of rats in mazes [39]. Scientists have since identified neurons in the rat's brain that suggest the existence of a Euclidean spatial model [18]. Because these neurons fire sequentially during sleep or at rest without visual input, scientists hypothesize that the rat uses them to represent space, to learn, and to plan [4]. The wide range of structures and content proposed for human cognitive spatial models includes single designated paths, graphs that record connectivity, *labeled graphs* (with metrics for distances and angles), and *surveys* (precise, allocentric metric maps) [40, 28]. Although it has been suggested that cognitive spatial models use metric distances and angles [16], more recent work indicates that cognitive maps have a non-metric, qualitative topological structure [14]. Other recent work suggests that the model a person learns is not a survey but a labeled graph [7, 8, 41]. The spatial models SemaFORR learns are most similar to labeled graphs where the relative size and position of features are recorded but precise dimensions are not. The exact nature of human cognitive maps, however, remains an important open problem in spatial cognition [42].

To reason about space for navigation, people use a variety of well-documented approaches: reactivity [36], planning [17], and satisficing heuristics [10]. These heuristics are "good enough" rules for decision making in any world, typically triggered by either percepts or an internal signal. Despite pedestrians' individual experiences and physiology, striking regularities appear in the ways that they formulate navigation heuristics [44]. These regularities include how people understand distance and direction, perceive proximity as dependent on context, and view direction as closely related to geometry. SemaFORR's reasoning mechanism incorporates many such heuristics, as well as reactivity and planning.

2.2 Spatial Models and Reasoning for Robots

To represent space for robot navigation, a controller requires some model of the robot's world. If that model must be a survey and one is not provided, the robot can methodically travel its world to create one with the state-of-the-art algorithm SLAM (Simultaneous Localization and Mapping) [30]. *Localization*, the ability to know where one is in the world, is a key challenge, because different locations may provide similar percepts (e.g., when one faces into a corner). Moreover, robots are subject to both sensor error (percepts that provide a noisy version of the ground truth) and actuator error (imprecise command execution). As the robot travels, SLAM localizes while it builds a survey of the obstructions it detects. SLAM is probabilistic, that is, it provides only likelihoods for the robot's location. In addition, a robot that relies on a SLAM-generated metric map must contend with both the errors present during the map's construction and the errors in the robot's current localization and sensing. More recently, to facilitate communication about navigation, semantic mapping has been used to represent space and apply qualitative labels to the environment [22].

Other work on spatial representations for robot navigation has emphasized hierarchical aspects. Prototype, Location, and Associative Networks (*PLAN*) represented a cognitive map with a hierarchical structure from the (egocentric) perspective of the robot [6]. In contrast, the Spatial Semantic Hierarchy (*SSH*) built an allocentric cognitive map with hierarchical metric and topological representations [23]. SSH also incorporated representations of partial knowledge and uncertainty. It was tested as a robot controller in simulation for indoor and outdoor environments, and on a physical robot in an office environment [3]. SemaFORR is hierarchical as well, both in its reasoning structure (described in Section 3.2) and in its ability to combine atomic spatial affordances into higher level ones (described in Section 3.1).

Research has also sought to adapt human-like internal representations of the environment, Thrun's robot controller integrated a grid-based metric map with a topological one [38]. A grid-based map used Bayesian updating to determine the probability that a grid cell was occupied, and the topological map partitioned the grid cells into connected regions at narrow passages, such as doors. Thrun also adapted humans' use of landmarks to guide navigation [37]. His Bayesian approach learned the location of landmarks, trained an artificial neural network to recognize them, and then used them to localize. Another, similar approach used a multi-layer representation: a global metric map, a navigation graph, a topological map, and a conceptual semantic map [43]. This approach used its spatial map for natural language dialogue with a human, and so is closest to our own. SemaFORR, however, does not require a pretrained classifier to build its model.

To move the robot from its current location to some *target* (desired location), the robot's controller must reason about space. SLAM informs a controller but does not navigate. A modern robot first formulates a *plan*, a sequence of locations (*waypoints*) from its current location to its target, in a SLAM-generated map. The robot then travels to each waypoint in turn. The granularity of the planning map, actuator error, or dynamic obstacles, however, often cause the plan to fail. In that case, the controller could repair the plan or construct a new one. Instead, SemaFORR has multiple planners and recourse to multiple heuristics for local search when its plan fails.

2.3 Spatial Models and Reasoning for Humans and Robots Together

A *natural explanation* gives transparent, intelligible, human-friendly reasons for behavior in natural language. This enables the robot to gain social acceptance and reduces confusion about the robot's abilities [24]. Explanations compare counterfactual cases, selectively include causes, and recognize that the interlocutor is a social being with her own beliefs and intentions [29].

To generate descriptions of a robot navigator's behavior, many researchers have relied on detailed, relatively opaque logs of the robot's experience [25, 35]. Natural language descriptions of a robot's travelled path have addressed abstraction, specificity, and locality [34, 32], and sought to improve sentence correctness, completeness, and conciseness [2]. Those approaches, however, used a labeled map to generate descriptions and did not explain the robot's reasoning. Other work visually interpreted natural-language navigation commands with a semantic map that showed the robot's resulting action [31]. More recently, some work has selected potentially suboptimal plans [15, 5] or behaviors [20] that are more readily understandable to humans. In contrast, our work with WHY, described in Section 4, answers questions to explain the robot's reasoning and behavior in natural language, but does not influence the robot's decisions.

3   SemaFORR

FORR (FOr the Right Reasons) is a general architecture for learning and problem solving [12]. SemaFORR is a FORR-based robot controller for autonomous navigation, where the *task* of the robot is to travel to a target. SemaFORR currently assumes perfect localization. (Future work could adapt SemaFORR to contend with noisy localization from SLAM.)

The robot's world is indoors, and the robot's sole sensor is a range finder that supports only two spatial dimensions. Thus, at any moment, the robot has a *pose* $< x, y, \theta >$ in an allocentric coordinate system, where (x, y) is the robot's *location* and θ is its *orientation* with respect to the origin. In the work reported here, a simulator provides the robot's current pose, and a range finder gauges distances to the nearest obstruction in multiple directions. SemaFORR's knowledge store holds the robot's target, its learned cognitive spatial model, its plan, a small action repertoire (turns, forward moves, and a pause), and a log of *decision points*, the poses and sensor readings when SemaFORR chose an action or formulated a plan in the current task.

3.1   Learned Cognitive Spatial Model

With SemaFORR as its controller, the robot has no access to a SLAM-generated survey. Instead, as the robot's navigation experience accumulates over a set of tasks, SemaFORR learns a cognitive spatial model on a footprint of a new world, using only its perceptual history and actions. The foundation of this model is a set of atomic *spatial affordances*, static features of the world expected to facilitate navigation there. Affordances generalize over the robot's experience and may not be architecturally precise, that is, the shape of the learned affordances may not match the physical architecture of the environment. Affordances are learned at the end of a task, from the log of the robot's poses and sensor readings, without any reference to the metric map. This cognitively-based approach learns from the robot's egocentric perspective as it travels, and does not assume any global knowledge of the world. The original affordances in the model were paths, regions, trails, conveyors, and a skeleton [13]. This paper introduces two new affordances to the model, doors and hallways, both of which address a laser scanner's limited range and its discrete approximation of continuous space. Examples of all affordances appear in Figure 1.

One important feature of the spatial model is the way atomic features support the development of higher-level ones. For example, a *path* is the ordered sequence of decision points logged for a task. While any contiguous subsequence of a path supports travel, paths are overly specific and may include errors one would want to avoid. A *trail* is a refined version of a path, also represented as a sequence of decision points, but is typically more direct

(a) Paths (b) Trails and conveyors (c) Minor diagonal hallways

(d) Regions with exits and doors (e) Skeleton
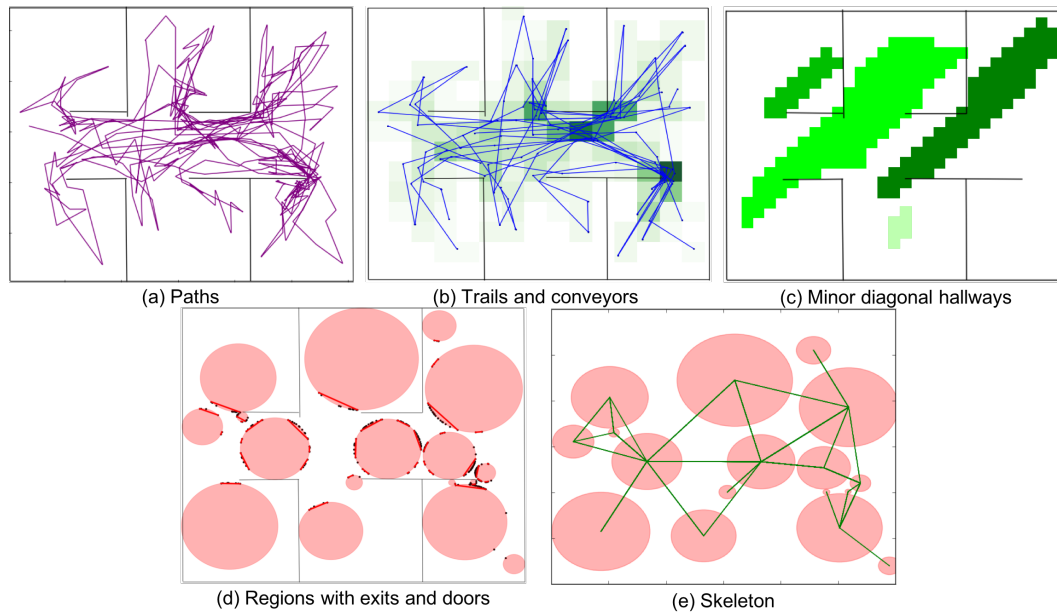
Figure 1 Affordances in a learned spatial model after visiting 40 randomly-assigned locations in a simple world (a) paths taken by the robot (b) trails refined from paths and overlaid on conveyors shaded by their count (c) minor diagonal hallways (d) regions with exits (points on the perimeter) and doors (secants) (e) the skeleton. Horizontal and vertical lines are not part of the model; they represent physical walls.

than the path from which it is derived. Trails also facilitate the construction of *conveyors*, cells in a $2m \times 2m$ grid superimposed on the footprint of the world. Conveyors tally the frequency with which trails have passed through them; those with high counts are likely to facilitate navigation because of the world's topology. A *region* represents unobstructed space as a circle whose center is a decision point and whose radius is the smallest distance to an obstacle detected there. Regions grow and shrink as the robot changes its pose. The *skeleton* is a graph whose nodes represent regions and whose edges represent the ability to move from one region to another. A path or a trail that moves from one region to the next induces an edge in the skeleton. Further details appear in [13].

Another important feature of the model is that many of the affordances are incremental. For example, a new incremental affordance is a *door*, an arc that affords access to a region along its perimeter. (For clarity in Figure 1, doors are drawn as secants on their endpoints.) Each location where a path or a trail crossed a region's perimeter is recorded as an *exit* for that region. To use an exit effectively, however, the robot's heading must align precisely with that exit. As exits accumulate, SemaFORR learns doors, generalizations about the region's circumference. A pair of exits is said to be *nearby* one another when the arc between them is less than ε. Algorithm 1 shows the pseudocode that learns doors. It moves along the circumference of any region with more than one recorded exit until it encounters a consecutive pair of nearby exits. When it finds such a pair, it records the arc between them as a door, and continues to extend the current door as long as the next exit is nearby its most recent addition. Otherwise, the algorithm resumes search for the next door. Doors for a region, along with any unincorporated exits, are recorded in the knowledge store. Data from subsequent tasks adds new exits to existing doors, identifies new doors, and merges them as necessary.

■ **Algorithm 1** SemaFORR's door-learning algorithm.

Input: *Regions, Exits*
Output: *Doors*
$Doors \leftarrow \varnothing$
for *each region R with more than one exit* **do**
 Select an exit e
 $D_e = \{e\}$
 $start \leftarrow e$
 $e' \leftarrow \varnothing$
 while $start \neq e'$ **do**
  Move clockwise from e to the next exit e'
  if e' *is within ε of e* **then**
   $D_e \leftarrow D_e \cup \{e'\}$
   $e \leftarrow e'$
  else
   if $|D_e| > 1$ **then**
    $Doors \leftarrow Doors \cup \{D_e\}$
   $e \leftarrow e'$
   $D_e = \{e'\}$
  end
 end
end
return *Doors*

The other new higher-level feature of the spatial model is a hallway. Intuitively, a hallway is a relatively straight, narrow, continuous area with both length and width. Figure 2 illustrates how horizontal hallways develop in the footprint of a simple world. Algorithm 2 is pseudocode for SemaFORR's hallway-learning algorithm. To begin, the algorithm forms a *segment* from every pair of consecutive poses in a path and the percepts at their endpoints. It then labels each segment (as horizontal, vertical, major diagonal, or minor diagonal), partitions the segments by their label, and performs the same five steps within each subset (e.g., Figure 2(a)). Step 1 identifies segment pairs (*parents*) that are most similar to one another. To do so, it calculates the similarity of each possible pair of segments, based on the distance between their midpoints and the difference in their angles. Parents are those more than three standard deviations above the mean similarity for their common label. (If none are detected, this criterion is iteratively reduced from 3σ by 0.25σ until at least one pair is found or 0 is reached. In our experience, most parents lie above 1.5σ.) Step 2 generates potential building blocks for hallways. Each pair of parents determine a *child* segment midway between them. If the child shares its parents' label, and their percepts indicate that both ends of the child would be visible from their four endpoints (i.e., the child does not pass through a wall), both parents and their child become *candidates* (Figure 2(b)). Step 3 constructs a *heatmap*, a $1m \times 1m$ grid on the footprint of the world. Initially, cells have value 0; each candidate then increments the values in the corresponding grid cells (Figure 2(c)). To smooth the heatmap, the algorithm searches for cells whose neighbors' values indicate that they should join a hallway. If a cell has value 0 but the values in at least 70% of its (at most 8) immediate neighbors meets a threshold τ (here, 1), that cell's value is set to 1 (Figure 2(d)). (Although this process is recursive, in our experience there are rarely more than two
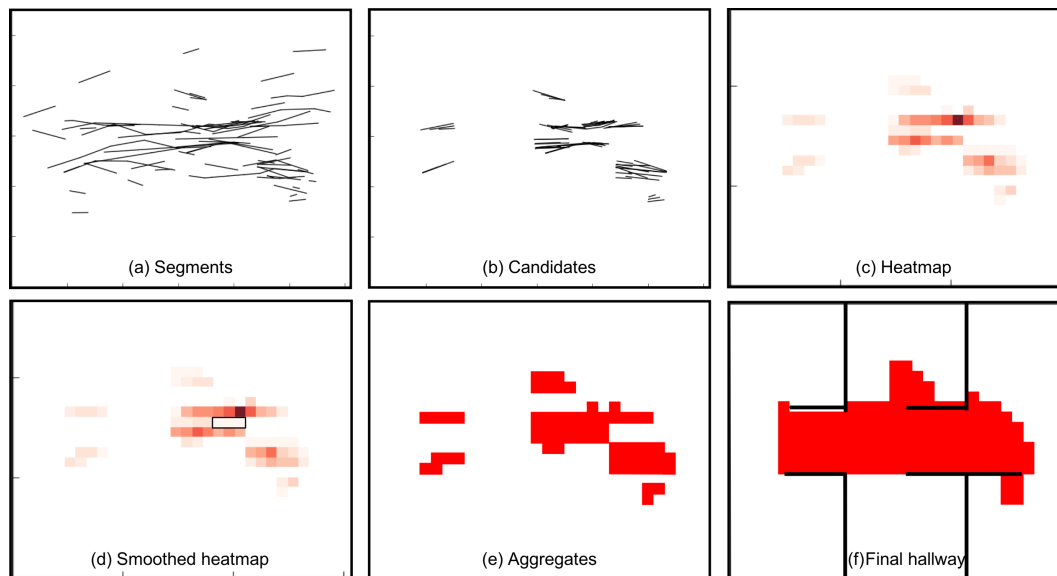
Figure 2 Stages in the development of horizontal hallways in a simple world after 40 tasks (a) segments (b) candidates (c) the heatmap (d) the smoothed heatmap with added cells indicated by the rectangle (e) aggregates (f) final horizontal hallways superimposed on the true map.
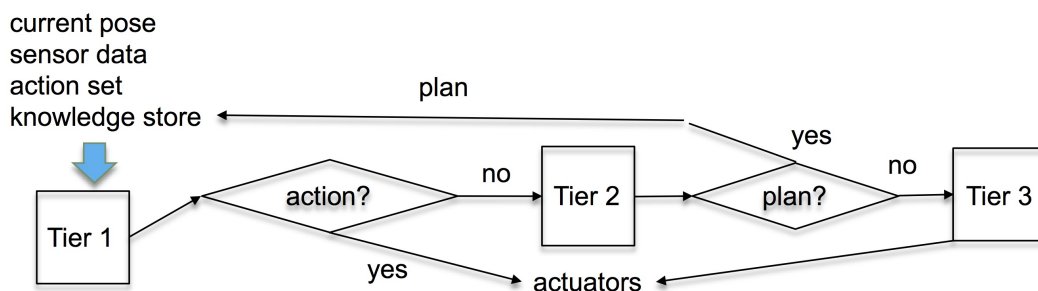
Figure 3 SemaFORR reasons with a hierarchy of Advisors.

iterations.) Step 4 uses depth-first search to find *aggregates*, connected components formed by cells with non-zero values in the heatmap (Figure 2(e)). Step 5 merges any two aggregates when each would be visible to the other, and repeats the smoothing process (Figure 2(f)). Finally, the algorithm records in the knowledge store, but does not merge, differently labeled hallways that intersect with one another.

3.2 Reasoning

As a FORR-based system, SemaFORR defines domain-specific "right reasons" called *Advisors*. An Advisor is a procedure that generates *comments*, opinions on how to navigate. Each Advisor has its own *rationale* (e.g.,"avoid walls" or "go to unfamiliar locations"), a narrow perspective on the degree to which a possible action supports or opposes success on the task. Table 1 lists the Advisors used in the work reported here.

FORR represents decision making as a combination of reaction, deliberation, and heuristic choice. To integrate those approaches, SemaFORR organizes its Advisors into the three-tier hierarchy of Figure 3. Advisors in tier 1 are reactive; they respond quickly and are assumed

■ **Algorithm 2** SemaFORR's hallway-learning algorithm.

Input: paths, laser scan history, smoothing threshold τ
Output: $Hallways$
$LineSegments \leftarrow Segment(\text{paths})$
$CardinalDirections \leftarrow Partition(LineSegments)$
for *each set of segments* $\in CardinalDirections$ **do**
    Calculate pairwise similarity for all segments in the set
    $Parents \leftarrow$ segments with similarity above dynamically-selected threshold
    $Candidates \leftarrow \{\ \}$
    for *pair* $\in Parents$ **do**
        Compute *child*
        if *child's direction* $= pair$'s direction $\land Visible(child, pair)$ **then**
            $Candidates \leftarrow Candidates \cup \{child, parent_1, parent_2\}$
    end
    $HeatMap \leftarrow ComputeHeatmap(Candidates)$
    $SmoothedHeatMap \leftarrow Smooth(HeatMap, \tau)$
    $Aggregates \leftarrow ConnectedComponents(SmoothedHeatMap)$
    $MergedAggregates \leftarrow MergeVisible(Aggregates)$
    $SmoothedAggregates \leftarrow Smooth(MergedAggregates, \tau)$
    $Hallways \leftarrow Hallways \cup SmoothedAggregates$
end
return $Hallways$

---

to be correct. A tier-1 Advisor can mandate an action (e.g., move directly to a visible target) or veto any number of actions (e.g., those that would move into a wall). Advisors in tier 2 are deliberative; each of them constructs a plan from the robot's current location to its target. Advisors in tier 3 are heuristics that comment on possible actions.

To control a robot, SemaFORR executes its sense-decide-act loop. Given its knowledge store and the data sensed by its most recent laser scan, SemaFORR moves through the Figure 3 hierarchy. In tier 1, if the target is in view and an action would immediately drive the robot to it, the Advisor VICTORY selects that action. Otherwise, if there is a current plan and the next waypoint in that plan is in view, ENFORCER selects the action that would immediately drive the robot to it. If an action was selected, the decision cycle ends, and SemaFORR sends the selected action to the robot's actuators. Otherwise, AVOIDOBSTACLES and NOTOPPOSITE veto any action that would cause a collision or return the robot to its last heading, respectively, and decision making proceeds to tier 2.

Tier 2 plans only once, at the beginning of a task, and provides waypoints for the entire task. A graph planner has an edge-weighted *cost graph* that reflects the planner's particular objective. The classic example is A*, which builds its cost graph from a grid superimposed on a map of the world, where each node represents the center of a grid cell and an edge represents unimpeded access between two cells with weight equal to the Euclidean distance between their centers. This allows A* to build shortest-path plans. SemaFORR has three planners, each of which exploits a particular category of spatial affordances: regions, hallways, or conveyors. Each planner represents its objective by adjustments to distance-based edge weights in its cost graph. For example, REGIONPLAN starts with the A* cost graph but then modifies each edge weight e between two nodes as described in Table 2. This creates a bias for paths that travel through regions.

■ **Table 1** SemaFORR's Advisors and their rationales.

Tier 1, in order	
VICTORY	Go toward an unobstructed target
ENFORCER	Go toward an unobstructed waypoint
AVOIDOBSTACLES	Do not go within ε of an obstacle
NOTOPPOSITE	Do not return to the last orientation
Tier 2 planners	
CONVEYORPLAN	Reduce cost-graph edge weights through conveyors
HALLWAYPLAN	Reduce cost-graph edge weights in hallways
REGIONPLAN	Reduce cost-graph edge weights in regions and near doors and exits
Tier 3 heuristics	
Based on commonsense reasoning	
BIGSTEP	Take a long step
CURIOSITY	Go to never visited locations
ELBOWROOM	Get far away from obstacles
ENFILADE	Go toward recent positions
EXPLORER	Go to currently unfamiliar locations
GOAROUND	Turn away from nearby obstacles
GREEDY	Go close to the target
VISUALSCAN	Turn in place to examine the world
Based on the spatial model	
ACCESS	Go to regions with many doors
CONVEY	Go to frequent, distant conveyors
CROSSROADS	Go to highly connected hallways
ENTER	Go into the target's region via an exit
ENTERDOOR	Go into the target's region via a door
EXIT	Leave a region without the target via an exit
EXITDOOR	Leave a region without the target via a door
FOLLOW	Use hallways to approach the target
LEASTANGLE	Leave a region in the target's direction
SPATIALLEARNER	Go to unmodeled locations
STAY	Stay within a hallway
TRAILER	Use a trail segment to approach the target
UNLIKELY	Avoid dead-ends in the skeleton

Algorithm 3 is pseudocode for tier 2. To resolve conflicts among its planners, each of SemaFORR's planners evaluates the plans of the others from its own perspective. Let C_{ij} be the cost of plan P_i from Advisor A_i as evaluated in Advisor A_j's cost graph. SemaFORR norms C_{ij} values in [0,10] for each i, scores plan P_i as $\sum_j C_{ij}$, selects the plan with the lowest score, places it in the knowledge store, and ends the decision cycle. Figure 4 illustrates this with three plans to travel from the lower left corner to the target (star). Each planner has produced a plan biased toward its particular objective. In Figure 4, when the plans are evaluated in the cost graphs of all three planners, REGIONPLAN has the lowest total cost because it is also relatively short (A*'s objective) and passes through a hallway (HALLWAYPLAN's objective).

If a plan is in place but multiple possible actions survive tier 1's filter, decision making passes to tier 3's heuristic Advisors. Each Advisor's rationale is deliberately narrow (e.g., "go to unfamiliar locations"), represented as a function that assigns individual *strengths*

Table 2 How REGIONPLAN exploits its spatial affordances to modify the A* cost graph.

Condition	Modified edge weight
Starts and ends in a region	$0.25e$
Only one end in a region and within 0.5 meters of a door and an exit	$0.50e$
Only one end in a region and within 0.5 meters of a door or an exit	$0.75e$
Only one end in a region but not near a door or exit	$1.00e$
Neither end in a region	$10.00e$

Algorithm 3 SemaFORR's Tier 2 procedure.

Input: *current pose, target location, spatial model, A* cost graph*
Output: *SelectedPlan*
for *each Advisor $A_i \in$ Tier 2* **do**
    Set A_i's cost graph to a copy of *A**'s cost graph
    Update A_i's cost graph based on A_i's objective and the spatial model
    Use A* search to find lowest cost plan P_i in A_i's cost graph
end
for *each Advisor A_j* **do**
    for *each plan P_i* **do**
        $C_{ij} \leftarrow$ cost of plan P_i in A_j's cost graph
    end
end
Normalize plan costs C_{ij} in [0,10]
for *each plan P_i* **do**
    $Score_i = \sum_j C_{ij}$
end
$SelectedPlan \leftarrow argmin_{i \in Tier2} Score_i$
return *SelectedPlan*

normalized in [0,10] to any subset of the remaining actions. Strengths above 5 represent support for the action; those below 5 represent opposition to it. For example, CURIOSITY supports actions that encourage the robot to travel to places in the environment it has never visited in any task, EXPLORER supports actions to unvisited locations in the current task, and SPATIALLEARNER supports actions to locations that are not included in regions, conveyors, or hallways. To capitalize on the synergy among multiple heuristics, *voting* selects the action with the maximum total score from all tier-3 Advisors, ends the decision cycle, and sends the selected action to the robot's actuators.

4   Natural Explanations

WHY uses SemaFORR's knowledge store and its Advisors' comments at a decision point to explain (and, to a limited extent, discuss) the robot's behavior. Elsewhere we have sketched our general approach to natural explanation, but only when the robot confronts crowds of moving people [21]. This work addresses explanations that reference SemaFORR's spatial model, Advisors, and reasoning structure. To answer questions, WHY identifies the Advisors that drove its decisions; their rationales are the reasons for SemaFORR's behavior.

Throughout this section, $\mathcal{N}$ represents functions that map any value to natural language.

	A* cost	Region cost	Hallway cost	Total cost
A* Plan	0	10	10	20
RegionPlan ·········	4	0	6	10
HallwayPlan — —	10	8	0	18
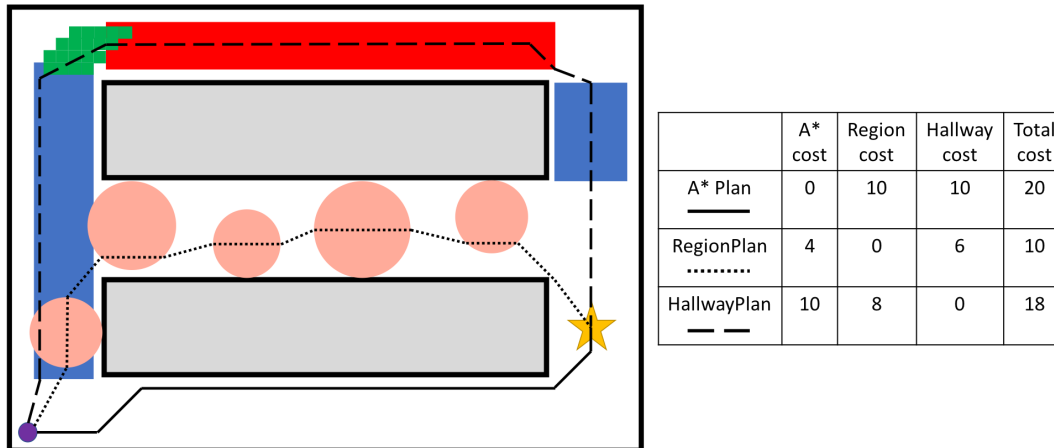
■ **Figure 4** SemaFORR's tier 2 votes to select the plan with lowest cost in a world with two large obstructions.

For example, an action a is described in natural language by $\mathcal{N}(a)$ and the rationale of Advisor A by $\mathcal{N}(A)$. WHY also calculates a variety of metrics that monitor aspects (e.g., confidence or enthusiasm) of the decision process. To explain these real values, WHY maps an ordered partition of each metric's range into natural language (also denoted by $\mathcal{N}$). For example, a metric $m \in (-\infty, +\infty)$, could be partitioned as $\{(-\infty, 0), [0, +\infty)\}$, with $\mathcal{N}(m < 0) \rightarrow$ "a little" and $\mathcal{N}(m \geq 0) \rightarrow$ "a lot." These ranges allow WHY to hedge in its responses, much the way people explain their reasoning when they are uncertain [27].

To generate an explanation, WHY completes templates with its $\mathcal{N}$ functions and appropriate punctuation and conjunctions. All examples in this section were drawn from the experiments described in Section 5. This section first describes WHY for behavior determined by tier 1 or tier 3, and then for plans determined by tier 2.

4.1 Explanations for Single Actions

"Why did you decide to do that?" This questions a particular action a. Algorithm 4 is pseudocode to produce a reply. In response, WHY takes as input the current pose, target location, spatial model, and the Advisors' comments. When SemaFORR makes a decision in tier 1, it is either because VICTORY or ENFORCER mandated it, so that WHY uses the template

I could see our [target/waypoint] and $\mathcal{N}(a)$ would get us closer to it.

or because AVOIDOBSTACLES vetoed all actions but the pause, so that WHY uses the template

I decided to wait because there's not enough room to move forward.

The inherent uncertainty and complexity of a tier-3 decision, however, requires this template's more nuanced explanation:

Although [$\mathcal{N}(\rho_{ai})\mathcal{N}(A_i)$ for A_i that oppose a],
I decided to $\mathcal{N}(a)$ because [$\mathcal{N}(\rho_{ai})\mathcal{N}(A_i)$ for A_i that support a].

WHY includes only those tier-3 Advisors with strong opinions about a, compared to other actions. Let μ_i be the mean comment strength across all actions and σ_i its standard deviation.

■ **Algorithm 4** WHY's explanation procedure for single actions.

Input: *current pose, target location, spatial model, Advisors' comments*
Output: *explanation*
switch *mode(decision)* **do**
   case *tier 1 decides action* **do**
     │ *explanation* ← sentence based on VICTORY or ENFORCER
   case *only 1 unvetoed action remains after tier 1* **do**
     │ *explanation* ← sentence based on vetoes from AVOIDWALLS
   otherwise do
     Compute relative support for tier-3 Advisors' strengths
     Categorize the support level for the chosen action
     Complete template for each Advisor with its support level and rationale
     explanation ← combined completed templates
   end
end
return *explanation*

For comment strength c_{ia} from Advisor A_i on action a, A_i's *relative support* is defined as $\rho_{ia} = (c_{ia} - \mu_i)/\sigma_i$. Because A_i has a strong opinion about a relative to the other actions only if $|\rho_{ia}|$ is large, WHY excludes A_i from its explanation if $\rho_{ia} \in (-0.75, 0.75]$. The first line in the template uses $\mathcal{N}(A_i)$ and $\mathcal{N}(\rho_{ia})$ phrases only if $\rho_{ia} \leq -0.75$; the line is omitted if no Advisors opposed a strongly enough. The second line uses $\mathcal{N}(A_i)$ and $\mathcal{N}(\rho_{ia})$ phrases only for $\rho_{ia} > 0.75$, For example, if GREEDY supports a forward move of $1.6m$ so strongly that $\mathcal{N}(\rho_{ia})$ is "I really want" but EXPLORER opposes that move, and $\mathcal{N}(a)$ is "move forward a lot," WHY would explain "Although I don't want to go somewhere I've been, I decided to move forward a lot because I really want to get close to our target."

"What action would you take if you were [here]?"   WHY substitutes the alternative pose [here] for the robot's current one, and has SemaFORR recompute its decision from the current spatial model to produce hypothetical comments. WHY then treats this as a "why did you decide" question, but substitutes "I would" for "I decided to."

"How sure are you that this is the right decision?"   This asks about the robot's *confidence*, that is, how much it believes its decision will help it reach the target. Decisions in tier 1 are by definition highly confident, so the template for VICTORY or ENFORCER is

> Highly confident, since [our target/the next waypoint in our plan] is in sensor range and this would get us closer to it.

and for AVOIDOBSTACLES the template is

> Highly confident, since there is not enough room to move forward.

Again, tier-3's uncertainty and complexity require more nuanced language. Confidence Λ_a relies on two metrics: γ_a, the extent to which the tier-3 Advisors agree with one another in their opinion of an action, and β_a, SemaFORR's *overall support* for its chosen action compared to other actions. Let $S_a = \sum_{i=1}^{v} c_{ia}$ be the total strength of possible action a when v tier-3 Advisors comment. Then the *level of agreement* on a among all Advisors is the Gini impurity of S_a, $\gamma_a = 2 \cdot (S_a/10v) \cdot (1 - S_a/10v)$, where values near 0 indicate a

high level of agreement in support or opposition and values near 0.5 indicate disagreement or lack of a strong opinion [19]. For example, if four Advisors assign equally supportive scores $[10, 10, 10, 10]$ to action a and divergent scores $[0, 0, 10, 10]$ to action a', then $\gamma_a = 0.0$ captures the agreement and $\gamma_{a'} = 0.5$ the disagreement. Overall support β_a for a compared to other actions is $\beta_a = (S_a - \mu_S)/\sigma_S$, where μ_S and σ_S are the mean and standard deviation, respectively, of S_a across all actions a. To gauge the robot's *confidence level* Λ_a, WHY weights the level of agreement and overall support equally, with $\Lambda_a = (0.5 - \gamma_a) \cdot \beta_a$, The template is

I'm $\mathcal{N}(\Lambda_a)$ sure
because $\mathcal{N}(\gamma_a)$ $\mathcal{N}(\beta_a)$
even though [$\mathcal{N}$ for whichever of γ_a or β_a is lower than Λ_a],
[$\mathcal{N}$ for whichever of γ_a or β_a is higher than Λ_a].

To complete it, WHY retrieves ordered labels for each of $\mathcal{N}(\Lambda_a)$, $\mathcal{N}(\gamma_a)$, and $\mathcal{N}(\beta_a)$. If γ_a and β_a have the same label as Λ_a, WHY uses only the first two lines. For example, "I'm really sure because I've got many reasons for it. I really want to do this the most." If only one of γ_a and β_a match Λ_a's label, WHY completes only the first line and the agreeing phrase in the second. For example, "I'm not sure because my reasons conflict." Finally, if neither γ_a nor β_a matches with Λ_a, WHY completes the first, third, and fourth lines. For example, "I am only somewhat sure because, even though I've got many reasons, I don't really want to do this the most."

"Why not do* [something else]*?" A person makes decisions with her own mental model of the world. When her decision conflicts with another's, she tries to understand why they made a different decision. To explain SemaFORR's preference for action a over an alternative b, the template for VICTORY or ENFORCER is

I decided not to $\mathcal{N}(b)$ because [I detect our target/this follows our plan]

and for AVOIDOBSTACLES or NOTOPPOSITE the template is

I decided not to $\mathcal{N}(b)$ because [$\mathcal{N}(A_i)$ for A_i that vetoed b].

The other possibility is that b scored lower in tier 3 than a did. How much SemaFORR prefers a to b is based on the difference in the two actions' overall support $\beta_a - \beta_b$. Only tier-3 Advisors with a *clear preference* for a over b (defined by $\rho_{ia} - \rho_{ib} \notin [-1, 1]$) are used to complete this template:

I thought about $\mathcal{N}(b)$
because it would let us [$\mathcal{N}(A_i)$ for A_i that prefer b],
but I felt $\mathcal{N}(\beta_a - \beta_b)$ strongly about $\mathcal{N}(a)$
since it lets us [$\mathcal{N}(A_i)$ for A_i that prefer a].

The second line is included only if any Advisors showed a clear preference for b. For example, if GREEDY preferred a, while EXPLORER preferred b, one explanation is "I thought about b because it would let us go somewhere new, but I felt slightly more strongly about a since it lets us get closer to our target."

■ **Algorithm 5** WHY's explanation procedure for plans.

Input: *robot's pose, target location, Advisors' comments, objectives $\mathcal{O}_s$ and $\mathcal{O}_q$*
Output: *explanation*
Compute plans: P_q based on $\mathcal{O}_q$ and P_s based on $\mathcal{O}_s$
Compute perspectives: $\Delta_q = C_{sq} - C_{qq}$ and $\Delta_s = C_{ss} - C_{qs}$
switch *mode(Δ_q, Δ_s)* **do**
   case $\Delta_q = \Delta_s = 0$ **do**
      | *explanation* $\leftarrow$ sentence based on template for equivalent plans
   case $\Delta_s < 0$ *and* $\Delta_q = 0$ **do**
      | *explanation* $\leftarrow$ sentence based on $\mathcal{O}_s$ (e.g., follows hallways)
   case $\Delta_s < 0$ *and* $\Delta_q > 0$ **do**
      | *explanation* $\leftarrow$ sentence based on $\mathcal{O}_s$ and $\mathcal{O}_q$ (e.g., follows hallways and length)
end
return *explanation*

4.2   Explanations for Plans

Explanations for a plan assume an alternative objective. Assume SemaFORR's current plan was produced by planner P_s with objective $\mathcal{O}_s$ in its cost graph, and that the questioner reasons instead with P_q and $\mathcal{O}_q$. Let C_{ij} be the cost of planner P_i's plan in the cost graph of planner P_j. WHY addresses the differences in the perspectives of P_s and P_q as $\Delta_q = C_{sq} - C_{qq}$ and $\Delta_s = C_{ss} - C_{qs}$. WHY's responses are based on the robot's pose, the Advisors' comments, the target, and objectives $\mathcal{O}_s$ and $\mathcal{O}_q$. As a running example, assume $\mathcal{O}_q$ is "take the shortest path" and $\mathcal{O}_s$ is "take the hallways." WHY translates objective $\mathcal{O}$ as $\mathcal{N}(\mathcal{O})$; in the example, this would be "short"and "follows hallways," respectively.

"Why does your plan go this way?"   could be asked anywhere along the robot's intended path. Algorithm 5 is pseudocode for WHY's explanation procedure. Based on the values for Δ_q and Δ_s, there are several possible cases, each with its own language template. If both are 0, then the plans equally address the two objectives, and WHY explains:

> I decided to go this way because I think it's just as $\mathcal{N}(\mathcal{O}_s)$ and equally $\mathcal{N}(\mathcal{O}_q)$.

Otherwise, the plans differ with respect to one or both objectives. If Δ_s is negative (e.g., P_s is more aligned with hallways), then WHY uses the template

> Although there may be a $\mathcal{N}(\Delta_q)$ $\mathcal{N}^*(\mathcal{O}_q)$ way,
> I think my way is $\mathcal{N}(\Delta_s)$ $\mathcal{N}^*(\mathcal{O}_s)$.

where $\mathcal{N}^*(\mathcal{O})$ is a comparator for $\mathcal{O}$ (e.g., "shorter" or "better at following hallways"). For example, an explanation could be "Although there may be a somewhat shorter way, I think my way is a lot better at following hallways." WHY omits the first line in the template if $\Delta_q = 0$. Other cases, where Δ_q is negative or Δ_s is positive, cannot occur because each planner is optimal with respect to its own objective.

"What makes your plan better than mine?"   If Δ_q and Δ_s are both 0, then WHY replies, "I think both plans are equally good." Otherwise, WHY responds with the template

> I think my way is better because it's $\mathcal{N}(\Delta_s)$ $\mathcal{N}^*(\mathcal{O}_s)$.

For example, an explanation could be "I think my way is better because it's a lot better at following hallways."

"What's another way we could go?"   In response, WHY applies the template

> We could go that way since it's $\mathcal{N}(\Delta_q)\ \mathcal{N}^*(\mathcal{O}_q)$ but it could also be $\mathcal{N}(\Delta_s)\ \mathcal{N}'(\mathcal{O}_s)$.

where $\mathcal{N}'$ denotes an opposite comparator (e.g., "longer" or "farther from known hallways"). For example, an explanation is "We could go that way since it's somewhat shorter but it could also be a lot farther from known hallways."

"How sure are you about your plan?"   WHY analyzes and explains its confidence in its objective with the template

> I'm $\mathcal{N}(P_s)$ sure because
> my plan is $\mathcal{N}(\mathcal{O}_s)\mathcal{N}^*(\mathcal{O}_s)$ and only $\mathcal{N}(\mathcal{O}_q)\mathcal{N}'(\mathcal{O}_q)$ than your plan.
> even though my plan is $\mathcal{N}(\mathcal{O}_s)\mathcal{N}^*(\mathcal{O}_s)$, it is also $\mathcal{N}(\mathcal{O}_q)\mathcal{N}'(\mathcal{O}_q)$ than your plan.
> my plan is $\mathcal{N}(\mathcal{O}_q)\mathcal{N}'(\mathcal{O}_q)$ and only $\mathcal{N}(\mathcal{O}_s)\mathcal{N}^*(\mathcal{O}_s)$ than your plan

WHY retrieves $\mathcal{N}(P_s)$, its confidence in SemaFORR's plan P_s based on Δ_s and Δ_q. To compute confidence, the values for Δ_s and Δ_q are first partitioned into three intervals each. The Cartesian product of the two partitions results in nine possible combinations. Finally, $\mathcal{N}(P_s)$ applies one of the labels ["really", "only somewhat", "not"] to each intersection. If $\mathcal{N}(P_s) =$ "really," WHY uses the second line in the template; if $\mathcal{N}(P_s) =$ "only somewhat," it uses the third line; otherwise it uses the fourth. For example, "I'm really sure because my plan is a lot better at following hallways and only somewhat longer than your plan."

5   Empirical Design and Results

The results reported here were run in simulation with Fetch Robotics' robot Freight, whose laser range finder reports 660 distances within $25m$, along a 220° arc at a rate of 15 times per second. The robot's world, shown in Figure 5, was the fifth floor of a building that occupies an entire Manhattan block (approximately $110m \times 70m$). It includes the jogs, narrow doorways, and support columns (which appear as small circles) of the original architectural floorplan. Moreover, Figure 5's four horizontal parallel hallways, and its three parallel vertical ones, provide multiple alternate routes to most targets. Nonetheless, the extent and accuracy of SemaFORR's model will be dependent upon where the robot has traveled. An example of the model learned after 40 tasks in this world appears in Figure 6.

During this experiment, the simulator localizes the robot directly within Figure 5 and reports the percepts it would experience; SLAM is not used. An *experiment* was a sequence of 40 preselected, randomly chosen targets to visit (*tasks*). To encourage a variety of challenges, there were 5 such experiments, each with a different set of 40 targets. The robot always began an experiment in the same pose and addressed its tasks in their given order. Each task after the first began wherever the previous one had ended. If the robot did not reach its target after 500 decision steps, it *failed* that task and began to address the next task from its current pose. Evaluation metrics were total (wall clock) travel time in seconds, total travel distance in meters, percentage of successful tasks, and *coverage*, the fraction of the world's footprint covered by the spatial model, as evaluated in a $1m \times 1m$ grid.

We tested SemaFORR with the full spatial model, all the Advisors in Table 1, and the procedure to select a plan in Algorithm 3. We also tested ablated versions that kept tier 1 but dropped Advisors from other tiers. The *model-free* version had only an A* planner in tier 2 and the commonsense Advisors in tier 3; it entirely ignores spatial models. Two other versions, RegionFocused and HallwayFocused, had only the planner for their affordance

Figure 5 Test world for SemaFORR and Why.
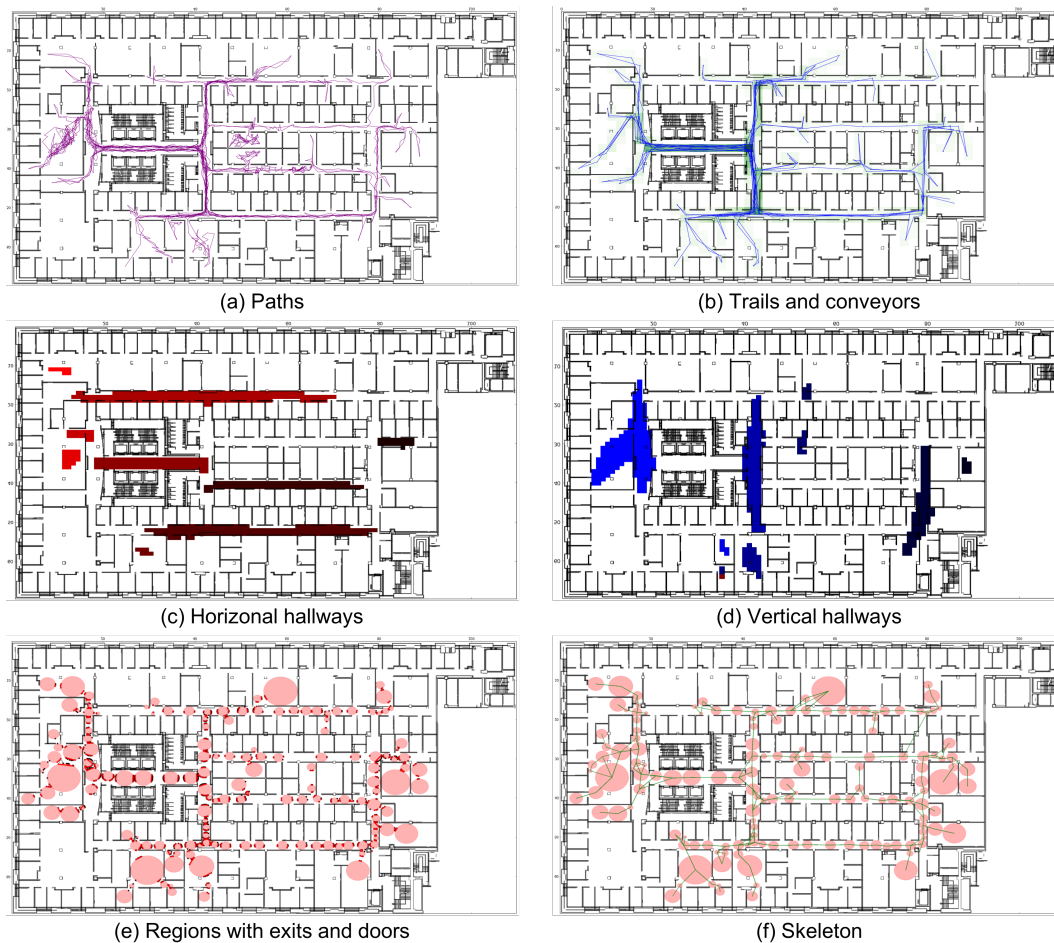
Table 3 Performance of SemaFORR and ablated versions.

Navigator	Travel Time	Distance	Success Rate	Avg. Coverage	Final Coverage
Model-free	3538.61	3475.35	89.40%	——	——
RegionFocused	**2846.13**	**3087.67**	**95.00**%	7.69%	12.36%
HallwayFocused	**3142.10**	3317.62	**92.90**%	6.21%	10.92%
SemaFORR	**2791.26**	**3194.94**	**93.90**%	10.95%*	17.65%*

(RegionPlan or HallwayPlan), the commonsense Advisors, and any tier-3 Advisors that used their affordance (regions, doors, and exits for RegionFocused, and hallways for HallwayFocused). Both planning versions use affordances to represent unobstructed space in the environment and the connectivity of the space. By ablating these versions, the experiment is able to tease out the difference in these approaches, their ability to represent connectivity, and their usefulness for planning.

SemaFORR averaged 137.48 decisions per task, and each decision required 0.04 seconds. After its one-time planning, SemaFORR made about 64% of its decisions in tier 1 and 36% in tier 3. Tier 2 selected on average 38.46% of its plans from RegionPlan, 28.21% from HallwayPlan, and 35.90% from ConveyorPlan. The spatial model required about 9 seconds to learn and revise at the end of each task.

The results in Table 3 report average performance across 25 runs (5 iterations on 5 sets of 40 targets each). Data in boldface indicates statistically significant improvements compared to the model-free version. Both RegionFocused and SemaFORR produced plans that allowed the robot to travel a shorter distance than the model-free version. All three alternatives to the model-free version enabled the robot to reach its targets more quickly and succeed more often ($p = 0.05$). The only statistically significant differences (denoted by an asterisk) between SemaFORR and RegionFocused lie in their coverage: SemaFORR's coverage is greater than that of RegionFocused, both during an experiment (measured after each of the 40 tasks and averaged) and at its completion (after 40 tasks). Although some region-related Advisor is deemed supportive in 63.17% of all explanations, SemaFORR draws on a richer set of reasons from its full spatial model and, in the end, has learned more about its world.

Why's tables for $\mathcal{N}$ generate distinct natural explanations that simulate people's ability to vary their explanations based on their context [26]. To examine its explanations, we ran an experiment with HallwayPlan for P_s and A* for P_q. The system learned the full spatial model as it navigated to 80 targets and answered every question described in Section 4 at each decision point. Why averaged less than 7 msec to compute each explanation. The results in Table 4 show that this approach is also nuanced, with many unique explanations per question.

(a) Paths

(b) Trails and conveyors

(c) Horizontal hallways

(d) Vertical hallways

(e) Regions with exits and doors

(f) Skeleton

Figure 6 An example of SemaFORR's learned spatial model in the test world.

The Coleman-Liau index measures text readability [9]; it gauged WHY's explanations at approximately a sixth-grade level, and thus readily understandable to a layperson.

6 Discussion

SemaFORR can serve as a robot controller for autonomous navigation in simulation, as it was used here, or on the floor. It can also merely observe and comment upon the behavior of a robot that has a range sensor but navigates with a different controller. Moreover, it can be used in dynamic worlds where it learns and exploits crowd models, in tiers 2 and 3 [1].

Learning a cognitive spatial model takes experience. If the robot does not travel within sensor range of an area, it will have no model for it. For example, in preliminary work we implemented TRAILPLAN, a planner that that relied only on trails in tier 2, and tested an ablated version called TrailFocused. That approach quickly preferred to reuse just a few early trails and therefore explored, and learned, very little. This considerably degraded the coverage of its learned spatial model; TrailFocused repeatedly failed and was eliminated from the study. Although CONVEYORPLAN developed more credible plans, the ablated version, ConveyorFocused, experienced similar difficulties and so was not evaluated separately.

Table 4 Analysis of explanation results by tier.

	Tier 1	Tier 2	Tier 3
Number of explanations	100,797	9,796	23,966
Average computation time (msec)	0.62	5.59	0.50
Number of unique phrasings			
Why did you do that/Why does your plan go that way?	13	6	2,655
How sure are you?	3	6	7
Why not do something else/What's another way to go?	20	6	12,592
What makes your plan better than mine?	—	4	—
Total	36	22	15,254
Average readability			
Why did you do that/Why does your plan go that way?	4.32	5.89	5.15
How sure are you?	7.49	7.30	8.36
Why not do something else/What's another way to go?	5.42	5.35	6.55
What makes your plan better than mine?	—	7.25	—
Overall	5.48	6.45	6.60

Instead of learning from navigation experience, one could simply position the robot in multiple locations throughout an architectural drawing or a SLAM-based map. The resultant model, however, may not detect useful, task-oriented affordances. In contrast, SemaFORR's model reflects the robot's experience, and the controller can resort to its commonsense heuristics in areas without coverage. Alternatively, an offline process could initialize SemaFORR's model and then be augmented and modified as the robot travels.

In realistic worlds, planners are essential. SemaFORR without any planners failed on most tasks in Figure 5, and so learned little or no spatial model. A graph-based planner with too coarse a grid can also fail, because sequences of waypoints in its cost graph become less reliable. Our robot is nearly as broad as some doorways; it can only leave a room if it approaches the door at just the right angle. As a result, we used a relatively fine grid, which produces a large graph (approximately 85,000 vertices and 170,000 edges). A* is optimal because its heuristic is admissible and consistent. Without such a heuristic, SemaFORR's model-based planners use Dijkstra's algorithm [11], whose theoretical time complexity is the same as A*'s, but whose average case performance is worse. As a result, the model-based planners required significantly more time (about 1 minute versus 15 seconds) than A*. Nonetheless, navigation with them proved more successful.

SemaFORR's affordance-based planners consider distance but do not assume that all unobstructed grid cells have identical features. In our experiments, A* plans tended to hug the walls and travel through tight spaces (e.g., narrow hallways), where turns were difficult and the robot often became stuck. For a robot with fragile or unstable cargo, the smoothness of a hallway or the range of available actions within a region may also be important.

SemaFORR's spatial model is hierarchical, graph-oriented, and has well-defined semantics, all features observed in the models that people generate. There are, however, no landmarks and its graphs are not labeled. Current work investigates ways to accelerate model-based planning, including admissible heuristics that would support A* in model-based cost graphs. Future work includes landmarks, other sensors, extended dialogue (e.g., queries to the user), and human subjects to gauge the quality of WHY's explanations and the reasonableness of its current values for $\mathcal{N}$. Meanwhile, SemaFORR demonstrates the power of a cognitive spatial model to inform both planning and user-friendly explanations, and to support autonomous navigation through the complexities of a large realistic world.

References

1 A. Aroor, S. L. Epstein, and R. Korpan. Online Learning for Crowd-sensitive Path Planning. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1702–1710, 2018.

2 D. P. Barrett, S. A. Bronikowski, H. Yu, and J. M. Siskind. Driving Under the Influence (of Language). *IEEE Trans. on Neural Networks and Learning Systems*, pages 1–16, 2017.

3 P. Beeson, J. Modayil, and B. Kuipers. Factoring the mapping problem: Mobile robot map-building in the hybrid spatial semantic hierarchy. *The International Journal of Robotics Research*, 29(4):428–459, 2010.

4 L. Buhry, A. H. Azizi, and S. Cheng. Reactivation, replay, and preplay: How it might all fit together. *Neural Plasticity*, 2011, 2011.

5 T. Chakraborti, S. Sreedharan, S. Grover, and S. Kambhampati. Plan explanations as model reconciliation. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 258–266. IEEE, 2019.

6 E. Chown, S. Kaplan, and D. Kortenkamp. Prototypes, location, and associative networks (PLAN): Towards a unified theory of cognitive mapping. *Cognitive Science*, 19(1):1–51, 1995.

7 E. R. Chrastil and W. H. Warren. Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cgnition*, 39(5):1520–1537, 2013.

8 E. R. Chrastil and W. H. Warren. From cognitive maps to cognitive graphs. *PloS One*, 9(11):1–8, 2014.

9 M. Coleman and T. L. Liau. A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology*, 60(2):283–284, 1975.

10 J. A. Conlin. Getting around: making fast and frugal navigation decisions. *Progress in Brain Research*, 174:109–117, 2009.

11 E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

12 S. L. Epstein. For The Right Reasons: The FORR Architecture For Learning In A Skill Domain. *Cognitive Science*, 18(3):479–511, 1994.

13 S. L. Epstein, A. Aroor, M. Evanusa, E. I. Sklar, and S. Parsons. Learning Spatial Models for Navigation. In *Proceedings of the 12th International Conference on Spatial Information Theory - Volume 9368*, COSIT 2015, pages 403–425. Springer-Verlag New York, Inc., 2015.

14 P. Foo, W. H. Warren, A. Duchon, and M. J. Tarr. Do humans integrate routes into a cognitive map? Map-versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2):195–215, 2005.

15 M. Fox, D. Long, and D. Magazzeni. Explainable Planning. In *IJCAI-17 Workshop on Explainable AI (XAI)*, 2017.

16 C. R. Gallistel. *The organization of learning*. The MIT Press, 1990.

17 R. G. Golledge. Human Wayfinding and Cognitive Maps. In *Wayfinding Behavior: Cognitive mapping and other spatial processes*, pages 5–45. Hopkins University Press, 1999.

18 T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436:801–806, 2005.

19 T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer, 2nd edition, 2009.

20 S. H. Huang, D. Held, P. Abbeel, and A. D. Dragan. Enabling robots to communicate their objectives. *Autonomous Robots*, 43(2):309–326, 2019.

21 R. Korpan, S. L. Epstein, A. Aroor, and G. Dekel. WHY: Natural Explanations from a Robot Navigator. In *AAAI Fall Symposium on Natural Communication for Human-Robot Collaboration*, 2017.

22 I. Kostavelis and A. Gasteratos. Semantic mapping for mobile robotics tasks: A survey. *Robotics and Autonomous Systems*, 66:86–103, 2015.

23 B. Kuipers. The spatial semantic hierarchy. *Artificial intelligence*, 119(1-2):191–233, 2000.

24 T. Kulesza, S. Stumpf, M. Burnett, S. Yang, I. Kwan, and W. K. Wong. Too Much, Too Little, or Just Right? Ways Explanations Impact End Users' Mental Models. In *IEEE Symposium on Visual Languages and Human-Centric Computing*, pages 3–10. IEEE, 2013.

25 C. Landsiedel, V. Rieser, M. Walter, and D. Wollherr. A Review of Spatial Reasoning and Interaction for Real-World Robotics. *Advanced Robotics*, 31(5):222–242, 2017.

26 B. F. Malle. How People Explain Behavior: A New Theoretical Framework. *Personality and Social Psychology Review*, 3(1):23–48, 1999.

27 R. Markkanen and H. Schröder. *Hedging and Discourse*, volume 24 of *Research in Text Theory*. Walter de Gruyter, 1997.

28 T. P. McNamara. Mental representations of spatial relations. *Cognitive psychology*, 18(1):87–121, 1986.

29 T. Miller. Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence*, 2018.

30 M. Montemerlo and S. Thrun. FastSLAM 2.0. *FastSLAM: A scalable method for the simultaneous localization and mapping problem in robotics*, pages 63–90, 2007.

31 J. Oh, T. M. Howard, M. R. Walter, D. Barber, M. Zhu, S. Park, A. Suppe, L. Navarro-Serment, F. Duvallet, A. Boularias, et al. Integrated Intelligence for Human-Robot Teams. In *International Symposium on Experimental Robotics*, pages 309–322. Springer, 2016.

32 V. Perera, S. P. Selveraj, S. Rosenthal, and M. Veloso. Dynamic Generation and Refinement of Robot Verbalization. In *25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 212–218. IEEE, 2016.

33 M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. ROS: An Open-source Robot Operating System. In *Proceedings of the ICRA Workshop on Open Source Software*, volume 3, page 5. IEEE, 2009.

34 S. Rosenthal, S. P. Selvaraj, and M. Veloso. Verbalization: Narration of Autonomous Mobile Robot Experience. In *Proceedings of the International Joint Conference on Artificial Intelligence*, volume 16, pages 862–868, 2016.

35 R. Scalise, S. Rosenthal, and S. Srinivasa. Natural Language Explanations in Human-Collaborative Systems. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 377–378. ACM, 2017.

36 H. J. Spiers and E. A. Maguire. The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology*, 28(3):232–249, 2008.

37 S. Thrun. Finding landmarks for mobile robot navigation. In *Proc. of the IEEE International Conference on Robotics and Automation*, volume 2, pages 958–963. IEEE, 1998.

38 S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71, 1998.

39 E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948.

40 B. Tversky. Cognitive maps, cognitive collages, and spatial mental models. In *European Conference on Spatial Information Theory*, pages 14–24. Springer, 1993.

41 W. H. Warren, D. B. Rothman, B. H. Schnapp, and J. D. Ericson. Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition*, 166:152–163, 2017.

42 S. M. Weisberg and N. S. Newcombe. How do (some) people make a cognitive map? Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(5):768–785, 2016.

43 H. Zender, O. M. Mozos, P. Jensfelt, G. Kruijff, and W. Burgard. Conceptual spatial representations for indoor mobile robots. *Robotics and Autonomous Systems*, 56(6):493–502, 2008.

44 B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3931–3936. IEEE, 2009.

Schematic Maps and Indoor Wayfinding

Christina Bauer
Chair of Information Science, Universität Regensburg, Regensburg, Germany
christina2.bauer@ur.de

Bernd Ludwig
Chair of Information Science, Universität Regensburg, Regensburg, Germany
bernd.ludwig@ur.de

—— **Abstract** ——

Schematic maps are often discussed as an adequate alternative of displaying wayfinding information compared to detailed map designs. However, these depictions have not yet been compared and analyzed in-depth. In this paper, we present a user study that evaluates the wayfinding behaviour of participants either using a detailed floor plan or a schematic map that only shows the route to follow and landmarks. The study was conducted in an indoor real-world scenario. The depictions were presented with the help of a mobile navigation system. We analyzed the time it took to understand the wayfinding instruction and the workload of the users. Moreover, we examined how the depictions were visually perceived with a mobile eye tracker. Results show that wayfinders who use the detailed map spend more visual attention on the instructions. Nevertheless, the depiction does not help to solve the task: they also needed more time to orient themselves. Regarding the workload and the wayfinding errors no differences were found.

1 Introduction

Maps are the main means of displaying information during a wayfinding task [22]. However, these depictions can vary a lot in terms of details, scale, etc. [13]. Schematic maps, i.e. maps that only contain information like the route to follow without a detailed depiction of the environment, are discussed as an adequate representation to convey wayfinding information [9]. In this context, it is still an open question whether these maps contain enough information to solve a wayfinding task and how these maps are perceived - especially in indoor environments:

> "Mobile maps can differ in scale, content, and style. As a result, the effectiveness of different types of maps (ranging from sketch or schematic map to topographic map or other detailed map) for indoor route communication should be evaluated. However, little work has been done on that." [13, p. 312]

Therefore, we addressed the research question whether schematic maps are an efficient means of presenting wayfinding information compared to commonly used detailed map depictions. For this purpose, we used a mobile navigation system in a real-world indoor scenario. To gain a deeper understanding of the visual perception of the different map designs, we moreover analyzed the gaze behaviour of participants with a mobile eye tracker. The remainder of this paper is structured as follows. First, we give an overview of the related work concerning schematic map design and mobile eye tracking during a wayfinding task. Then, we describe our experiment followed by the results. Lastly, we discuss our findings and future work.

2   Related Research

2.1   Wayfinding and Schematic Maps

Wayfinding is the part of navigation that requires substantial cognitive processes and spatial reasoning to orient oneself [22]. For this purpose, wayfinders often search for salient objects, i.e. landmarks in their environment that help to identify their own position relative to these points [11]. Therefore, a wayfinding aid should help the user to find these objects. Consequently, we included landmarks in our map designs. Moreover, we conducted a prestudy to identify suitable objects.

The wayfinding task can basically be solved without an aid, e.g. with the help of the cognitive map of a person [1]. However, especially in unfamiliar areas or if the destination is not known, persons need the help of an aid [35]. In this context, maps are the most common means of presenting spatial information [22]. As already pointed out, these depictions can vary a lot in terms of displayed content, especially in indoor environments [13], which are the focus areas of this study. However, there are no real design guidelines for the creation of wayfinding maps [18]. Consequently, a framework is missing that informs which elements should be presented to solve a specific wayfinding task under certain circumstances [8]. In this context, schematic maps are often discussed as an adequate presentation of wayfinding information [9, 20]. However, to the best of our knowledge, these depictions have not been compared and analyzed in-depth yet. These maps abstract the depiction and try to convey only the information needed for wayfinding. Commonly, this includes the route to follow and landmarks [16]. In contrast to this, detailed maps depict the environment in much more detail. Every possible path is visualized and more information about the environment is given. In indoor environments for instance, rooms, staircases and the closing direction of doors are often presented (see e.g. [24]).

Our research focus lies on displaying different map depictions with the help of a mobile navigation system. Previous studies in these research areas show that especially landmarks should be depicted more clearly in mobile map designs to support wayfinding (see e.g. [4]). Moreover, it is once again recommended to simplify map depictions to avoid that wayfinders focus on interface elements that are not immediately relevant for their current task [29].

2.2   Eye Tracking in the Field of Wayfinding Research

It is common to analyse the time it takes to accomplish the task (see e.g. [34]) or to measure subjective feelings like the experienced workload (see e.g. [28]) to evaluate mobile navigation systems. We additionally used a mobile eye tracker to analyse the gaze behaviour of the participants. This variable allows to analyze cognitive processes during a wayfinding task [15]. Although it is relatively common to analyze gazes in wayfinding research, this measurement method has some drawbacks. The studies are often conducted in the lab, mainly due to the complications caused by direct sunlight [10]. Since the post experiment annotation process is often cumbersome, frequently small sample sizes with less then 20 participants are used [7]. Studies in indoor environments are rare, only Schnitzler et al. [30] analyzed the use of paper and digital maps in this context. Their results showed no differences between the two depictions.

To overcome all of this research gaps, i.e. small sample sizes and scarce indoor field studies, we conducted a large scale user study with 118 participants in a complex indoor area. The participants used a mobile navigation system and accomplished a wayfinding task in the field.
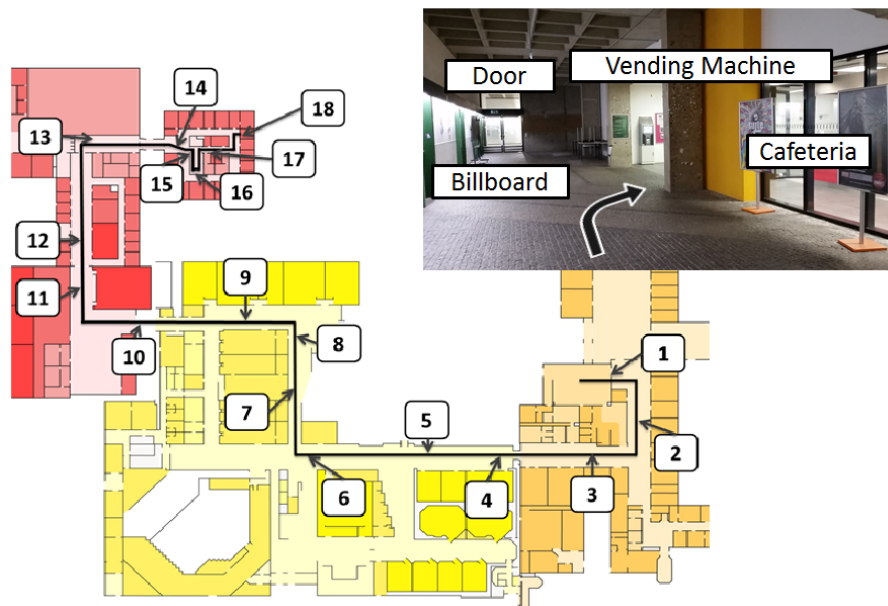
Figure 1 Test route with different colours representing different buildings. Arrows indicate where the participants received new navigation instructions. The example illustrated step 2 with predefined landmarks for the prestudy.

3   Study

In order to address our research question how the wayfinding behaviour differs if schematic maps are used compared to detailed map designs, we conducted a study with 118 participants and a between group design, i.e. participants only navigated with one of the depictions. The following sections describe the study set-up in detail, focusing on the chosen test route, the participants and the interface design. Moreover, necessary annotations are described.

3.1   Test Route

The study took place in a large-scale university building. The test route was about 375 meter long and led through three different buildings (see Figure 1). The first two buildings mainly consist of open spaces such as halls or big corridors. The last part of the route was located at an office building and therefore was dominated by narrow hallways and more changes of direction. All in all, the route consisted of nine changes of direction and three floor changes. In order to identify landmarks that could be displayed in the different maps a prestudy was conducted. As a first step, decision points along the route were determined. For this purpose, potential and "real" decision points, i.e. points were a change of direction was necessary or potentially possible (see e.g. [19]) were taken into account. Moreover, at several route points spatial barriers such as doors and stairs had to be crossed. Therefore, an instruction could be necessary at these points and they were included. This resulted into 18 steps where an instruction should be provided. For every of these points a set of four landmarks was predefined according to the findings of Viaene et al. [33] and Ohm et al. [27]. This resulted in a test sample that mainly consisted of doors, stairs and furniture. Afterwards, 87 participants (44 male, 74 students, mean age = 23.12, SD = 4,46) rated the salience of every object using the questionnaire of Kattenbeck [14]. The participants were

■   **Table 1** Identified landmarks and mean salience rating on a 5-point Likert scale (L = length of the route part in meter).

Step	Instruction	Landmark	Rating	L
1	Turn right after the stairs.	stairs	4.36	15.8
2	Turn right after the cafeteria.	cafeteria	4.16	19.1
3	Go straight ahead through the door.	door	3.44	25.9
4	Go straight ahead and pass the billboard.	billboard	4.14	24.4
5	Go straight ahead and pass the door.	door	3.74	27.1
6	Take the stairs on the right.	stairs	1.70	20.2
7	Go straight ahead and down the stairs.	stairs	3.88	25.5
8	Turn left in front of H6.	room H6	3.59	14.5
9	Go straight ahead and pass H9.	room H9	3.41	33.5
10	Turn right at the stairs.	stairs	3.96	22.8
11	Go straight ahead through the door.	door	3.67	17.3
12	Go straight ahead through the door and turn right.	door	3.31	37.8
13	Go straight ahead through the door.	door	3.63	29.9
14	Go straight ahead through the door.	door	3.86	13.1
15	Turn right and go up the stairs.	stair	4.02	6.5
16	Go through the door on the right.	door	3.34	9.5
17	Go straight ahead through the door.	door	3.77	8.5
18	Your destination is on the right.	destination	no rating	7.2

positioned at all points illustrated in Figure 1 and a map fragment showing the objects that had to be rated and the route was shown to them. As a result, the most salient object for every scene could be determined and was displayed in the map. Figure 1 also illustrates an example (step 2). Here, for instance, the cafeteria was rated as the most salient landmark (mean = 4.16 on a 5-point Likert scale) compared to the other landmarks (door mean = 3.00, billboard mean = 2.20, vending machine mean = 3,60). An overview of the chosen landmarks, their ratings and the formulated instructions is given in Table 1. Please note that at step 6 a landmark with a low rating was chosen. These stairs had to be climbed and were therefore included in the instructions. These led to wayfinding problems described in Chapter 4.

3.2   Interface Design

The prototypes were implemented in Android. The interface is subdivided into four main sections (see Figure 2). The upper right part displays a text instruction which indicates the route to take. The instructions were generated by the test designers and followed a fixed structure. They incorporated the landmark identified in the prestudy (see Section 3.1) and a simplified direction instruction, thus only referring to "left", "right" and "straight ahead" (see an overview in Table 1). In the upper left corner an arrow illustrates the direction to take, here again only showing the directions "left", "right" and "straight ahead". In addition, the landmark used to give a wayfinding instruction is displayed using an icon positioned relative to the arrow according to the current route segment. Both, the text instruction and the simplified arrow were displayed according to the recommendations of Butz et al. [5] and Kray et al. [17]. These two elements did not differ for the two test groups.
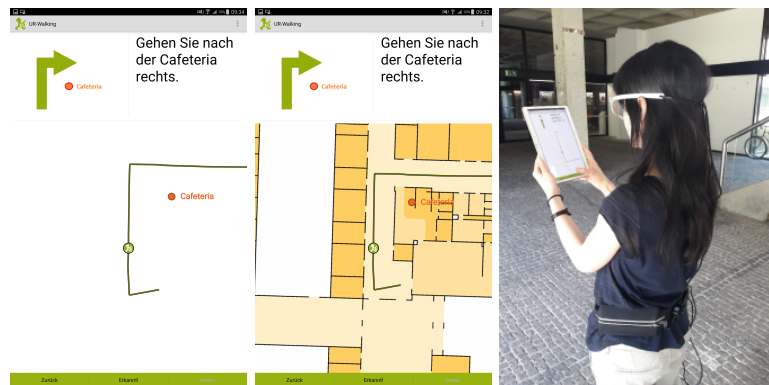
Figure 2 Interface designs (left: schematic map; middle: detailed map, right: experimental set-up). The screenshots show step 2 of the test route (see Figure 1 and Table 1). The instruction says "Go right after the cafeteria".

The main and biggest part of the screen in the middle shows the map fragment. Here, the designated current position of the user is indicated with a green manikin. The schematic map only shows the route to follow, the position of the user and the landmark, which is considered to be the minimal amount of information needed to solve a wayfinding task [32]. The detailed map is designed according to Butz et al. [5] and is a common visualisation in related studies (see e.g. [24]). Only indoor information is displayed, which means that e.g. trees and benches outdoors are not visualized. Rooms and hallways are displayed in different colours inspired by the visualisation in Schnitzler et al. [30]. Except for the landmark identified in the prestudy, no (additional) landmarks were displayed. The maps were designed for the study purpose and therefore especially for a wayfinding task. The route to follow was visualised in both map designs. The interface was "zoomable", however, the initial zoom level was fixed so that the landmarks were visible for every step. Unfortunately, no localization technique was available for this study. Therefore, the participants had to request the next navigation instruction by clicking on a "Next"-button located at the bottom right of the screen. It was possible to see previous screens using the "Back"-button (bottom left). Between these two buttons an interface element labeled "Recognized" was located. This button had to be pressed as soon as the participant had understood the instruction. By this, we wanted to record the time needed for orientation independently of the time needed for movement. This reflects the division of the navigation process in wayfinding and locomotion described by Montello [21]. The Next-button was activated only after the Recognized-button was clicked. The recognition time was considered as one of our main dependent variables.

The accuracy of the sensors used to determine the orientation of the users decreased to an insufficient level, which is a frequently reported problem in indoor ares (see e.g. [6]). Therefore, the map fragment was always oriented in direction of movement, which is preferred by users compared to north-up maps (see e.g. [31]).

3.3 Procedure and Annotation

The experiment took place in a university building during the lecture period between 10am and 16pm. The participants were picked up outside of the building and then led to the starting point of the test route. Before the experiment started they were asked to fill in the sense of direction self-assessment questionnaire of Münzer and Hölscher [23].

Figure 3 Examples of the annotation process. Left: Reference views (the boxes are labeled with "Landmark" and "Environment"). Right: Recording. Top: Gaze on a landmark referenced in the maps. Middle: Gaze on an object in the environment not referenced. Bottom: Gaze on the map element of the screen. Circular markers represent gazes.

In addition, demographic data and familiarity with smart phones and pedestrian navigation systems was collected. After this, the eye tracker was put on and calibrated using the one-point calibration. The calibration process was repeated if a gaze offset was detected. The achieved mean tracking ratio was 94.87 %. An example for the experimental set-up is shown in Figure 2. The application was started and its handling was explained using the first screen at the starting point of the route. Consequently, this step was not taken into account in the analysis. Particular focus was drawn on the explanation of the purpose of the "Recognized"-button (see Chapter 3.2), since the time measured with this interface element is one of our main dependent variables.

A between subject design was applied so that 59 participants navigated with the schematic map and 59 with the detailed map. A balanced distribution of men an women among the two prototypes was ensured.

If no more questions aroused, the test run was started. The destination was not communicated to avoid that participants could find their way without the wayfinding aid using their cognitive maps. This procedure was also applied e.g. by Münzer and Stahl [24]. The participants did not receive any additional help. If someone took a wrong turn at a decision point this was recorded as an error and the person was informed and guided back to the route. At the destination the eye tracker recording was stopped and the device was packed away. Finally, the participants had to fill in the NASA-TLX questionnaire, which measures the workload of a task (see [12]). Questions concerning the usefulness of the maps and the landmarks were asked in addition.
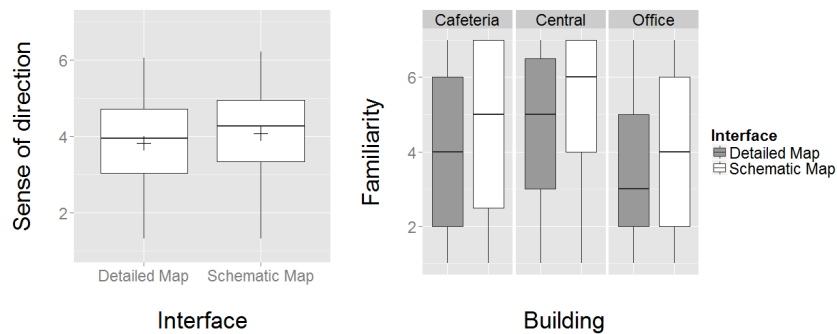
■ **Figure 4** Sense of direction and familiarity with the test route of the participants measured with a 7-point Likert scale with higher values representing higher sense of direction respectively familiarity. The familiarity is split according to the three building of the test route.

After the experiment the eye tracking data was annotated with the help of the software of the manufacturer (SMI BeGaze 3.7). The eye tracker recording shows a video of the environment and the detected gaze (see Figure 3, right). This data was mapped on so-called reference views, which represent areas that are of interest for analysis. For this study, we annotated all gazes on the screen and distinguished the different areas "arrow", "text instruction" and "map". Moreover, we annotated gazes on the referenced landmark and other objects in the environment using "placeholder elements", i.e. labeled boxes (see Figure 3, left). Gazes only needed for locomotion, such as looking at the floor, were not considered.

3.4 Participants and Devices

The test sample consisted of 118 participants (60 male), most of them being students (110 participants). Their mean age was 23.36 years (SD = 5.00; minimum: 18 years, maximum: 54 years). Due to the eye tracker used, persons who need glasses were not allowed to participate. The subjects were very familiar with the use of smart phones (mean 5.78 on a 7-point Likert scale with higher values representing higher familiarity; SD = 1.78), but rather unfamiliar with pedestrian navigation systems (mean 3.25; SD= 1,67). Their sense of direction measured with the questionnaire of Münzer and Hölscher [23] did not differ between the two groups (t(115) = 1,105; p = 0,272; see Figure 4, left). The familiarity with the test route was distributed heterogeneously (see Figure 4, right), but did not differ amongst the two test groups taking into account the mean for all three buildings (Z = 1.03; p = 0.301).

The eye tracker used was the "SMI Eye-Tracking Glasses 2", which records gazes with a 60 Hz rate. In order to increase the accuracy of the detected gazes on the screen, the navigation prototypes were displayed on a Samsung Galaxy Tab S (screen diagonal = 26.7 cm). Other studies showed that gazes on smart phones cannot be recognized with a satisfying accuracy (see e.g. [26]).

4 Results

In the next sections the results are reported. We analyzed whether differences in wayfinding behaviour could be observed if users navigated with the schematic or the detailed map depictions. The first step was used to explain the procedure and is therefore not taken into account. In addition, the last step is not considered. Here, the instructions only referred to the destination, thus not including to a specific landmark.

4.1   Errors and Time Needed for Orientation

The experimenter took a note every time a participant took a wrong turn at a decision point. This variable is almost equally distributed among the two map versions. With each interface 10 participants had problems to find their way without additional help (see Table 2). One person made an error twice with the detailed map. Another schematic map user got lost three times. Thus, a fist insight is that wayfinding problems are not mainly caused by the map used. In fact, the situation seems to have a high impact on wayfinding performance. Especially at the beginning of the navigation process, users had problems. The initial wayfinding phase is very demanding [30], and therefore future work should address how to assist persons at this point.

The interface only displayed one landmark per step. At step 14, however, the instructions referred to tow landmarks ("Go up the stairs and through the door"). The schema was disrupted, because the two participants of the conducted prestudy with the interface stated that the steps would otherwise be too small. This led to a relatively high amount of errors, showing that only referring to one landmark could enhance navigation efficiency as described in [2].

In total the wayfinders needed 7 minutes and 46 seconds to accomplish the task. The total navigation time does not differ significantly among the two map groups ($Z = -1.58$; $p = 0.115$). However, the time needed to orient oneself determined by the Recognized-button (see Chapter 3.2) differs significantly ($Z = -2.50$; $p = 0.013$): schematic maps users have slightly lower mean values (mean schematic map = 7.23; mean detailed map = 7.39). A detailed overview for every step is depicted in Figure 5. The plot also shows – like the navigation errors already indicated – that the orientation time highly depends on the wayfinding situation and the map material only marginally influences performance. Especially the visibility of the landmark has a high impact. Due to the fact that the participants had to decide themselves when they want to see the next instruction by clicking on the Next-button, some landmarks were (not yet) visible, as some wayfinders demanded the instruction earlier than expected. For steps 6 and 10 this led to longer orientation times. An example is given in Figure 6.

4.2   Gaze Behaviour

To gain a deeper understanding how the map material was perceived, we analysed the gaze behaviour of the participants during the wayfinding task. We distinguished gazes on the three areas of the screen, i.e. the map, the arrow and the text instruction. In addition, gazes on the referenced landmarks and the environment were annotated. The results show no differences concerning the gaze duration on the environment, neither on the landmark ($Z = -0.711$; $p = 0.477$), nor the environment ($Z = -0.027$; $p = 0.979$). No differences were found concerning the arrow displayed in the upper left corner of the screen ($Z = -1.07$; $p = 0.286$). This element was hardly consulted at all (mean detailed map = 0.17 seconds; mean schematic = 0.15 seconds) and is therefore probably not necessarily needed.

■ **Table 2** Number of participants who took a wrong turn separated by route parts. Steps that are not reported did not led to errors. Two participants got lost more than one time.

Map/Step	2	3	5	6	7	8	10	11	14	15
Schematic	1	2	1	0	1	2	1	1	3	0
Detailed	3	1	0	2	0	1	0	0	2	2
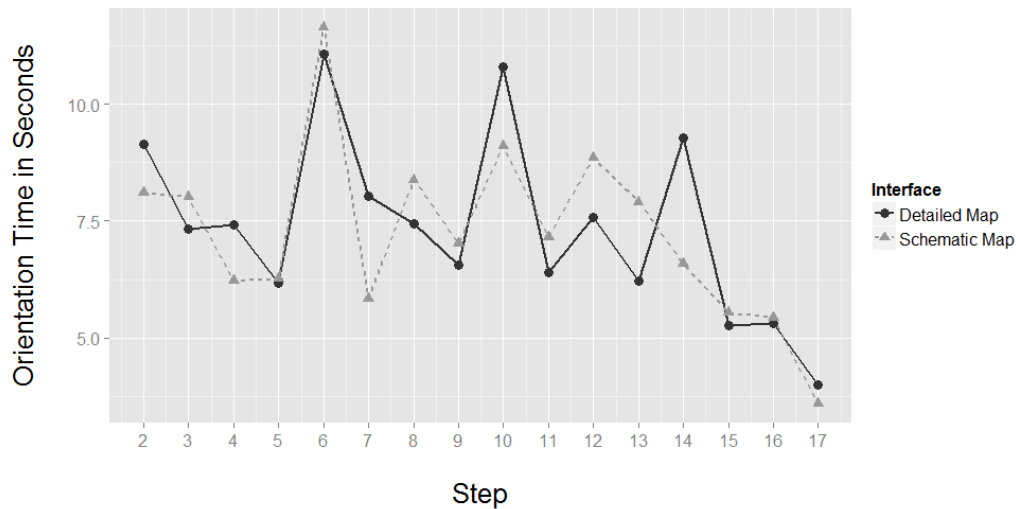
Figure 5 Time needed for orientation divided by step.
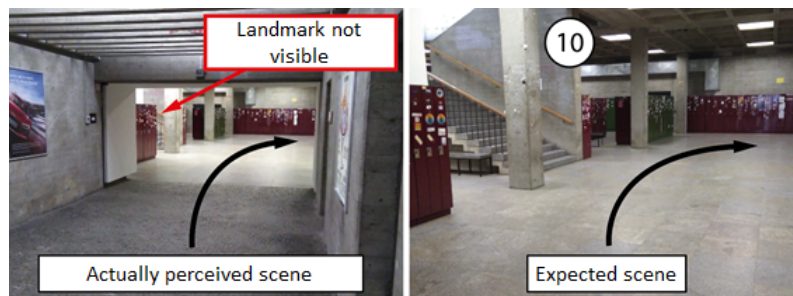
Figure 6 Bad visibility of landmark leads to longer orientation times. The example shows step 10. The instruction says "Turn right at the stairs".

However, significant differences were found for the fixation duration on the text element (t(116) = -2.24; p = 0.027; r = 0.201). Participants navigating with the detailed map looked longer at this element (mean detailed map = 1.74 seconds; mean schematic = 1.41 seconds), even though this element did not differ among the groups. Moreover, detailed map users spent significantly more visual attention on the map element (Z = -3.67; p < 0.001; r = 0.337, see Figure 7 for a detailed overview). Since this depiction contains more visual information, this was expectable. Nevertheless, the results of the orientation times show that the additional information does not lead to faster self-localisation or better performance. In fact, longer fixations on the screen seem to reflect orientation problems: the recognition time and fixation time on the map correlate significantly ($r_{sDetailed}$ = 0.360; $r_{sSchematic}$ = 0.230; $p_{Detailed\&Schematic}$ < 0.001).

We additionally analysed the revisits on the screen. Revisits show how often users look (once again) at the interface after they looked at the environment. This is therefore a hint for disorientation and searching for additional information. This variable also revealed significant differences (Z = -6.75; p < 0.001). Wayfinders using the detailed map return more often to the map (mean detailed map = 3.74, mean schematic = 3.01).

The gazes also reflect orientation problems due to a bad visibility of landmarks. For the critical steps 6 and 10 the fixation duration on the map (see Figure 7) and the revisits
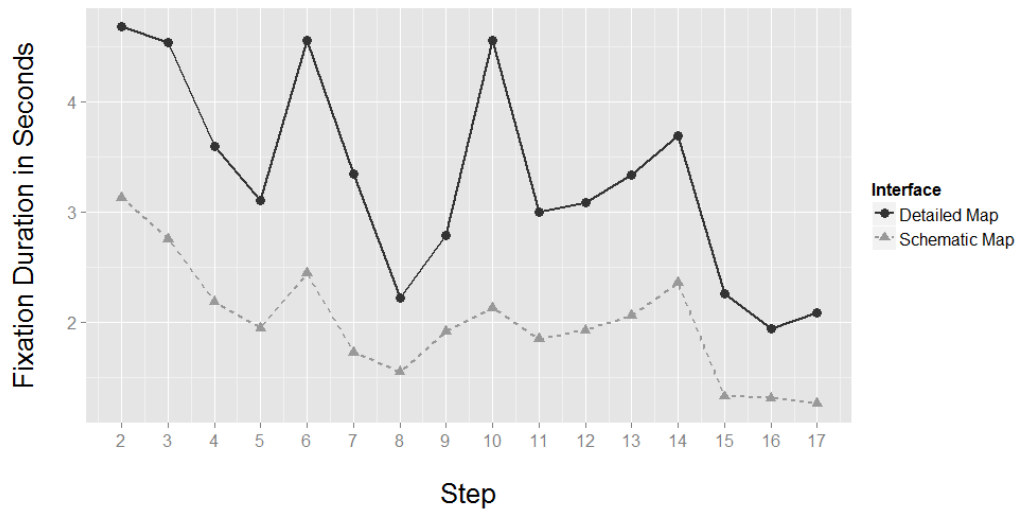
Figure 7 Fixations on the map divided by step.

are higher compared to the rest of the steps. In these situations users seem to search for information on the map in order to solve their orientation problems. Moreover, the initial orientation problems at the beginning of the route are also observable.

4.3   Questionnaire

At the end of the experiment participants were asked to fill in the NASA-TLX, which is a questionnaire often used in the context of the evaluation of pedestrian navigation systems to assess the workload of a task (see e.g. [10, 28]). Neither the overall workload ($Z = 0.238$; $p = 0.812$), nor the separate dimensions of the questionnaire differed among the two map groups ($p > 0.05$). Figure 8 shows an overview of the results. It also shows that the dimension "mental demand" was rated as the most challenging. The high amount of "outliers" at the performance dimension also show that several users stated that they were not confident with their own performance.
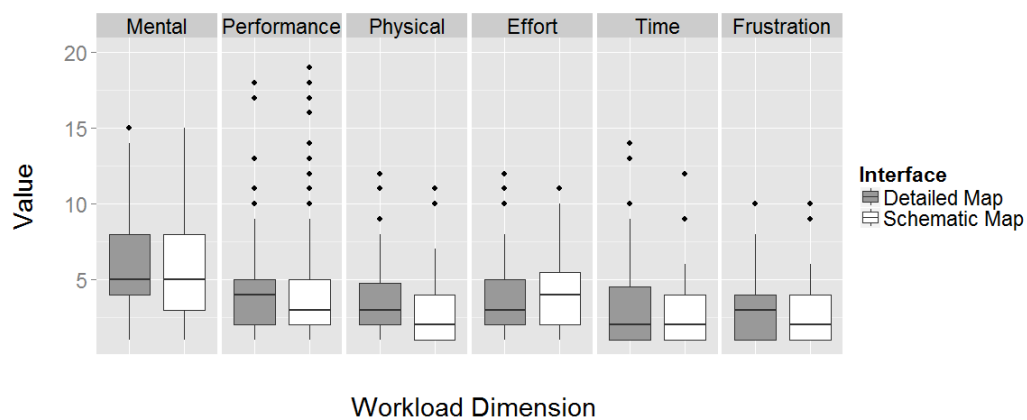
Figure 8 Workload dimensions measured with the NASA-TLX.

In addition, the wayfinders were asked to rate the helpfulness of the map depiction on a 7-point Likert scale. The ratings did not differ for the two map designs (Z = -0.165; p = 0.869) and most of the participants were very content with the visualisation (mean detailed map = 5.78, mean schematic map = 5.64). Furthermore, the participants had to rate if the chosen landmarks supported orientation. Even though several participants had observable orientation problems with some landmarks that were not always visible, the majority of the users found that the landmarks were very helpful (mean schematic map = 6.24, mean detailed map = 6.27 on a 7-point Likert scale).

5 Conclusion and Future Work

In this paper we presented a study that examined the wayfinding behaviour of pedestrians navigating with a schematic map compared to participants that additionally used a detailed map displayed on a mobile navigation system for indoor environments. In addition, both map designs depicted landmarks to support orientation. The landmarks were collected during a prestudy. The results show that wayfinders spend more visual attention on the text instruction and especially on the map material if they use the detailed map while meanwhile the orientation time slightly increases. In addition, they look more often again at the screen after they looked at the objects in the environment. The self-reported workload and satisfaction with the displayed elements did not differ for the two map designs. Wayfinders who used a schematic map did e.g. not endure higher mental demand or had to invest more effort to solve the task.

All in all, the test persons were able to solve the wayfinding task more quickly with the schematic map. Therefore, we conclude that this depiction leads to more efficient orientation and is an efficient means of displaying wayfinding information. However, the absolute differences of the time needed for orientation are very small. The detailed map is therefore still a good navigation aid. The main advantage of a schematic map is that the wayfinders have more "free" visual resources that could be used to explore the environment. Nevertheless, it is still an open question whether this map material allows the wayfinders to focus more on the environment, since we could not find any differences concerning the gazes in the real world. This question could be addressed in future work, e.g. by examining whether the users can draw more detailed sketch maps after the task. Moreover, our study clearly showed that the landmarks used to guide the pedestrians and especially their visibility have a great impact on wayfinding efficiency and interactions with the display. Therefore, the main future research direction should focus on means to adequately convey landmark information under varying conditions of the environment. If e.g. a reliable indoor localisation technique is available, the system could only refer to landmarks that are certainly visible at the current position of the wayfinder. In this context, it is also important to analyse at which decision points a map is actually needed. Most of the participants spent approximately the same amount of visual attention on the text instruction and the map. At some points during the route a text instruction could be enough to solve the wayfinding task. On the other hand, a map provides more information about the environment and could help to maintain orientation at complex decision points. Another critical situation is the initial orientation phase. In this context, displaying e.g. an overview map could help the pedestrians to gain a better understanding of the route they have to take.

Furthermore, future research should examine whether our findings are applicable for outdoor environments. In this context, e.g. Bienk et al. [3] showed that the preferred depiction depends on the sense of direction of the wayfinders, whereby persons with a "good"

sense of directions benefit from more abstract depictions. In a previous study we could also show that bad-oriented users profit from detailed map material [25]. Consequently, the influence of the characteristics of the wayfinder and the navigation situations should be examined in more detail.

References

1 Gary L. Allen. Cognitive Abilities in the Service of Wayfinding: A Functional Approach. *The Professional Geographer*, 51(4):555–561, 1999.

2 Christina Bauer, Manuel Müller, and Bernd Ludwig. Indoor Pedestrian Navigation Systems: Is More Than One Landmark Needed for Efficient Self-localization? In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, MUM '16, pages 75–79, New York, NY, USA, 2016. ACM.

3 Stefan Bienk, Markus Kattenbeck, Bernd Ludwig, Manuel Müller, and Christina Ohm. I want to view it my way: Interfaces to mobile maps should adapt to the user's orientation skills. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*, pages 34:1–34:9, New York, NY, USA, 2013. ACM.

4 Anders Bouwer, Frank Nack, and Abdallah El Ali. Lost in Navigation: Evaluating a Mobile Map App for a Fair. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, pages 173–180, New York, NY, USA, 2012. ACM.

5 Andreas Butz, Jörg Baus, Antonio Krüger, and Marco Lohse. A Hybrid Indoor Navigation System. In *Proceedings of the 6th International Conference on Intelligent User Interfaces*, IUI '01, pages 25–32, New York, NY, USA, 2001. ACM.

6 Andreas Ettlinger, Hans-Berndt Neuner, and Thomas Burgess. Smartphone Sensor-Based Orientation Determination for Indoor-Navigation. In Georg Gartner and Haosheng Huang, editors, *Progress in Location-Based Services*, pages 49–68. Springer, Cham, 2017.

7 Conrad Franke and Jürgen Schweikart. Mental representation of landmarks on maps: Investigating cartographic visualization methods with eye tracking technology. *Spatial Cognition & Computation*, 17(1-2):20–38, 2017.

8 Christian Freksa. Spatial aspects of task-specific wayfinding maps. In *Visual and spatial reasoning in design*, pages 15–32, Sydney, 1999. Key Centre of Design Computing and Cognition, University of Sydney.

9 Georg Gartner and Verena Radoczky. About the Role of Cartographic Presentation for Wayfinding. In Emmanuel Stefanakis, Michael P. Peterson, Costas Armenakis, and Vasilis Delis, editors, *Geographic Hypermedia: Concepts and Systems*, pages 381–398, Berlin, Heidelberg, 2006. Springer.

10 Ioannis Giannopoulos. *Supporting Wayfinding Through Mobile Gaze-Based Interaction*. Dissertation, ETH-Zürich, 2016.

11 Reginald G Golledge. Human wayfinding and cognitive maps. In Reginald G Golledge, editor, *Wayfinding behavior: Cognitive mapping and other spatial processes*, pages 5–45, 1999.

12 Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Peter A. Hancock and Najmedin Meshkati, editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 139–183. North-Holland, 1988.

13 Haosheng Huang and Georg Gartner. A Survey of Mobile Indoor Navigation Systems. In Georg Gartner and Felix Ortag, editors, *Cartography in Central and Eastern Europe: CEE 2009*, pages 305–319, Berlin, Heidelberg, 2010. Springer.

14 Markus Kattenbeck. *Empirically Measuring Salience of Objects for Use in Pedestrian Navigation*. Dissertation, Universität Regensburg, 2016.

15 Peter Kiefer, Ioannis Giannopoulos, and Martin Raubal. Where Am I? Investigating Map Matching During Self-Localization With Mobile Eye Tracking in an Urban Environment. *Transactions in GIS*, 18(5):660–686, 2014.

16 Alexander Klippel, Kai-Florian Richter, and Stefan Hansen. Wayfinding Choreme Maps. In Stéphane Bres and Robert Laurini, editors, *Proceedings of the 8th International Conference on Visual Information and Information Systems*, pages 94–108. Springer, Berlin, Heidelberg, 2006.

17 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.

18 Alexandra Lorenz, Cornelia Thierbach, Nina Baur, and Thomas H. Kolbe. Map design aspects, route complexity, or social background? Factors influencing user satisfaction with indoor navigation maps. *Cartography and Geographic Information Science*, 40(3):201–209, 2013.

19 Kristin L. Lovelace, Mary Hegarty, and Daniel R. Montello. Elements of Good Route Directions in Familiar and Unfamiliar Environments. In Christian Freksa and David M. Mark, editors, *Proceedings of the International Conference on Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, pages 65–82, Berlin, Heidelberg, 1999. Springer.

20 David M. Mark, Christian Freksa, Stephen C. Hirtle, Robert Lloyd, and Barbara Tversky. Cognitive models of geographical space. *International Journal of Geographical Information Science*, 13(8):747–774, 1999.

21 Daniel R. Montello. Navigation. In A. Miyake and P. Shah, editors, *The Cambridge Handbook of Visuospatial Thinking*, pages 257–294. Cambridge University Press, New York, NY, US, 2005.

22 Daniel R Montello and Corina Sas. Human factors of wayfinding in navigation. In W. Karwowski, editor, *International Encyclopedia of Ergonomics and Human Factors*, pages 2003–2008. CRC Press/Taylor & Francis, London, England, 2006.

23 Stefan Münzer and Christoph Hölscher. Development and validation of a self-report measure of environmental spatial strategies. *Diagnostica*, 57(3):111–125, 2011.

24 Stefan Münzer and Christoph Stahl. Learning Routes from Visualizations for Indoor Wayfinding: Presentation Modes and Individual Differences. *Spatial Cognition & Computation*, 11(4):281–312, 2011.

25 Christina Ohm, Stefan Bienk, Markus Kattenbeck, Bernd Ludwig, and Manuel Müller. Towards interfaces of mobile pedestrian navigation systems adapted to the user's orientation skills. *Pervasive and Mobile Computing*, 26:121–134, 2016.

26 Christina Ohm, Manuel Müller, and Bernd Ludwig. Evaluating indoor pedestrian navigation interfaces using mobile eye tracking. *Spatial Cognition & Computation*, 17(1-2):89–120, 2017.

27 Christina Ohm, Manuel Müller, Bernd Ludwig, and Stefan Bienk. Where is the landmark? Eye tracking studies in large-scale indoor environments. *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research co-located with the 8th International Conference on Geographic Information Science (GIScience 2014)*, 2014.

28 Timo Partala and Miikka Salminen. User Experience of Photorealistic Urban Pedestrian Navigation. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 204–207, New York, NY, USA, 2012. ACM.

29 Arto Puikkonen, Ari-Heikki Sarjanoja, Merja Haveri, Jussi Huhtala, and Jonna Häkkilä. Towards Designing Better Maps for Indoor Navigation: Experiences from a Case Study. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia*, MUM '09, pages 16:1–16:4, New York, NY, USA, 2009. ACM.

30 Verena Schnitzler, Ioannis Giannopoulos, Christoph Hölscher, and Iva Barisic. The Interplay of Pedestrian Navigation, Wayfinding Devices, and Environmental Features in Indoor Settings. In *Proceedings of the 9th Biennial ACM Symposium on Eye Tracking Research & Applications*, ETRA '16, pages 85–93, New York, NY, USA, 2016. ACM.

31 Nanja J. J. M. Smets, Guido M. te Brake, Mark A. Neerincx, and Jasper Lindenberg. Effects of Mobile Map Orientation and Tactile Feedback on Navigation Speed and Situation Awareness. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '08, pages 73–80, New York, NY, USA, 2008. ACM.

32   Barbara Tversky and Paul U. Lee. Pictorial and Verbal Tools for Conveying Routes. In
Christian Freksa and David M. Mark, editors, *Proceedings of the International Conference
on Spatial Information Theory. Cognitive and Computational Foundations of Geographic
Information Science*, pages 51–64, Berlin, Heidelberg, 1999. Springer.

33   P. Viaene, A. Vanclooster, K. Ooms, and P. De Maeyer. Thinking aloud in search of landmark
characteristics in an indoor environment. In *Ubiquitous Positioning Indoor Navigation and
Location Based Service (UPINLBS)*, pages 103–110, 2014.

34   Dirk Wenig, Alexander Steenbergen, Johannes Schöning, Brent Hecht, and Rainer Malaka.
ScrollingHome: Bringing Image-based Indoor Navigation to Smartwatches. In *Proceedings of
the 18th International Conference on Human-Computer Interaction with Mobile Devices and
Services*, MobileHCI '16, pages 400–406, New York, NY, USA, 2016. ACM.

35   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.

Dyadic Route Planning and Navigation in Collaborative Wayfinding

Crystal J. Bae (ORCID)
University of California, Santa Barbara, USA
cbae@ucsb.edu

Daniel R. Montello
University of California, Santa Barbara, USA
montello@ucsb.edu

Abstract

The great majority of work in spatial cognition has taken an individual approach to the study of wayfinding, isolating the planning and decision-making process of a single navigating entity. The study we present here expands our understanding of human navigation as it unfolds in a social context, common to real-world scenarios. We investigate pedestrian navigation by pairs of people (dyads) in an unfamiliar, real-world environment. Participants collaborated on a task to plan and enact a route between a given origin and destination. Each dyad had to devise and agree upon a route to take using a paper map of the environment, and was then taken to the environment and asked to navigate to the destination from memory alone. We video-recorded and tracked the dyad as they interacted during both planning and navigation. Our results examine explanations for successful route planning and sources of uncertainty in navigation. This includes differences between situated and prospective planning – participants often modify their route-following on the fly based on unexpected challenges. We also investigate strategies of social role-taking (leading and following) within dyads.

1 Introduction

Wayfinding as a cognitive process is necessarily situated in a social world, whether someone is explicitly traveling with other people, following route directions, using socially-created signs or maps, or following the physical traces of others to direct their travel [5]. Wayfinding consists of all the acts associated with planning the way between an origin and a destination, including remembering routes, recognizing landmarks, and orienting oneself within the environment [8]. People often need to find their way through an environment while co-present with other people, making decisions jointly.

The majority of prior work in spatial cognition has taken an individual approach to the study of wayfinding, isolating the planning and decision-making process of a single person, animal, or robot as the unit of study. We know for instance how a single person looks at a map and plans a route [21], and we know about choice behaviors at decision points along a route [23, 34]. But limited prior research supports how navigation may work for pairs or groups of people, such as strategies that contribute to success in these interactions or the unique challenges and behavioral effects facing multiple people wayfinding together. Our

project has implications for the design of both physical and digital navigation aids, expanding what we know about the information needs of multiple people working in conjunction on a wayfinding task.

2 Prior Work

Wayfinding is a complex act that depends on our mental representations of physical environmental spaces that we experience directly or indirectly [24]; in many cases, we learn environments both directly, by traveling through the environment, and indirectly, via symbolic media such as maps or language. Both the planning and enactment of a route through navigation are important wayfinding processes that are often social. Analyzing navigation behavior "in the wild" [15] within a more realistic social setting, versus in a controlled laboratory or virtual setting, is important for the construct and ecological validity of our work. While a wealth of information informs our understanding of wayfinding, a small but growing body of work forms the basis for our knowledge about social interaction in human wayfinding and navigation.

2.1 Route Planning

People commonly give route directions by providing a sequentially-structured set of instructions used to identify a route from an origin to a destination [35]. Investigations into direction-giving allow us to define the structure of a complete set of route instructions, what is at the core of a route plan, and what makes for more or less effective route directions [1,6,22]. The establishment of common ground discussed in the route directions literature is also important to people working together in planning and in active navigation.

Studies by Hölscher et al. [14] show a profound difference between situated and prospective planning, wherein participants often modify their route-following *in situ*. The authors also highlight differences between the construction of routes for oneself and for others: Effective routes planned for others are simple (with few direction changes) and contain distinctive landmarks; those planned for oneself are attractive, fast, direct, and not too busy. Additionally, route plans intended for others include more detailed descriptions to establish common ground between planner and addressee. This suggests that verbalized plans of intended behavior often differ from real-world behavior, highlighting a need for more situated studies. Our work looks at these behaviors in planning and during real-time navigation with a partner.

2.2 Navigation

Navigation along a route, as opposed to only planning a route, presents contextual challenges of remembering the route plan, understanding correspondence of the plan to the experienced physical environment, self-localizing and maintaining one's orientation, judging distances, and (often) coordinating one's spatial knowledge with others. Spatial disorientation and misorientation are common problems threatening any navigation activity. According to Montello [25], geographic *disorientation* occurs when people believe they are unsure of their location or heading or which way to go to reach a destination (what people mean by explicitly expressing they "are lost"). When people are geographically *misoriented*, in contrast, they are objectively not where they think they are or are not going the correct way towards the destination, regardless of their awareness.

Environmental factors like low visibility, poor signage, and outdated maps often present real-world challenges to orientation and wayfinding. Fortunately, people have many available strategies to overcome being lost, such as moving in a specific direction, sampling routes

from a location, and backtracking [13]. However, the way individuals employ these strategies may only partially inform strategies at the group level. For groups, social factors could either cause problems like disagreement between navigational partners, or could provide valuable aid in dealing with unexpected problems. We look at wayfinding challenges as well as strategies enacted by people at the dyad level.

2.3   Group Navigation

There is recent enthusiasm around the social dimensions of wayfinding [5], though not traditionally explored by spatial cognition researchers. One distinctive example was Hutchins' work on "cognition in the wild," [15] which studied the navigation of a U.S. Naval crew as socially-distributed cognition, situated in the real world, rather than as an independent mental act. Hutchins proposed that group cognition in humans may have qualitatively different properties than individual cognition. This provides support for the ecological validity of conducting such a study in the real-world versus in a lab or virtual environment.

One important finding from He et al. [11] is that better navigators appear to adjust their route directions to the navigational ability of their partner. In their study on route direction-giving and -receiving by pairs using mobile phones for communication, they found that participants with a better sense of direction were better equipped to adjust how they provided navigational instructions. They were able to do so both because they stored more information about the environment they had traversed, and because they were more attuned to their partner's informational needs. Their study shows that flexibility in social coordination between members of a dyad may help overcome the disadvantages of being a poor individual navigator. Pairs perform differently than individuals not only due to differences in their spatial abilities but also because of their interpersonal route communication. Our work builds on this using pairs of people working synchronously in a wayfinding task to explore how people communicate when navigating together.

Our study uses the *dyad* as the unit of analysis, a pair of individuals who work together toward a shared goal. The dyad is considered the simplest-sized social group. Simmel's work on social geometry states that as each individual person is added to a group, different social behavioral dynamics emerge, such as a triad's tendency to act more as a dyad plus an individual, and a four-person group to divide into two dyads [33]. Specific to dyads, Reilly et al. [28] characterized the social roles adopted within pairs during navigation. These roles include, but are not limited to, roles such as leader and follower, or independent versus collaborative participants. We use this as a starting point to look at differences in how dyads act more or less collaboratively during both planning and navigation.

2.4   Social Interaction Analysis

The close investigation of social interaction that we employ in this project is *Conversation Analysis* (CA). A key feature of this approach [9, 31, 32] is its concern with conversational talk as it unfolds within a socially-shared context. CA as applied to situated navigation gives us methods of understanding how the project of wayfinding is constructed and maintained in real time (e.g. [10]). When multiple people navigate together, they must orient themselves with regards to the physical environment as well as coordinate their spatial knowledge to establish a shared reality within which they can work [27].

Many behavioral studies are predesigned to record certain expected behaviors, wherein the topics of observation are determined beforehand (i.e., they are top-down). On the other hand, CA gives us a bottom-up opportunity to learn the strategies people employ to form

common ground, for example using place labels to establish shared understanding [30]. By examining the talk immediately following an action, we observe how participants jointly understand and respond to what is being done. In the case of navigation, a person may see their partner pause at a juncture and use that opening to provide instruction. We see that they read the pause as an expression of uncertainty and as a point of potential intervention. People clearly orient themselves not only to the spatial task of navigation but also to the social task of shared understanding.

We demonstrate the value of incorporating the methods of CA to understand social actions and strategies relevant to wayfinding. By observing both route planning and in-person navigation, we compare how navigational plans are proposed ahead of time (prospectively) to how they are enacted in the physical environment (situatively). Close analysis of navigational performance by different dyads helps us explain how social interaction contributes to success or failure in solving wayfinding problems such as recovering from being lost. We focus on the issues of leadership, knowledge alignment, and personal characteristics.

3   Method

This work investigates route planning and navigation by dyads in a novel environment. Participants making up the dyads did not previously know each other and had little or no prior knowledge of the study site. To investigate both prospective co-planning of routes and situated co-navigation, the study consisted of two phases: (1) the planning of a route between an origin and destination in a nearby neighborhood, done in a separate lab room, and (2) the subsequent navigation of the route within the environment. We integrate the conceptual and methodological research traditions of geography and sociology, which generally apply group-level analyses, and psychology, which conventionally examines the individual.

3.1   Research Questions

The research questions we address are:
1. How do differences in sense of direction and personality among individual navigators relate to dyadic route planning and travel, examined both as overall characteristics of dyads and as differences between dyad members?
2. Do dyads' prospective planned routes through a novel environment differ from their routes as enacted *in situ*, and if so, how?
3. How do dyads coordinate their knowledge and behavior in a real-world environment to navigate efficiently, such as by adopting social roles within the dyad?

3.2   Participants

A total of 30 pairs of people (60 individuals) were recruited from a subject pool of university students enrolled in introductory Geography classes. However, as these courses fulfill several general requirements, very few students in the subject pool were Geography majors. The average age of participants was 19.5 years old ($SD = 2.1$), which is representative of our subject pool. So that our results would not involve any effects of prior social role-taking, we tested pairs who did not previously know each other. We assessed prior familiarity by asking participants about it at the start of the study session. In 27 of the dyads, participants first met as part of participating in this study; in 3 dyads, members had briefly met in a classroom context, but none considered themselves more than acquaintances. Each dyad was tested at a separate time (i.e. not concurrently).

3.3 Individual Difference Measures

We summarize the wide differences in peoples' individual abilities [16] in terms of three factors important to our research agenda: sense of direction, personality, and gender. We examine whether patterns of social interaction and wayfinding differ as a function of the dyads' overall levels of the factors, or as a function of the relative match or mismatch of these factors between members of the dyads.

Sense of Direction (SOD). Directly relevant to real-world navigational ability is "sense of direction" (SOD), the ability to locate and orient oneself with respect to an environmental space. We assessed SOD with the Santa Barbara Sense of Direction Scale (SBSOD [12]), which asks people to rate their agreement with a variety of navigation-related statements, such as "I can usually remember a new route after I have traveled it only once" and "I have trouble understanding directions." Agreement is expressed on a Likert scale from 1 (strongly agree) to 7 (strongly disagree), with positively worded statements reverse-coded so that a higher score indicates a better reported sense of direction. A summary of our participants' scores on the SBSOD scale are presented in Table 1.

▪ **Table 1** Means on SBSOD and Big Five Inventory for Individual Dyad Members ($N = 60$).

Measures	All Members [Range]	Females ($N = 43$)	Males ($N = 17$)
SBSOD	3.9 [1.6–6.6]	3.8	4.2
Extraversion	3.3 [1.5–5.0]	3.3	3.4
Agreeableness	4.0 [2.3–5.0]	4.2	3.8
Conscientiousness	3.6 [1.2–4.8]	3.6	3.4
Neuroticism	2.8 [1.4–4.6]	2.9	2.6
Openness	3.5 [2.1–5.0]	3.5	3.6

Personality. Personality may account for some of the differences in social interaction style, engagement with novel environments, and leadership. Prior work has attempted to delineate the complex relationship between personality factors and spatially-relevant measures such as SOD, starting with Bryant's seminal work [2, 4]. We assessed personality using the "Big Five" Inventory (BFI), a widely-used measure for a modern framework of personality factors [17, 18]. The Big Five factors are widely used and accepted, based on decades of research [7], and include the dimensions of *Extraversion*, *Agreeableness*, *Conscientiousness*, *Neuroticism*, and *Openness to Experience*. Of these, Extraversion has been most consistently shown to correlate with leadership behavior, followed by Conscientiousness and Openness [19]. Respondents express their level of agreement with 44 statements on a 5-point Likert scale. For a given dimension, scores range from the lowest score of 1.0 to the highest of 5.0. A summary of our participants' scores on each dimension is presented in Table 1.

Gender. Gender has been shown to have a reliable relationship with aspects of spatial ability and style, including survey-based over route-based navigation [3, 20, 26]. Comparison across gender pairings therefore has the potential to capture considerable variation in spatial performance and strategy and in social interaction and role-taking. Scores on the SBSOD measure and the BFI measures of personality, grouped by gender, are shown above in Table 1. Dyads were fairly evenly distributed between female-female ($N = 15$), and female-male ($N = 13$) pairs. Unfortunately, we had very few male-male ($N = 2$) pairs, typical for the gender breakdown in our subject pool.

3.4   Materials

Test Neighborhood.   The study site is a residential suburban neighborhood approximately 2.5 km from campus (see Figure 1). Although there is public access, the neighborhood has only two entrances (to the north and west) and a number of traffic control measures (lower speeds and speed bumps), so it is not conducive to through-traffic. The layout is complex enough to pose a moderate level of wayfinding challenge, with a mostly circular street structure, smaller streets and cul-de-sacs branching off of the main access, and a central open space with interior footpaths. There is little elevation change throughout, so no locations provide visual access to the entire layout. This suburban neighborhood differs from a typical urban environment in that it has minimal visual differentiation in the form of landmarks and no regular street grid pattern. It differs from a more rural environment in that there are no long-distance vistas available within the neighborhood. We recognize that our results may be limited to this type of environment, leaving room to expand this line of research to a variety of environmental forms.

We selected a neighborhood that our pool of participants would likely be unfamiliar with, to ensure no advantage on the task based on prior knowledge. At the beginning of the study session, participants rated their prior familiarity with this neighborhood while looking at an overview map of the larger region. All participants included in the study rated their prior familiarity with the test neighborhood as either "very unfamiliar" or "unfamiliar," which meant that most had never previously been inside the neighborhood; those that had were further questioned to ensure this knowledge was minimal.

Map for Route Task.   The planning phase involved a paper map of the study area, which is shown scaled-down in Figure 1. We created this map by selecting a custom area using the InkAtlas tool[1] from OpenStreetMap[2] base map data, including street, footpath, bike path, and building features, and editing it in Adobe Illustrator to include task instructions, a map key, and origin and destination locations for the task.

3.5   Procedure

The individual spatial ability and personality measures described above were administered using an online or pen-and-paper based questionnaire at sign-up. The main data on route planning and navigation were collected in-person as follows:

Prospective Planning.   The two members of a dyad met independently at the lab. They were told they would work together to plan a shortest-path pedestrian route between a given origin and destination in a neighborhood near campus, and that afterwards, they would be taken to the neighborhood to walk their route. They were given the paper map shown in Figure 1 with the start and destination locations clearly marked. Participants were instructed to remember their planned route, as they would not have use of the map itself during their walk. Each dyad was given 10′ (10 minutes) to complete the task, including both deciding upon their route and committing it to memory. We video-recorded dyads' interaction during the planning process.

---

[1] `https://inkatlas.com`
[2] Map data copyrighted OpenStreetMap contributors and available from `https://www.openstreetmap.org`
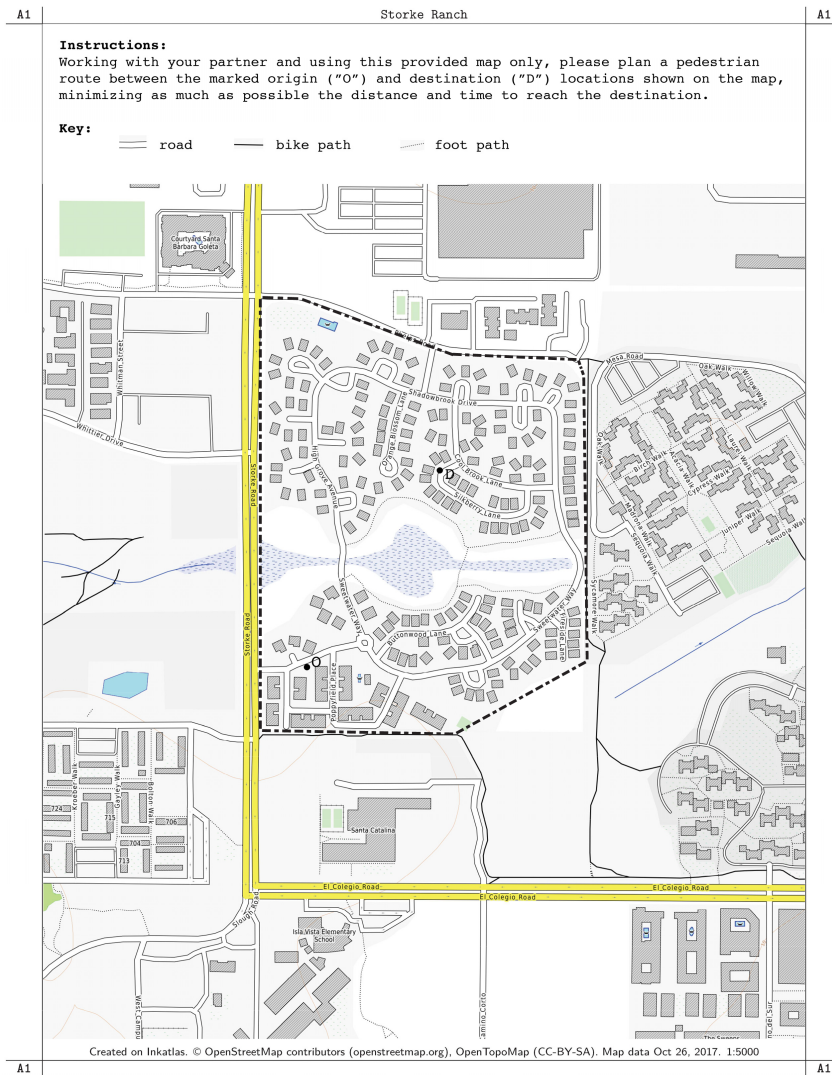
Figure 1 Map for planning with task instructions, marked origin and destination points, and key. The dashed line (not present on maps shown to dyads) shows the extent of the test neighborhood.

After planning, each member was separately asked to produce a drawing of the route ("route sketch") on a copy of the same base map and verbally describe the route they had planned with their partner. This was video-recorded for comparison within each pair (level of agreement within the dyad) and with the route as enacted by the dyad in the next phase (prospective versus situated navigation). Once the pair completed these route sketch and verbal description tasks, they were driven by the researcher to the start location for the situated navigation.

Situated Navigation. The navigation phase took place immediately following the planning phase, at the route origin in the study neighborhood. Dyads were instructed to work with their partner to walk to the destination, minimizing the time and distance to reach the destination as best they could. Importantly, they were told they did not have to take the same route as planned in the first phase. Each participant wore a chest-mounted video

camera (GoPro Hero 3+, a lightweight camera typically used for action sports) that recorded their speech, some of their hand gestures, and their approximate views. The researcher additionally observed, GPS-tracked, and video-recorded dyads using a handheld camcorder, but did not assist the dyads in any way to wayfind (i.e., gave no advice).

This phase of the study stopped either when the dyad reached and identified the destination successfully, unsuccessfully identified the destination point on three attempts (went to the wrong destination), or exceeded the maximum time allotted (30′). We counted as an attempt when both members of the dyad identified to the researcher that they believed they were standing at the destination. The researcher then reported whether they had correctly identified the destination, and if not, how many attempts they had remaining. After this phase, the researcher walked the participants to a nearby location within the study neighborhood to individually complete a follow-up questionnaire noting their leadership, following, or collaboration during the task; any deviations from the planned route; and any other unexpected occurrences during navigation.

4   Results and Discussion

We present overall task success for the dyads in the navigation task, relating navigational performance to difference measures for personality and spatial ability. Next, we summarize the effects of route selection and dyads' correspondence between their planned and enacted routes. We then look more closely at the enactment of leadership within dyads, and examine a specific case of dealing with uncertainty during decision-making.

4.1   Navigational Performance

We use both time and distance as a measure of navigational performance on this task, as dyads were asked to minimize both when navigating to the destination location. Time was highly correlated with distance traveled, $r = .94$, $p < .001$, for all dyads. Generally, those dyads who took more time in navigation were those who walked further, but this is not a perfect correlation due to slight differences in time spent pausing and in walking speed. Our initial measure of success was whether dyads navigated correctly to the destination location within three attempts and 30 minutes (30′). However, only one dyad failed to reach the destination within three attempts, and even they made all 3 attempts within 30′. This means 29 of 30 dyads reached the destination within three attempts. Of those who eventually found the destination, 26 dyads (87%) correctly reached and identified the destination on their *first* attempt.

Given the high eventual success rate, we distinguish the dyads who correctly reached the destination on the first attempt as "successful" and those who did not (including the dyad that never succeeded) as "failed." All 4 failed dyads were female-female pairs. The average navigation time by the successful dyads ($N = 26$) was 9′ 48″ ($SD = 4′ 05″$), the shortest lasting 5′ 10″ and the longest 22′ 55″. In contrast, the failed dyads ($N = 4$) took on average 22′ 28″ total, but averaged 14′ 06″ to their first (incorrect) attempt.[3] Successful dyads also traveled a shorter distance during navigation, averaging only 0.93 km, as compared to failed dyads, who averaged 1.28 km to their first attempt.

---

[3]   Subsequent comparisons involving time or distance traveled are based on time or distance to the first attempted destination, whether it was correct or incorrect.

Though each dyad was allowed 10 minutes for planning prior to navigation, none required the entire time. The average planning time across all dyads was only 3′ 25″, and time for planning ranged from 1′ 15″ to 7′ 40″. Successful dyads planned for longer (average of 3′ 32″) than did dyads who failed (2′ 41″). Of course, a sample size of 4 is too small for meaningful significance tests, but it is still suggestive to note that failed dyads took 4′ 18″ longer and walked 0.35 km further to reach their first attempted destination than did successful dyads, though successful dyads spent 51″ longer to plan.

4.2 Individual Differences

To assess sense of direction and personality for each dyad, we compared SBSOD scores and BFI scores on each dimension with navigational success using both the averages of members' individual scores and the differences between them (see Table 2 below). Again, for distance and time measures we use the distance and time to dyads' first attempt during navigation. We also report personality factors averaged from BFI scores for each dyad and their relation to distance and time to the first attempted destination. We found no reliable correlations between navigational time or distance and mean SBSOD or BFI personality factors.

The direction of correlation appeared to be positive for SBSOD, meaning higher SBSOD scores (suggesting better average sense of direction) may have related to travelling longer distances and taking more time to navigate (poorer performance). Comparing successful dyads to failed dyads, we find that mean SBSOD scores for successful dyads were actually 0.6 points poorer than for failed dyads. However, we would require a larger sample to verify these interpretations. This suggests the navigational advantage of better individual sense of direction scores may not apply at the dyad level due to the influence of social interaction. For instance, differences in personality may cause a dyad to have issues reaching consensus in their navigational decisions even where each individual may have a generally good sense of direction.

■ **Table 2** Means on SBSOD and Big Five compared with Navigational Performance.

Measures	All Members [Range]	Correlation with Distance	Correlation with Time
SBSOD	3.9 [1.6–6.6]	0.14	0.20
Extraversion	3.3 [1.5–5.0]	0.11	0.04
Agreeableness	4.0 [2.3–5.0]	-0.15	-0.13
Conscientiousness	3.6 [1.2–4.8]	0.15	0.15
Neuroticism	2.8 [1.4–4.6]	0.13	0.18
Openness	3.5 [2.1–5.0]	-0.12	-0.14

For further comparison, we assessed individual difference scores in terms of their mismatch between dyad members. We did this by calculating the absolute differences between members' scores on each measure (shown in Table 3 above). Although not quite reaching significance, dyads with greater differences in the members' SBSOD scores appeared to travel a shorter distance ($r = $-0.24, $p = 0.19$) and take less time ($r = $-0.29, $p = 0.12$) to their first attempt. This is consistent with the notion that having a member with better sense of direction helps the dyad navigate more effectively, but especially when the other member is content to accede decisions to the member with better sense of direction (suggested by work such as He et al. [11]).

Table 3 Difference Scores on SBSOD and Big Five compared with Navigational Performance.

Measures	All Members [Range]	Correlation with Distance	Correlation with Time
SBSOD	1.3 [0.2–3.7]	-0.24	-0.29
Extraversion	1.0 [0.1–3.5]	0.33	0.32
Agreeableness	0.6 [0.0–2.0]	0.10	0.09
Conscientiousness	0.8 [0.0–2.4]	-0.24	-0.20
Neuroticism	0.9 [0.0–2.0]	-0.06	-0.13
Openness	0.7 [0.1–1.9]	-0.14	-0.11

For personality, we found marginally significant correlations between difference in Extraversion and navigational performance ($r = 0.33$, $p = 0.07$ for distance and $r = 0.32$, $p = 0.09$ for time). That is, dyads with greater difference in members' Extraversion tended to travel longer and take more time navigating. We speculate that this could relate to leadership conflicts in groups with differing Extraversion; we examine leadership below. Differences in dyad members' personality scores on the other dimensions did not appear to correlate with performance. This points to our need to further investigate strategies used by dyads in planning and navigation that could contribute to success.

4.3   Adherence to Route Plans

We analyzed route plans as drawn and described by dyads and found high agreement within pairs. Most dyads ($N = 23$) agreed completely on their route plan, with each person reporting the same route as their partner in the individual descriptions of the route via the route sketches and verbal descriptions. In the 7 cases where they drew or described different routes, those routes had only a slight divergence (such as taking the first turn rather than the second onto the same street). In 3 cases, dyads prospectively planned a main route and an alternate route, and both members reported the two routes.

Figure 2 Five most popular route plans.

Figure 3 Overlay of all enacted routes.

A map displaying the five most commonly-planned routes by the dyads in this study is presented in Figure 2. These plans were compiled from the route sketches and checked against the video-recorded descriptions. Route plans not shown were minor variations on those shown, and were described by only 1 or 2 dyads in the study. Labels given to the planned routes are Route A ($N = 12$, shown in blue) which goes all the way around on the main road, Route B ($N = 7$, in green) which takes the footpath, Route C ($N = 7$, in yellow)[4] which takes the footpath and anticipates the shortcut, Route D ($N = 5$, in orange) which takes the footpath and passes by the shortcutting opportunity, and Route E ($N = 4$, in red) which plans a shortcut through a place where it is not possible.[5]

Dyads were instructed to take the best possible route to reach the destination location and not bound to follow their originally planned route. They therefore had the option of taking alternate routes or shortcuts but were not primed by the researcher to look for them. To measure the match between planned and enacted routes, we compare dyads' descriptions of routes during the planning phase with their recorded tracks of routes walked in the navigation phase. We processed minor noise in the GPS tracks by snapping the tracks to the road and path network using ArcGIS Desktop 10.6, while retaining any backtracking or significant divergence by comparing the tracks with the video recordings. In cases where the tracks were of poor quality or failed to record properly, routes were traced by hand based on the video recording only.

An overlay of all traveled paths by dyads during the navigation phase is shown in Figure 3 above. Darker colored lines represent segments that more dyads walked on; lighter colored lines are less-traveled paths. The most popular routes included the northern segment of the footpath and the main road running counter-clockwise through the neighborhood. Therefore, spatial strategies in this study appeared to sort into two main groups, those dyads taking the footpath and those following the main road.

To compare actual traveled distance to distance of the planned route, we computed a ratio of the distance of the route taken divided by the distance of the planned route:[6]

$$Distance \ Ratio = Distance \ of \ Enacted \ Route \ / \ Distance \ of \ Planned \ Route$$

With this ratio, 0.5 represents a dyad who walked only half as far as they had planned, such as by taking a shortcut; 1.0 represents a perfect match, where the dyad walked the same distance as the planned route (though not necessarily following the same route); 2.0 represents a dyad who walked twice as far as planned; and so on. The resulting ratios ranged from 0.67 to 4.33, with an average of 1.34 ($SD = .75$); this mean is significantly longer than 1.0, $t(29) = 2.49$, $p < .01$. Dyads thus walked longer overall on the enacted route than they had planned to walk, with one walking a distance over four times as long.

From participant responses to the follow-up questionnaire, we find that many were conscious of deviation from their original plan. In half the dyads ($N = 15$), one or both members mentioned taking a different path. Their explanations attribute these deviations to a variety of causes, which we categorized as "lost", "alternate", or "shortcut". In order of declining frequency, dyads explained deviations as due to: unexpected problems (such as disorientation, turning the wrong way, or overshooting), taking a planned alternate route based on decisions during active navigation, or recognizing and taking a shortcut to the destination. Table 4 gives examples of these explanations.

---

[4] This is the shortest possible (legal) route.
[5] Numbers do not sum to 30, as some dyads reported two alternate plans.
[6] In cases where the dyad decided on and reported more than one route option, the distances of those planned routes were averaged.

■ **Table 4** Question was posed as "Did you and/or your partner take a path that was different from your planned route in any way? Describe if so.".

Coded Explanation	Count of Dyads	Example Explanation
lost	8	"Yes, we weren't sure about a few of the turns and overshot them so we had to backtrack."
alternate	4	"We had 2 paths planned out. We found out that the plan A doesn't work, so we took the plan B."
shortcut	3	"Yes, instead of going all the way up the footpath we discovered a shortcut."

Overlap between Planned and Enacted Routes.   To further compare prospectively-planned routes to routes enacted during navigation, we defined route overlap using the recorded routes and route plans as coded in our GIS. For each dyad, we extracted the overlapping segments between the enacted and planned routes using the ArcGIS Intersect tool. We then calculated route overlap by dividing the total distance of the overlapping segments by the distance of the route as actually walked by the dyad:[7]

Route Overlap = Distance of Overlapping Segments / Distance of Enacted Route

In cases where dyads took the route they planned without any deviations, planned and enacted routes completely overlap (100%); in cases where dyads took completely different routes, overlap is 0%. In our study, percentage route overlap ranged from 100.0% to 11.9%; the average across all dyads was 69.1% ($SD = 32.4\%$). One third of all dyads ($N = 10$) followed their route exactly as planned and reported with 100.0% overlap. Route overlap correlated negatively with time to first attempt, $r(28) = -0.59$, $p < .001$, and with distance to first attempt, $r(28) = -0.48$, $p < .01$, suggesting that dyads reached their first attempted destination more quickly and directly if they more closely followed their original plan. Navigational performance therefore differed not only in total time and distance of travel, but also in terms of directness (as a result of more or less adherence to route plan).

Route Selection Strategy.   The particular route selected during the planning phase appears to be the strongest predictor of whether or not dyads successfully reached the destination without getting lost. The most common route choice, Route A (refer back to Figure 2), involved taking the main road counter-clockwise through the neighborhood and included the fewest number of turns. Correspondingly, the dyads who planned this route were more likely to closely follow it ($N = 12$, average 89.0% overlap) than were dyads who planned other routes ($N = 18$, average 55.8% overlap); they were also more likely to follow the route exactly without going off course (9 of 12 dyads). There were no gender differences between those who took this route versus other routes.

Review of the video recordings made during planning show that some, but not all, dyads explicitly decided to take a route with fewer turns because it was easier to remember and held less risk of getting lost. We think this points to the influence of route simplicity on navigational success. More complex routes have more turns to remember (or misremember), making them inherently more difficult to follow in a task that did not allow much opportunity

---

[7] Where two different routes were described by dyads after planning (such as the case above where the dyad "had 2 paths planned out"), the planned route more closely matching the enacted route was used to derive the overlapping segment.

to rehearse the planned route. Additionally, with more decision points to recognize, there is greater chance of travelers missing a cue in the environment while navigating *in situ*. We are performing more in-depth coding to characterize the nature of how different types of route plans were assessed relative to one another by dyads during planning.

4.4 Social Leadership and Decision-Making

In their follow-up questionnaire, individuals were asked (separately) to state who acted more as the navigational leader during the task. Of the 30 dyads, 18 agreed that "neither was clearly leading more," 5 agreed that "one was leading more," and in the remaining 7, the two members disagreed about leadership. In the 5 dyads where one member claimed they were leading more, the partner agreed. Interestingly, in all 7 of the "mismatch" cases, one person claimed "neither was clearly leading more" while their partner claimed that the first person was leading more. Perhaps people are hesitant to claim that they are leading more – that it is more socially acceptable to claim equal collaboration in the dyad rather than assert leadership (at least in the context of dyads whose members did not formerly know one another). This highlights a shortcoming of self-assessment; we follow this up below by coding conversational behaviors to assess leadership and following versus collaboration in navigation.

Individual and Dyad-Level Differences. At the dyad level, Conscientiousness significantly differed between the 12 groups with a stated leader and those 18 without ($t(17) = 2.17$, $p < .05$). Those dyads with a self-reported leader/follower dynamic had an overall lower score (0.4 less) on Conscientiousness than those who reported a collaborative dynamic, and tended to have a larger mismatch (1.0 difference) between dyad members' Conscientiousness scores. No other individual difference measure appeared significant. We also looked at individual-level leadership scores[8] in relation to SOD and personality, and found no significant relationships.

Although Conscientiousness was significantly related to leadership at the dyad level, individual scores on Conscientiousness did not correlate with a tendency for an individual to lead. To not see effects of Extraversion and possibly SOD seems surprising, since we expect these differences to relate to the emergence of a leader within a group; for instance, Judge et al. [19] showed Extraversion to significantly relate to leadership. The adoption of leadership roles is likely to be context-specific: navigational leadership may be more likely to express itself in a larger group, where there is more potential advantage to having a strong leader and potentially cumulative inefficiency in considering each members' suggestions.

Talk During Navigation. As another measure of leadership versus collaboration in navigation, we examined talk during navigation and calculated a ratio of navigationally-relevant talk between the two members of each dyad. In our exploratory assessment of the collected video-recordings, we noted that if one person made most of the wayfinding decisions, that person generally spoke more about the navigation than their partner, who affirmed or accepted their partner's suggestions. In dyads that looked to be more collaborative in their decision-making, we observed that this was more of an equal exchange, with both partners discussing their available options and neither "dominating" the conversation. To

---

[8] Scores were assigned wherein stronger reports suggesting a given member was leading corresponded with a higher score: "0" for those who reported their partner led, "1" for each if both agreed neither was leading more, and "2" for those who claimed to lead or were identified by their partner as leading.

quantify these observations with the transcribed video recordings of the navigation task, we summarized the total time each member contributed navigationally-relevant talk to the conversation. This provides a high-level view of comparative participation in the wayfinding, as another indicator of leadership.

Using all transcribed talk for each pair during navigation, we filtered out only the navigationally-relevant talk. Navigationally-relevant talk included all talk relevant to decision-making, identifying landmarks, remembering the route plan, or commentary on the current physical environment or the route. We excluded "getting to know you" talk, casual chat about interests, classwork, or weather, and anything that did not appear to contribute to wayfinding. We calculated a "talk ratio" equal to the duration of relevant talk by the partner who contributed less to the wayfinding divided by the duration of relevant talk by the partner who contributed more. This gave us values between 0 and 1 for each dyad, where values closer to 1 would describe more equal durations of relevant talk between the members, a value of 0.5 would represent a case in which one member talked twice as much as their partner, and values closer to 0 would describe situations where one member dominated most of the relevant conversation. For our 30 dyads, these values averaged 0.71 and ranged from a pair in which one person talked almost four times as much about the navigation as their partner (0.28) to a pair which was virtually equal (0.97).

Talk ratios corresponded with self-reported leadership, where dyads with a clear leader averaged a talk ratio of 0.65 and those who did not report a clear leader averaged 0.76. These means were significantly different, $t(21) = 2.1$, $p < .05$, meaning those who did not report leadership within the dyad did indeed have more equal durations of relevant talk than those with a reported leader. Especially in dyads with less collaborative talk ratios, the reported leader was consistently the one who talked at greater length over the entire task, with most navigation talk consisting of directives by the leader and often simple clarifications or affirmations by the follower. See Appendix A, Examples 1 and 2, for two short excerpts from dyads with a low talk ratio that demonstrate this. This suggests either that navigational leadership in a dyad is indeed associated with a less equal ratio of relevant talk, or that a less collaborative talk ratio gives the impression of leadership even where there is none.

Uncertainty in Decision-Making.   As an example of the detailed analysis of interaction we are undertaking, we share the case of one "failed" dyad who took three attempts to reach the correct destination point. This pair (Dyad 2) was made up of two female participants with similar SOD and personality scores. They planned to take Route B (0.9 km long), but ended up walking more than twice what they had planned (1.9 km). The dyad traveled for $19' \ 35''$ to reach what they first thought was their destination, and traveled a total of $23' \ 45''$ to finally reach the correct destination.

Dyad 2 encountered trouble throughout the task, not only with remembering the route plan but also in managing their en-route decision-making. Though one member (A) reported afterward that her partner was leading, neither displayed strong leadership during navigation. The ratio of relevant talk between the two members was close to equal over the entire navigation (0.84), and from the coded video recordings it appears that neither person was predominantly leading. The decision-making in the dyad was mostly collaborative, where each attempted to establish consensus with her partner before proceeding. The following excerpt portion demonstrates, however, that this was often difficult (see explanations of coding symbols and the entire excerpt in Appendix A, Example 3):

```
01 B: we were supposed to make a le-
02 A: LEFT, huh? a LE^FT? [wait (.) THA^T way?]
03 B:                    [that's why I said through the-] through the-
04    that's why I SAID I was like, through the THI^NG (0.1)
05 A: HH.h are you SU^RE?
06 B: NO I dunno^ ((shields eyes, looks in same direction as partner))
07 A: NO we go... ((turns, brings hands together)) kay on the map it was...
```

Only three minutes in, the dyad is already off course and disoriented. Revisiting what went wrong (line 1), B suggests they should have gone left instead of right to find the footpath. When A questions B further (line 5) with "Are you sure?" her partner backs down with "No, I don't know," and they proceed to review their ongoing navigation from the beginning (line 7, continued in Appendix A, Example 3). After further review of their plan using the available communicative resources of speech, gesture, and body positioning, B shows impatience with their inability to figure out what went wrong. B interrupts with "All right, let's just see, whatever. We'll just go through the streets," (lines 51–52) and begins to walk away. This prompts A to follow along even while asking, "Well, what are the pathways supposed to look like?", something B would have no reason to know any better than her. Much later (not included in the excerpt), B attempts to use a stick to draw their plan in the soil; however, this is quickly abandoned as it does not appear to aid in mutual understanding.

This dyad's attempt to work collaboratively during navigation was handicapped by a 'divide and conquer' strategy for memorizing their route and by studying only the streets relevant to their plan. During planning, they focused exclusively on two street names that cued important turns on their route. When they encountered trouble committing both names to memory, they decided each person would focus on only one of the street names. Once in the actual environment, the dyad struggled with correspondence between their plan and those unstudied options. The dyad demonstrated uncertainty throughout the entire task and explicitly stated this in the follow-up questionnaire. One stated, "Most of the navigation I felt lost, at one point I knew for sure we were on the right path, but then [became] confused when I didn't see the way we planned to take." They also acknowledged disagreement at several points during the task, which is corroborated in the analysis of their decision-making. Our detailed example suggests that disagreement and miscommunication between dyad members presents a source of uncertainty and suboptimal navigational performance.

5   Summary and Conclusions

Our study makes a contribution to the empirical evaluation of wayfinding by explicitly considering social interaction. We present a comprehensive account of dyads working together to plan a navigational route through a new environment, then working together within a situated context to enact the planned (and sometimes misremembered) route. This scenario exemplifies *strong synchronous* social wayfinding in the framework by Dalton et al. [5], as dyad members directly interact with one another to make wayfinding decisions and accompany one another during the task in real time. This is one of the few empirical studies to date that has done so; others that have looked at strong synchronous wayfinding have generally used remote methods of communication [11, 28]. As stated above, there exists a body of work that looks at situations of asynchronous wayfinding (such as providing route directions [6]), but we also believe complementary work that would support this research agenda would focus on weak wayfinding scenarios, in which people follow social cues indirectly provided by others.

In our results, navigational performance did not seem to relate to gender pairings within dyads, though we recognize that the small number of male-male pairs in this study is a shortcoming. We do believe that future studies focused on comparing different gender pairings would make a valuable addition to the literature. Performance also did not relate much to the average sense of direction or personality scores of the dyads, suggesting more in-depth interactional analysis is necessary to determine the social contributions to successfully wayfinding in pairs. Difference scores on sense of direction and personality measures between the dyad members showed modest and marginal relationships with performance: Dyads with greater difference in members' SBSOD navigated more quickly and for less distance, while dyads with greater difference in Extraversion navigated more slowly and for more distance.

Most dyads walked further than planned, demonstrating challenges of accurately enacting a route plan *in situ*. The specific overlap between planned and enacted routes was nearly 70% and correlated strongly with time and distance walked to first attempt. In general, dyads who chose the simplest possible route to the destination were most likely to accurately walk the planned route. The cost associated with getting off-track when taking a complicated route reduced the advantage of planning a shorter route. Although selecting the simplest route to walk appeared to play a role in navigational success, dyads had various spatial and social strategies at their disposal to deal with uncertainties.

Self-reported leadership within dyads did not relate to individual Extraversion, but dyads with higher Conscientiousness did tend to work more collaboratively during navigation. However, as self-report falls short of assessing actual leadership verbalizations and other behaviors, we looked at individual members' contributions to navigation during the task as a "talk ratio" and found that navigation-related conversation was indeed more one-sided in dyads with a reported leader-follower dynamic.

However, detailed Conversation Analytic (CA) investigations into dyadic decision-making processes during navigation will help us illuminate the strategies employed in successful versus unsuccessful navigation. We plan to follow up with this in a future paper. As justification for this, we presented a detailed transcript of the interactions between the members of one dyad, suggesting that disagreements and miscommunications are an important source of uncertainty and contribute to poor navigational performance. Studying social navigation elucidates how people share knowledge in a task-oriented setting specific to wayfinding, establish social roles like leadership within groups, and deal with common challenges.

Our study focuses on dyads without prior familiarity with one another, but we acknowledge that social interactive aspects relevant to navigation may be more pronounced in familiar dyads. Ongoing work will present a similar navigational scenario but recruit dyads with existing social relationships. Whether accurate or not, existing notions about others' relevant navigational abilities should plainly influence group interaction. Established dyads are likely to have established patterns of interaction relevant to the domain of navigation and are likely to feel comfortable enacting those roles, so leadership may be more clearly expressed in such a comparison. We also plan to make a direct comparison between dyadic and individual navigators, to help elucidate differences in planning and dealing with uncertainty when one is working alone versus with others. Additionally, we will use the video-recorded interactions to produce a large collection of specific conversational actions relating to navigational leadership across dyads to form a generalizable account of how this type of leadership is enacted socially.

Our interest in studying navigation from a social interaction perspective is related to how people use and communicate spatial knowledge in a task-oriented setting, establish social roles within groups, and interact with one another to deal with common challenges such as uncertainty at decision points. The sources of potential uncertainty in wayfinding are many,

and further study applying these methods will allow us to investigate how people deal with these uncertainties in direct, situated interaction. Real-world navigation is a phenomenon that occurs within social contexts, often explicitly in conjunction with other people. Our work highlights the rich nature of observing people working together towards a navigational goal.

References

1   Gary L. Allen. Principles and practices for communicating route knowledge. *Applied Cognitive Psychology*, 14(4):333–359, 2000.

2   Kendall J. Bryant. Personality correlates of sense of direction and geographic orientation. *Journal of Personality and Social Psychology*, 43(6):1318–1324, 1982.

3   Emanuele Coluccia and Giorgia Louse. Gender differences in spatial orientation: A review. *Journal of Environmental Psychology*, 24(3):329–340, 2004.

4   David M. Condon, Joshua Wilt, Cheryl Ann Cohen, William Revelle, Mary Hegarty, and David H. Uttal. Sense of direction: General factor saturation and associations with the Big-Five traits. *Personality and Individual Differences*, 86:38–43, 2015.

5   Ruth C. Dalton, Christoph Hölscher, and Daniel R. Montello. Wayfinding as a social activity. *Frontiers in Psychology*, 10:1–14, 2019.

6   Michel Denis, Francesca Pazzaglia, Cesare Cornoldi, and Laura Bertolo. Spatial discourse and navigation: An analysis of route directions in the city of Venice. *Applied Cognitive Psychology*, 13(2):145–174, 1999.

7   J. M. Digman. Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1):417–440, 1990.

8   Reginald G. Golledge and Robert J. Stimson. *Spatial behavior: A geographic perspective*. Guilford Press, 1997.

9   Charles Goodwin and John Heritage. Conversation analysis. *Annual Review of Anthropology*, 19:283–307, 1990.

10   Pentti Haddington. Action and space: Navigation as a social and spatial task. In P. Auer, M. Hilpert, A. Stukenbrock, and B. Szmrecsanyi, editors, *Space in Language and Linguistics: Geographical, Interactional and Cognitive Perspectives*, pages 411–433. de Gruyter, Berlin, 2013.

11   Gengen He, Toru Ishikawa, and Makoto Takemiya. Collaborative navigation in an unfamiliar environment with people having different spatial aptitudes. *Spatial Cognition & Computation: An Interdisciplinary Journal*, 15(4):285–307, 2015.

12   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.

13   Kenneth A. Hill. *Lost person behavior*. National SAR Secretariat, Ottawa, Canada, 1998.

14   Christoph Hölscher, Thora Tenbrink, and Jan M. Wiener. Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition*, 121(2):228–247, 2011.

15   Edwin Hutchins. *Cognition in the wild*. The MIT Press, Cambridge, MA, 1995.

16   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(2):93–129, 2006.

17   Naumann L. P. John, O. P. and C. J. Soto. Paradigm shift to the integrative Big Five Trait taxonomy. *Handbook of Personality: Theory and Research*, page 114–158, 2008.

18   O. P. John and S. Srivastava. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research (2nd ed.), L. A. Pervin & O. P. John (Eds.), New York: Guilford Press*, pages 102–138, 1999.

19   Timothy A. Judge, Joyce E. Bono, Remus Ilies, and Megan W. Gerhardt. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4):765–780, 2002.

20   Carol A. Lawton and Janos Kallai. Gender differences in wayfinding strategies and anxiety about wayfinding : A cross-cultural comparison. *Sex Roles*, 47(9–10):389–401, 2002.

21   Amy K. Lobben. Tasks, strategies, and cognitive processes associated with navigational map reading: A review perspective. *The Professional Geographer*, 56(2):270–281, 2004.

22   Kristin Lovelace, Mary Hegarty, and Daniel R. Montello. Elements of good route directions in familiar and unfamiliar environments. In Christian Freksa and David M. Mark, editors, *COSIT '99, LNCS*, volume 1661, pages 65–82. Springer-Verlag Berlin Heidelberg, 1999.

23   Tobias Meilinger, Gerald Franz, and Heinrich H. Bülthoff. From isovists via mental representations to behaviour: First steps toward closing the causal chain. *Environment and Planning B: Planning and Design*, 39(1):48–62, 2012.

24   Daniel R. Montello. Navigation. In Priti Shah and Akira Miyake, editors, *The Cambridge Handbook of Visuospatial Thinking*, pages 257–294. Cambridge University Press, Cambridge, UK, 2005.

25   Daniel R. Montello. Cognition and spatial behavior. In *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, pages 1–20. John Wiley & Sons, Ltd, 2017.

26   Daniel R. Montello, Kristin L. Lovelace, Reginald G. Golledge, and Carole M. Self. Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, 89(3):515–534, 1999.

27   George Psathas. Some sequential structures in direction-giving. *Human Studies*, 9(2):231–246, 1986.

28   Derek Reilly, Bonnie Mackay, Carolyn Watters, and Kori Inkpen. Planners, navigators, and pragmatists: Collaborative wayfinding using a single mobile phone. *Personal and Ubiquitous Computing*, 13(4):321–329, 2009.

29   Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696–735, 1973.

30   Emanuel A. Schegloff. Notes on a conversational practice: Formulating place. In D. N. Sudnow, editor, *Studies in Social Interaction*, pages 75–119. MacMillan, The Free Press, New York, 1972.

31   Emanuel A. Schegloff. Conversation analysis and socially shared cognition. In L. B. Resnick, J. M. Levine, and S. D. Teasley, editors, *Perspectives on socially shared cognition*, pages 150–171. American Psychological Association, Washington, DC, 1991.

32   Emanuel A. Schegloff. Confirming allusions: Toward an empirical account of action. *American Journal of Sociology*, 102(1):161–216, 1996.

33   G. Simmel. *The sociology of Georg Simmel, Compiled and translated by K. H. Wolff*. Glencoe, IL: The Free Press, 1950.

34   Thora Tenbrink and Jan M. Wiener. Wayfinding strategies in behavior and language: A symmetric and interdisciplinary approach to cognitive processes. *Spatial Cognition*, 5:401–420, 2007.

35   Dieter Wunderlich and Rudolph Reinelt. How to get there from here. In Robert J. Jarvella and Wolfgang Klein, editors, *Speech, Place, and Action*, pages 183–201. John Wiley & Sons, Ltd, New York, 1982.

A   Navigational Transcript Excerpts

We follow basic conventions in Conversation Analysis, adapted from the guide by Sacks et al. [29]. This guide directs coders to spell transcribed utterances in a way that attempts to directly capture speech as produced rather than as properly spelled, aligns overlapping speech between two speakers [within brackets], uses colons to indicate the prolonging of a syllable, capitalizes louder speech, surrounds softer speech with °degree symbols°, and

represents upward inflections with ^. Gestures are described within ((double brackets)). Pauses lasting less than a tenth of a second are represented as (.); longer pauses are shown with the duration in tenths of a second in parentheses.

A.1   Dyad 24 Excerpt

Example from Dyad 24 (03′13″ to 03′21″), whose member A spoke 3.6 times as long about navigation than partner B (talk ratio = 0.28):

```
01 B: thi^s is...
02 A: or do you want me to che^ck like-
03 B: yeah, we can... check
04 A: yeah, we can check and then come back if we're not certain about it
```

A.2   Dyad 9 Excerpt

Example from Dyad 9 (00′14″ to 00′23″), whose member A spoke 2.4 times as long about navigation than partner B (talk ratio = 0.41):

```
01 B: what was the first street, Sweetwater?
02 A: ye::s:: I'm pretty sure it's this one
03 B: okay
04 A: this is the roundabout and we just go that way
05 B: okay
```

A.3   Dyad 2 Excerpt

Example from Dyad 2 (03′06″ to 04′24″):

```
01 B: we were supposed to make a le-
02 A: LEFT, huh? a LE^FT? [wait (.) THA^T way?]
03 B:                     [that's why I said through the-] through the-
04    that's why I SAID I was like, through the THI^NG (0.1)
05 A: HH.h are you SU^RE?
06 B: NO I dunno^ ((shields eyes, looks in same direction as partner))
07 A: NO we go... ((turns, brings hands together)) kay on the map it was...
08 B: ((turns around to face same direction as partner)) (0.4)
09    ah.hh (0.1)
10 A: °out of° Sweetwater...
11 B: yeah Sweetwater ((turns to face same way as partner))
12    and then there was a LOOP ((draws circle with finger, points forward))
13 A: and then you go
14    [you go around the loop] ((extends left arm with right arm to elbow))
15 B: [then after you barely ] wa^lk
16    yea^h we go arou^nd the LOOP
... 28 lines removed for space considerations ...
45 A: cuz we were supposed to go a- (0.6)
46 B: NO cuz if you go through tha-
47 A: it's either we go-
48    it's either we go tha^t way ((points straight out with left arm))
49    or we come this way and we wait for the... ((holds out right arm)) (0.3)
50    no cuz we were [supposed t-    ]
51 B:               [all right let's] just g- let's just-
52    let's just see, whatever (0.2) we'll just go through the streets
53 A: well, what- what are the pathways suppo-
54    [°walking pathways supposed to look like°   ]
```

```
55 B: [that's what I'm sayin like where are the p-] (0.8) pathway
56    (0.9) I don't know where the pathways were
57    (2.1)
58 A: I think they-
59    (0.5)
60 B: do you wanna go ba^ck?
61 A: Sweetwater... NO cuz if we woulda went tha^t way it woulda been
62    another stree::t
```

Assessing Spatial Information in Physical Environments

Vinicius M. Netto

Department of Urbanism, Universidade Federal Fluminense (UFF),
Rua Passo da Patria 156, Niteroi, Rio de Janeiro state, Brazil
http://www.socialfabric.city
vmnetto@id.uff.br

Edgardo Brigatti

Institute of Physics, Universidade Federal do Rio de Janeiro,
Av. Athos da Silveira Ramos, 149, Cidade Universitária, 21941-972, Rio de Janeiro, RJ, Brazil
edgardo@if.ufrj.br

Caio Cacholas

Programme of Graduate Studies, Universidade Federal Fluminense (UFF),
Rua Passo da Patria 156, Niteroi, Rio de Janeiro state, Brazil
caiocacholas@id.uff.br

Vinicius Gomes Aleixo

Programme of Graduate Studies, Universidade Federal Fluminense (UFF),
Rua Passo da Patria 156, Niteroi, Rio de Janeiro state, Brazil
viniciusaleixo@id.uff.br

Abstract

Many approaches have dealt with the hypothesis that the environment contain information, mostly focusing on how humans decode information from the environment in visual perception, navigation, and spatial decision-making. A question yet to be fully explored is how the built environment could encode forms of information in its own physical structures. This paper explores a new measure of spatial information, and applies it to twenty cities from different spatial cultures and regions of the world. Findings suggest that this methodology is able to identify similarities between cities, generating a classification scheme that opens up new questions about what we call "cultural hypothesis": the idea that spatial configurations find consistent differences between cultures and regions.

1 Introduction

One of the most interesting question in spatial information is *how the built environment could carry forms of information encoded in its own physical structures*. Of course the idea that cities can actually encode information is well known in urban theory. It is at the heart of Kevin Lynch's [12] pioneering work in 1960 on spatial elements guiding navigation in the environment, along with memory and representation, even though he did not quite use the term "information". In the 1980s, Amos Rapoport [16] (page 19) explicitly asserted, "physical elements of the environment do encode information that people decode". More recently, Haken and Portugali [5, 6] have seen information latent in street layouts and built form. The possibility that information could be encoded in physical structures is indeed very

interesting. If physical spaces materialise information, we could encode information in the built environment and decode it while living in it. Information materialised in physical space could be contextual resources useful to guide our actions. Research in spatial information has certainly dealt with this possibility focusing on how we decode information from the environment [22, 4].

A question yet to be fully explored is *how the built environment could encode forms of information in its own physical structures* in the first place, and *how such amounts of information could be measured empirically.* Approaches to information have adopted different measures to try to extract task relevant information [9, 23, 18]. In this paper, we further develop and apply a new approach [13] based on Shannon entropy to assess urban configurations as environmental information. With this aim in mind, we characterise the spatial information encoded in two-dimensional configurations of buildings, seen as the most elementary unit of the urban fabric. By analysing the cellular arrangement of buildings, we are able to assess the structures of urban blocks and streets. Information will be quantified measuring Shannon entropy [20, 21], operationally estimated by looking at the sequence of bits 1 and 0 representing built form cells and open space cells within sections of cities. Theoretically, this corresponds to analysing a 2D symbolic sequences of 1 and 0. In this context, information finds a very precise meaning: the entropy of the sequence, a measure of the surprise a source that produces the sequence causes in the observer [24]. In fact, physical arrangements characterised by higher levels of randomness, uncertainty or unpredictability are associated with high entropy. In contrast, the presence of regularities and patterns in urban structures will correspond to lower entropy, which means a higher predictability. Hypothetically, cities with ordered structures would help agents understand their environment, allowing them to make predictions about areas beyond their fields of visibility. Agents can make inferences, memorise layouts and navigate more easily from one place to another – say, grasping the pattern of blocks and intersections from local streets and inferring that some blocks away they will still have the same pattern (see [12, 7, 8, 1, 11] on legibility and intelligibility in urban structures; [15, 22, 4] on pattern recognition and spatial decision making).

Indeed, the layout of the environment encodes more information than two-dimensional configurations can express. However, we opted for an analytic approach able to sufficiently describe differences in information potentially encoded in urban built form – hence our reduction of urban form to cellular aggregations. We apply this approach to a number of empirical cases, namely 20 cities from three different regions of the world. We expect our measure to grasp "spatial signatures" of such cities, i.e. a measure able to point out differences and quantify similarities inherent to their spatial configurations. Thus we introduce a general concept of similarity between pairs of cities based on their informational content. This concept will allow to order our pool of cities and define a classification scheme which can help in verifying something that captures the imagination of many scholars: the possibility of finding consistent similarities in the configuration of cities from a same "spatial culture" or world region, along with consistent differences between cities from different cultures – what we call *"the cultural hypothesis"*.

2   Materials and methods

Our database includes 20 cities from the South of Europe, North America and Latin America. We selected cities based on (i) their importance for the region and country; (ii) a size compatible with our methodological requirements; (iii) availability of data. The first item has brought us to some well-known cities as emblematic cases.

The second aspect involved the selection of areas for the application of our measure. Selection of cities and sections was based on the identification of homogeneous areas, i.e. with spatial continuity of its urban fabric able to satisfy occupation rates close to 50%, which means avoiding large empty areas or rarefied patterns of urbanisation. These restrictions follow two critical considerations. The first is that our method is well fitted for estimating entropy for continuous urban areas. Second, and most importantly, it is interesting to decouple the analysis of city structures between small-scale, detailed and denser urban areas, and large-scale regional and peripheral urban areas. In fact, the two scales are different, and for this reason, they can be described using different methodologies. The first one is marked by specific features such as blocks and buildings, and the stratification of human interventions uniquely characterises resulting shapes and structures. The second scale is distinguished by sparse occupation, frequently with a scale-free character, where physical phenomena and constraints that geographical formations and barriers might play very relevant roles. In this work we will focus only on small scale, continuous urban areas.

The third criterion involved availability of spatial information on the configuration of cities. Many cities, particularly in Latin America, have incomplete information regarding buildings and their precise location, position and geometry (for instance, major cities like Lima and Bogota).

We extracted sections of cities from public map repository Google Maps API. We tested trade-offs between resolution and availability of data for distinct scales. We chose geographic areas of $9,000,000$ m^2, which were considered sufficient for representing the general spatial characteristics of dense urban areas regarding the configuration of buildings, urban blocks and street networks. Background picture bases of the selected cities were then prepared and exported in high resolution, filtering layers and converting entities representing buildings into solid raster cells. Images underwent a re-sizing process for 1000^2 cells and were converted to a monochrome system and then into a matrix of size 1000×1000 cells with binary numerical values.

We propose to assess Shannon entropy in the following cellular arrangements (Figure 1). Our approach uses a method generally applied for estimating the entropy of sequences of symbols encoded in one-dimensional strings [19]. For these data sets, the method consists of defining the block entropy of order n through

$$H_n = -\sum_k p_n(k) \log_2[p_n(k)], \tag{1}$$

where blocks are string segments of size n, and the sum runs over all the k possible n-blocks. Equation (1) corresponds to the Shannon entropy of the probability distribution $p_n(k)$. The Shannon entropy of the considered system [19, 10], which we indicate with h, is obtained from the following limit:

$$h = \lim_{n\to\infty} H_n/n, \tag{2}$$

which measures the average amount of randomness per symbol that persists after all correlations and constraints are taken into account. It was proven that the above limit exists for all spatial-translation invariant systems. More details about this approach can be found in [19, 10].

This method can be applied to our problem once we have defined the blocks for a two-dimensional matrix [3]. In this two-dimensional context, the most intuitive idea is to consider a block of size n as a square which contains n cells. To obtain the sequence of H_n also for

n values that do not correspond to squares, we considered blocks that interpolate perfect squares, as described in Figure 2. Note that there is no unique natural way to scan a 2D matrix. We tested our approach for different reasonable forms of constructing the blocks, and using different paths does not seem to strongly influence the estimation of H_n.

Equation 2 gives precisely the entropy for a theoretical infinite set of data. In real situations, where the data set is finite, our method estimates the probabilities of distinct arrangements of cells within blocks up to a certain size n, counting their frequencies, and then estimates the limit. Note that when working with two symbols, the estimation of H_n becomes not reasonable when $2^n \approx N$, where N is the number of elements in our data set [10]. Thus, in our case, this condition is verified for $n \approx 20$.

3 Results and discussion

We found empirically that, for all examined cases, the following ansatz provides an excellent fit: $a + b/n^c$. Even if we observed that the convergence is relatively slow, the fitted value of a gives a reasonable extrapolation of the Shannon Entropy h. As an example, we show the results for the estimation of H_n/n and the corresponding entropy h for the city of Los Angeles in Figure 3 (top left). Results for the estimation of the entropy h for the selected areas of the twenty sampled cities are shown in Table 1.
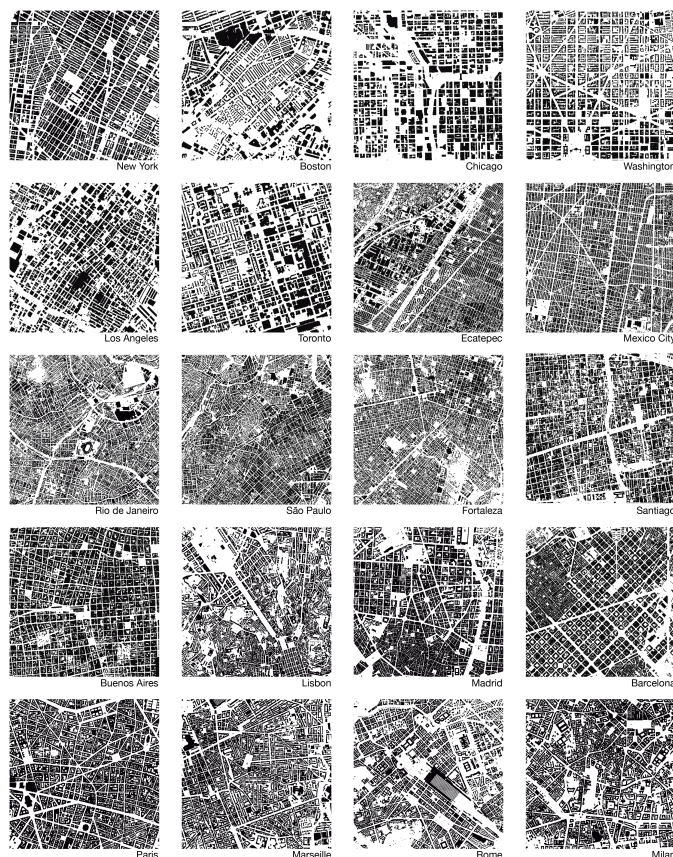
Figure 1 Spatial distributions in real cities ($9,000,000$ m^2 windows, 1000×1000 cells), extracted from Google Maps. These sections of emblematic cities are used to compute Shannon entropy.
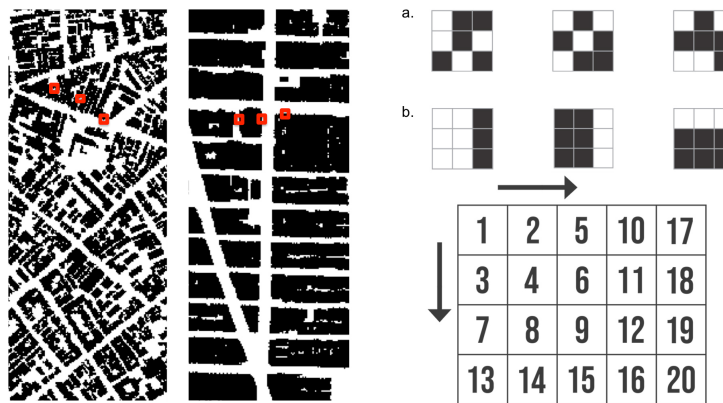
■ **Figure 2** Examples of blocks with nine cells are shown in red for selected areas in Rio and Manhattan, NY (left), and are amplified on the right. Blocks are constructed following the determined 1-D path represented on the bottom, right. Numbers indicate the order in which the cells are added to the block. The first block of size 1 corresponds to cell 1 and neighbouring cells are added in the corresponding order.

■ **Table 1** The estimated entropy h for selected areas of the sampled cities. The errors relative to the fitting procedure are of the order of 0.01.

Chicago	0.092	Paris	0.230
Washington	0.144	Madrid	0.232
Toronto	0.161	Marseille	0.250
LosAngeles	0.163	Barcelona	0.261
NewYork	0.167	Lisbon	0.261
Boston	0.171	Ecatepec	0.339
Buenos Aires	0.208	Mexico City	0.368
Santiago	0.216	Fortaleza	0.402
Roma	0.228	São Paulo	0.421
Milan	0.229	Rio de Janeiro	0.428

We proceeded to analyse these values in search of similarities and differences between the entropy signatures of the sampled cities. For clarity, we initially plotted the entropy values along a straight line (Figure 3, top right). Aiming to use these results for developing a classification scheme for our dataset, we performed a proximity network analysis based on the entropy values, to identify the possible formation of clusters of cities sharing similar values. Once we obtained the entropy h for all considered cities, we can quantify the levels of similarity defining a distance between cities i and j based on the values of h: $d_{ij} = |h_i - h_j|$. We created a matrix of distances for the twenty cities and then defined a network where cities are nodes, and edges (links between nodes i and j) are present only if the value of d_{ij} is smaller than a fixed threshold value of 0.03, which roughly corresponds to the 99% C.I. of the extrapolated values of h. The proximity network of cities can be seen in Figure 3.

We can clearly distinguish between five different, disconnected clusters. In the lower side of the cellular entropy spectrum, Chicago is the most ordered structure. It is the case to an extent that it appears on a class of its own, an isolated node with an entropy value h below 0.100. We have a strong cluster formed by other cities selected in North
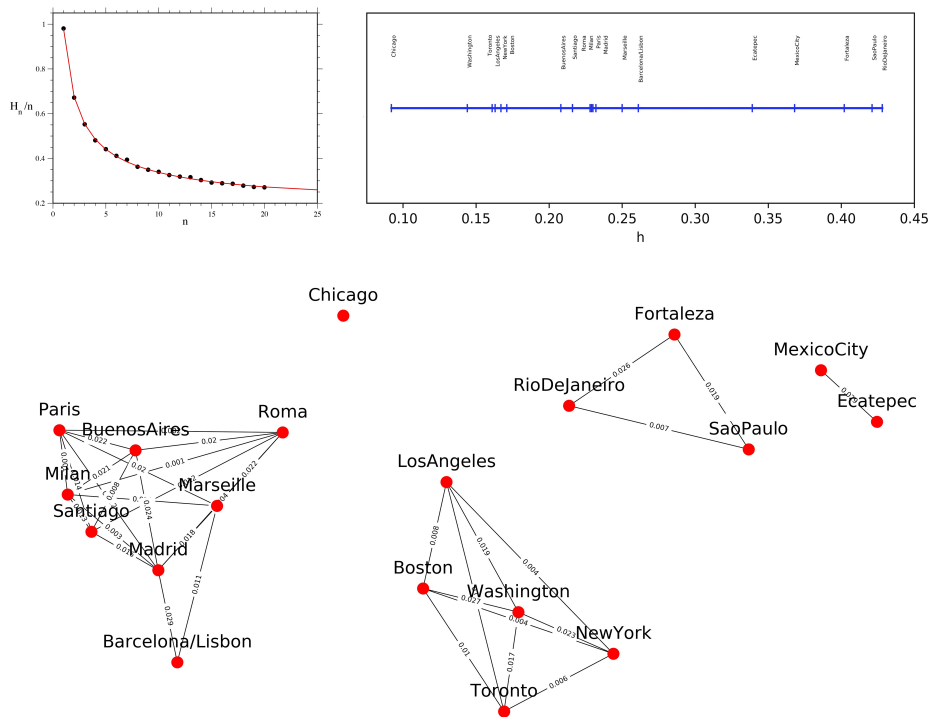
■ **Figure 3 top left**: An example of the estimated values of H_n/n for the city of Los Angeles. The continuous line represents the best fitting of our data using the function: $a + b/n^c$. The fitted values of a give an extrapolation of the Shannon Entropy h of the data set. All the analyzed cities present a very similar behaviour. **top right:** Estimated values of h for cities under analysis.
bottom: Proximity network of cities based on the value of h. The edge lengths are not proportional to the levels of proximity between entropy values. Distances are labelled over edges.

America (US/Canada), with h from 0.144 (Washington) to 0.171 (Boston). South American cities Buenos Aires and Santiago follow in a different cluster (entropy values above 0.200), along with European cities. Since other cities in Latin America have two clusters close in entropy levels (one with cities in Brazil and one with the cities in Mexico), this is an apparently surprising result – one that runs counter the cultural hypothesis of similarity in spatial signatures for cities within a same culture or region. This suggests that a single, general regionally-based explanation cannot respond to all cases. In fact, deeper historical contingencies including specific cultural factors might be forces at play. For instance, cities founded by Spanish colonizers in Latin America were often created as an orthogonal grid pattern, which were the case for Buenos Aires and Santiago. As these cities expanded, patchwork patterns were added around the central core's regular structure, adding entropy to the mix. Nevertheless, order in those spatialities is still felt today. Close to them, Southern European cities appear as a middle range cluster, with h from 0.228 (Rome) to Lisbon (0.261). Of course these cities have fully functional systems of urban blocks and street networks even though they are arranged as deformed grids, probably due to centuries of urbanization since the Middle Ages and ancient foundations. Finally, we have the cluster on the higher side of the cellular entropy spectrum, with South American cities in Brazil (Rio de Janeiro, Sao Paulo and Fortaleza). This finding is consistent with Medeiro's [2] analysis of Brazilian cities

as in average the most topologically fragmented among 164 cities in the world, and with Boeing's [1] finding about São Paulo having one of the three lower grid-order values in a sample with a hundred cities (which did not include Rio).

To sum up, in this paper, we attempted to understand how spatial information is encoded in built environments, looking into their physical configurations as cellular aggregations. We proposed a measure of physical information, part of a larger three-layered model of environmental information-interaction [14], and we applied this measure to a sample of twenty cities from different world regions. We found empirical signs that our entropy measure is powerful enough to capture consistent "spatial signatures" of different cities.

Although our sample is limited, our results are in general consistent with previous findings based on different spatial measures. Chicago and American/Canadian cities appear prominently as the most ordered in their cellular configurations, as they are once assessed as metrical, topological or fractal systems [2, 17], or in terms of grid order [1]. But there are interesting specificities identified by our approach. South American cities Buenos Aires and Santiago appear in a cluster with European cities. Since other cities in South America have a cluster of their own, this is an apparently surprising result – one that runs counter the cultural hypothesis.

Our method and findings about physical information signatures of different cities are equally suggestive considering the hypothesis of cultural and regional similarities that lies in the urbanistic imagination. On the one hand, *a general regionally-based theory* of regional morphogenesis of structurally similar cities seems unable to respond to every individual case. Deeper historical contingencies including specific cultural factors might actively shape cities. On the other hand, our results do not allow us to simply refute the cultural hypothesis. In general, the hypothesis stands, but it does require a more complex consideration: *culture matters, but cannot be reduced to regional borders.*

References

1 Geoff Boeing. Urban spatial order: street network orientation, configuration, and entropy. *SSRN Electronic Journal*, 2018.

2 Valério Augusto Soares de Medeiros. *Urbis Brasiliae. O Labirinto das Cidades Brasileiras*. Editora UnB, 2013.

3 David P Feldman and James P Crutchfield. Structural information in two-dimensional patterns: Entropy convergence and excess entropy. *Physical Review E*, 67(5):051104, 2003.

4 Simone Garlandini and Sara Irina Fabrikant. Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In *International Conference on Spatial Information Theory*, volume 5756, pages 195–211. Springer, 2009.

5 Herman Haken and Juval Portugali. The face of the city is its information. *Journal of Environmental Psychology*, 23(4):385–408, 2003.

6 Hermann Haken and Juval Portugali. *Information adaptation: the interplay between shannon information and semantic information in cognition*. Springer, 2014.

7 B Hillier. *Space is the machine*. Cambridge University Press, 1996.

8 Bill Hillier. The genetic code for cities: is it simpler than we think? In J. Portugali, H. Meyer, E Stolk, and E. Tan, editors, *Complexity theories of cities have come of age*, pages 129–152. Springer, 2012.

9 Didier G Leibovici. Defining spatial entropy from multivariate distributions of co-occurrences. In *International Conference on Spatial Information Theory*, pages 392–404. Springer, 2009.

10 Annick Lesne, Jean-Luc Blanc, and Laurent Pezard. Entropy estimation of very short symbolic sequences. *Physical Review E*, 79(4):046208, 2009.

11 R. Louf and M. Barthelemy. A typology of street patterns. *Journal of The Royal Society Interface*, 11(101):20140924, 2014.

12   Kevin Lynch. *The image of the city*, volume 11. MIT press, 1960.

13   Vinicius Netto, Edgardo Brigatti, João Meirelles, Fabiano Ribeiro, Bruno Pace, Caio Cacholas, and Patricia Sanches. Cities, from information to interaction. *Entropy*, 20(11):834, 2018.

14   Vinicius M Netto, Joao Meirelles, and Fabiano L Ribeiro. Social interaction and the city: the effect of space on the reduction of entropy. *Complexity*, 2017, 2017.

15   Juval Portugali. *Complexity, cognition and the city.* Springer Science & Business Media, 2011.

16   A. Rapoport. *The Meaning of the Built Environment.* Sage, 1982.

17   Mahbub Rashid. The Geometry of Urban Layouts. *ArchéoSciences / Journal of Archaeometry*, 2017.

18   Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of vision*, 7(2):1–22, 2007.

19   Thomas Schürmann and Peter Grassberger. Entropy estimation of symbol sequences. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 6(3):414–427, 1996.

20   Claude E Shannon. A Mathematical Theory of Communication. *Bell system technical journal*, 27:379, 1948.

21   Claude E Shannon. A Mathematical Theory of Communication. *Bell system technical journal*, 27:623, 1948.

22   Wilson P Tanner Jr and John A Swets. A decision-making theory of visual detection. *Psychological review*, 61(6):401–409, 1954.

23   Allison Woodruff, James Landay, and Michael Stonebraker. Constant information density in zoomable interfaces. In *Proceedings of the working conference on Advanced visual interfaces*, pages 57–65. ACM, 1998.

24   W. H. Zurek. *Complexity, Entropy and Physics of Information.* Addison-Wesley, Redwood City, CA, 1990.

λProlog(QS): Functional Spatial Reasoning in Higher Order Logic Programming

Beidi Li
Department of Engineering, Aarhus University, DK
beidi.li@eng.au.dk

Mehul Bhatt
School of Science and Technology, Örebro University, Sweden
http://www.mehulbhatt.org/
Mehul.Bhatt@oru.se

Carl Schultz
DIGIT, Aarhus University, DK
http://think-spatial.org
cschultz@eng.au.dk

—————— **Abstract** ——————

We present a framework and proof-of-concept implementation for functional spatial reasoning within *high-order logic programming*. The developed approach extends λProlog to support reasoning over spatial variables via Constraint Handling Rules. We implement our approach within Embeddable λProlog Interpreter (ELPI) and demonstrate key features from combined reasoning over spatial functions and relations. The reported research is an ongoing development of the *declarative spatial reasoning* paradigm.

1 Introduction

Declarative spatial reasoning denotes the ability to (declaratively) specify and solve real-world problems related to mixed geometric (i.e., quantitative) and qualitative visual and spatial representation and reasoning [3]; the paradigm emphasises diverse forms of reasoning capabilities (e.g., question-answering, learning, abduction) with a rich spatio-temporal ontology where aspects pertaining to space, time, events, actions, change, interaction, conceptual knowledge may be handled as first-class objects within a systematic formal artificial intelligence / *knowledge representation and reasoning* (KR) framework [2]. From the practical viewpoint of practical KR methods, this encompasses spatial reasoning with *answer set programming* [12, 14, 15], *constraint logic programming* [3, 10], and *inductive logic programming* [11]. This paper continues this line of work by developing a KR framework for reasoning in a seamless, integrated way over spatial *functions*, spatial relations, and KR-based domain-specific conceptual knowledge.

In many application areas where *space* plays a central role, such as architectural design or Constructive Solid Geometry, it is necessary to not only represent and reason about *relations* between spatial entities, but to also express and evaluate *functions* over spatial entities. For example, we may want to query the incidence relation between a point $(5, 5)$ and the

intersection of two polygons A, B. In the context of architectural design, polygons A and B may be used to represent the *visibility space* from which a sign and a landmark L_A, L_B are visible, and the point may represent an important *threshold* position where a person is expected to need to orient themselves as they enter a large open room:

```
?- A = polygon [(vertex 0 0), ...],
   B = polygon [(vertex 10 0), ...],
   incidence Relation (point 5 5) (intersect A B).
```

The query result is:

```
Relation = exterior.
```

That is, the point is *exterior* to the intersection region meaning that, in the context of architectural design, the sign and the landmark are not mutually visible from the threshold position, suggesting that an occupant may lose orientation at that critical location. For a second example in the context of Constructive Solid Geometry, suppose we have cube *Cube* that has side length 7 and whose centroid is located at point $(5, 5, 5)$, and sphere *Sphere* with radius 4 and centroid $(10, \_, 5)$, such that the Y coordinate of the centroid is unknown (i.e. the Y coordinate is an unbound real valued variable). These spatial entities may be defined by transforming (translating, scaling) primitive unit-sized entities e.g. a unit cube with side length 1 centred at point $(0, 0, 0)$, and a unit sphere with radius 1 centred at $(0, 0, 0)$. We then assert that *Cube* is topologically *part of* the *Sphere*:

```
?- Cube   = translate (vector  5 5 5) (scale 7 unit_cube),
   Sphere = translate (vector 10 _ 5) (scale 4 unit_sphere),
   topology part_of Cube Sphere.
```

The query result is:

```
false.
```

This means that no translation can satisfy the required topological relation, due to the cube being too large to be part of the sphere.

In this paper we show that reasoning over the combination of spatial functions and spatial relations overcomes numerical instability problems in certain well-defined cases (that would otherwise result in logical inconsistencies), and provides significantly more computationally efficient query answering. Returning to the architecture example above, suppose that the visibility spaces A, B are *disconnected*: the intersection of A and B will be the empty (void) region, to which every point is necessarily *exterior*. Therefore, the result that *Relation = exterior* is arrived at based on purely qualitative spatial reasoning, thus avoiding the need for potentially expensive and unstable numerical computations of polygon intersections, and point-region incidence checks.

In this paper we develop the foundations for reasoning about spatial functions in logic programming, λProlog($\mathcal{QS}$), based on the λProlog framework extended with constraint programming. Based on previous work we target a specific class of *qualitative spatial constraints* that we formulate in the framework of polynomial constraint solving [3, 14]. Our key contributions, and the novel features provided by our integration of spatial reasoning in λProlog, are:

- integrated reasoning about both spatial functions and spatial relations (Section 4);

- by representing spatial functions as abstract syntax trees we can avoid logical inconsistencies that arise from numerical instabilities when computing intermediate functions (Section 5).

- a proof-of-concept implementation of λProlog($\mathcal{QS}$)with query examples is available at: http://think-spatial.org/Resources/LamPrologQS.zip

2 Preliminaries: Lambda Prolog

Our λProlog($\mathcal{QS}$) system builds on lambda logic programming theory originally developed by Nadathur and Miller [9], and extended with constraint programming [7].

Prolog. [13] We assume basic familiarity with first-order logic. A *term* is either a variable, constant, or a compound term (or *predicate*) $f(t_1, \ldots, t_n)$ with functor f applied to terms $t_1, \ldots, t_n$. A Prolog program L_P consists of a finite set of universally quantified *rules* of the form $h \leftarrow b_1, \ldots, b_n$ such that h is a predicate, and the expression $b_1, \ldots, b_n$ is a conjunction of predicates (i.e. rules are Horn clauses). Prolog *facts* are rules of the form $h \leftarrow \top$. A *query* is a conjunction of predicates $b_1, \ldots, b_n$. A ground term is a term with no variables. The Herbrand universe U of L_P is the set of ground terms that can be made from the constants and function symbols of L_P. Let q be a query, then $q\theta$ is a conjunction of ground predicates resulting from an assignment of all variables in q to values from U. A *query* is a logical consequence of L_P if $\exists \theta(L_P \models q\theta)$.

λProlog. λProlog [9] is an extension of Prolog that supports λ-terms as data structures, and higher-order programming beyond what can be expressed using Horn clauses.[1] λ-terms include variables (e.g. x, y, z), constants (e.g. alphanumerical strings), function application $(s\ t)$ and abstraction $(\lambda x.s)$, where s, t are λ-terms. λ-terms enable high-order unification by λ-conversion and facilitate the manipulation of variable names and substitution. λPrologalso incorporates a GENERIC search operation for unification so that type errors detected during parsing are used to 1identify goals that will never succeed.

ELPI [4] is an implementation of λProlog extended with a constraint system based on the Constraint Handling Rules (CHR) language [5]. We implement spatial relations as CHR constraints in ELPI. The constraint system extension consists of a *constraint store* and CHR *rules*. Whenever a λ-term is added to the store, all CHR rules are checked to see if a λ-term match occurs, causing the rule to *fire*. Rules have the form:

$$\text{rule } t_{match} \setminus t_{remove} \mid t_{guard} \Leftrightarrow t_{add}$$

where $t_{match}, t_{remove}, t_{add}$ are λ-terms and t_{guard} is a condition that is either true or false. A rule is fired if t_{match} and t_{remove} are in the store, and t_{guard} is true. This causes term t_{add} to be added to the store, and t_{remove} to be removed from the store.

3 Spatial Representation and Reasoning

The qualitative spatial domain ($\mathcal{QS}$) that we focus on in our formal framework consists of the following ontology.

Spatial Domains. Domain entities in $\mathcal{QS}$ are as follows. A 2D *point* is a pair of reals x, y. A *3D point* is a triple of reals x, y, z. A *simple polygon* is a 2D spatial region (single piece, no holes) defined by a list of n vertices (points) $p_1, \ldots, p_n$ (spatially ordered counter-clockwise) such that the boundary is non-self-intersecting, i.e., there does not exist a polygon boundary

---

[1] In summary, Horn clauses in Prolog are replaced by Hereditary Harrop formulas in λProlog. The role of *resolution refutation* as the logical foundation for sound querying in Prolog is replaced by *sequent calculus* in λProlog.

edge between vertices p_i, p_{i+1} that intersects some other edge p_j, p_{j+1} for all $1 \leq i < j < n$ and $i + 1 < j$. A *simple polyhedron* is a 3D spatial region (single piece, no holes) defined by a set of 3D vertices (points) $V = p_1, \ldots, p_n$ and a set of faces $f_1, \ldots, f_m$ where each face is a triple of vertices $v_1, v_2, v_3 \in V$. A (general) *polygon* is a set of *boundaries* and a set of *holes* (each set of which are simple polygons) such that every hole is a non-tangential part of one boundary. A (general) *poyhedron* is a set of *boundaries* and a set of *holes* (each set of which are simple polyhedra) such that every hole is a non-tangential part of one boundary.

A spatial *object* $o \in O$ is a variable associated with a spatial domain D (e.g. the domain of $2D$ points). An *instance* of an object $i \in D$ is an element from the domain. Given $O = \{o_1, \ldots, o_n\}$, and domains $D_1, \ldots, D_n$ such that o_i is associated with domain D_i, then a *configuration* of objects ψ is a one-to-one mapping between object variables and instances from the domain, $\psi(o_i) \in D_i$.

For example, a variable o_1 is associated with the domain D_1 of 2D points. The point $(0, 1)$ is an instance of D_1. A configuration is defined that maps o_1 to $(0, 1)$ i.e. $\psi(o_1) = (0, 1)$.

Spatial Relations and Spatial Functions.   Let $D_1, \ldots, D_n$ be spatial domains. A spatial relation r of arity n $(0 < n)$ is defined as:

$$r \subseteq D_1 \times \cdots \times D_n$$

Given a set of objects O, a relation r of arity n can be asserted as a constraint that must hold between objects $o_1, \ldots, o_n \in O$, denoted $r(o_1, \ldots, o_n)$. The constraint $r(o_1, \ldots, o_n)$ is satisfied by configuration ψ if $(\psi(o_1), \ldots, \psi(o_n)) \in r$. For example, if dc is a topological relation *disconnected*, and O is a set of polygon objects, then $dc(o_4, o_9)$ is the constraint that polygons $o_4, o_9 \in O$ are disconnected. We define topological, size, and incidence spatial relations, as presented in Table 1.[2]

A spatial function f of arity $n - 1$ $(1 < n)$ is defined as:

$$f : D_1 \times \cdots \times D_{n-1} \to D_n$$

That is, each function maps $(n - 1)$ spatial entities to a (single) spatial entity. For example, if *translate* is a spatial transformation function, v is a vector $(5, 5)$ and T is a polygon with vertices $((0, 0), (10, 0), (5, 5))$ then (*translate v T*) evaluates to the polygon with vertices $((5, 5), (15, 5), (10, 10))$. We introduce the unique *void* spatial entity to ensure that spatial functions are closed over the spatial domains. For example, the intersection of two disconnected polygons is not itself a polygon, but rather the *void* spatial entity. Spatial functions defined in λProlog(QS) are presented in Table 1.

4   Spatial Functions in λProlog:

Using the λProlog type system we define fundamental spatial types *point*, *region*, and define vertices, simple polygons, and (general) polygons as functions:

```
%% defining instances of topological relationships (extract only)
type contact, disconnect, partial_overlap, part_of relation_topology.
%% defining specialised spatial domains        %% polymorphic typing through functions (extract only)
kind point  type.                              type vertex       real -> real -> point.
kind region type.                              type polygon      list point -> list point -> region.
kind relation type.                            type spatial_void region.
```

---

[2] *Discrete from* means that two regions do not share any interior point, *overlaps* means they share at least one interior point, and *disconnected* means they do not share any point including on the boundary.

■ **Table 1** $\lambda\mathrm{Prolog}(\mathcal{QS})$ relation predicates and functions.

$\mathcal{QS}$ **Relations**	**Description**
size: Relation × Region × Region	Size relations between regions: smaller, equisize, larger.
topology: Relation × Region × Region	Contact relations between regions: contact, disconnected, discrete_from, overlaps. partial_overlap, part_of, proper_part_of.
incidence: Relation × Point × Region	Incidence relations between points and regions: interior, on_boundary, exterior.
$\mathcal{QS}$ **Functions**	**Description**
centroid: Region → Point	Centre point of region. Centroids of polygons and polyhedra are the average of their vertices.
extent: Region → $\mathbb{R}$	Area for 2D regions, and volume for 3D regions.
translate: Point × Region → Region	Translates region by a vector defined by the given point.
scale: $\mathbb{R}$ × Region → Region	Scales region by the given positive factor about the region's centroid point.
union: Region × Region → Region	Union of two regions.
intersect: Region × Region → Region	Intersection of two regions.
difference: Region × Region → Region	Difference of two regions.

We define spatial functions and spatial relations to range over these types (Table 1):

```
%% signatures of spatial relations              %% signatures of spatial functions
type incidence  relation_incidence -> point  -> region -> o.   type centroid   region -> point.
type topology   relation_topology  -> region -> region -> o.   type extent     region -> real.
type size       relation_size      -> region -> region -> o.   type translate  point  -> region -> region.
                                                               type scale      real   -> region -> region.
                                                               type union      region -> region -> region.
                                                               type intersect  region -> region -> region.
                                                               type difference region -> region -> region.
```

We implement algebraic semantics of spatial relations in CHR. For example, no region is *disconnected* from itself (irreflexive), and if region A is *part of* region B, and region B is *part of* region C, then A must necessarily be a *part of C*:

```
rule(topology disconnected A B) | (A=B) <=> false.  %% disconnected is irreflexive
rule((topology part_of A B), (topology part_of B C)) | true <=> part_of A C.  %% part of is transitive
```

Combined reasoning about spatial functions and relations.   We use λ-terms to capture the higher-order abstract syntax of spatial functions, and *reduce* this structure by rewriting it in a simplified form based on the algebraic properties of those spatial functions. For example, the union of a polygon A with itself, expressed as the λ-term (*union A A*), is necessarily equivalent to A, and thus we can reduce (*union A A*) simply to A without any further geometric calculations. More generally, given two polygons A, B, then (*union A B*) can reduce to B when A is a *part of B*. Even more generally still, the arguments A, B need not be polygons but can be arbitrarily complex spatial λ-terms: if A is part of B then the term (*union A B*) can be reduced to B.

On the other hand, we can deduce that certain spatial relations must hold between the arguments of a function and the result of the function. For example, two non-void regions A, B must each necessarily be *part of* the union of A and B. Similarly, if regions A and B topologically *overlap* then the intersection of A and B must necessarily be part of A and part of B. By recursively stepping through a spatial λ-term, deducing the relations between its parts and simplifying, we can potentially reduce the λ-term at a purely symbolic

level. Once no further reductions can be made, λProlog(QS) evaluates the true numerical spatial functions (union etc.) using computational geometry libraries GPC[3] for polygons and PyMesh[4] for polyhedra. In the following section we demonstrate the power of this approach. The following code excerpt implements the above example cases, and the recursive *simplify* predicate for reducing spatial λ-terms:

```
%% Simplifying abstract syntax trees of spatial functions:
%%    (1) If A is part of B, then (union A B) reduces to B
simplify_ (union A B) B :- topology part_of A B.
%%    (2) If A is disconnected from B, then (intersect A B) reduces to the spatial void type
simplify_ (intersect A B) spatial_void :- topology discrete_from A B.

%% CHR rules for deducing spatial relations between function arguments and function evaluations:
%%    (1) A and B are each part of (union A B)
rule (deduce (union A B)) | true <=>
  topology part_of A (union A B), topology part_of B (union A B).
%%    (2) if A and B contact, then (intersect A B) is part of A and part of B
rule (deduce (intersect A B)) | (topology overlaps A B) <=>
  topology part_of (intersect A B) A, topology part_of (intersect A B) B.

%% Recursive definition of the simplify predicate
simplify (point X Y) (point X Y).   %% base case
simplify (polygon B H) (polygon B H).  %% base case
simplify (Op Left Right) Simp :-   %% recursive step
  simplify Left SLeft, simplify Right SRight,
  (deduce (Op SLeft SRight)), simplify_ (Op SLeft SRight) Simp.
```

5 Empirical Evaluation

In this section we demonstrate key features of our current implementation of λProlog(QS).

Ex1: Architectural Design. This example demonstrates how λ-term reduction based on combined reasoning over spatial functions and relations avoids potentially expensive geometric computations. A building consists of objects represented as facts in the knowledge base, including a landmark *statue* that is positioned in a central courtyard that is visible from many rooms, and a number of signs. Each object has a visibility space, i.e. a polygon describing the points on the floor plan from which an object can be seen (also referred to as the *isovist*). The building has numerous threshold positions from which building occupants are expected to need some orientation if they are unfamiliar with the building, such as the entrance to a large room. This is modelled as facts in λProlog(QS):

```
%% domain objects
landmark (id lm8263) (object_type statue).  sign (id sign73).
threshold_position (point 5.3 82.3).
%% 2D geometric representations of visibility spaces
visibility_space (id lm8263) (polygon [(vertex 52.3 56.0) ...]).
visibility_space (id sign73) (polygon[(vertex 32.3 281.0) ...]).
```

The architect wants to identify threshold positions from which the occupant does *not* have visible access to *both* the central statue and at least one sign.

```
?- threshold_position Position, landmark Statue (object_type statue),
   visibility_space Statue StatueVisibility,
   not((
       (sign Sign), (visibility_space Sign SignVisibility),
       incidence interior Position (intersect StatueVisibility SignVisibility)
   )).
```

Given such visibility constraints, a numerical program will need to compute the intersection of every pair of statue and sign visibility polygons to determine whether the threshold position

---

[3] http://www.cs.man.ac.uk/~toby/alan/software/
[4] https://pymesh.readthedocs.io/en/latest/

lies in their intersection. By contrast, λProlog($\mathcal{QS}$) directly reduces the intersection to the *void* spatial entity at a purely symbolic level when the visibility polygons are disconnected, thus avoiding potentially computationally expensive geometric calculations.

Ex2: Avoiding logical inconsistencies from numerical instability.   This example demonstrates how λProlog($\mathcal{QS}$) guarantees logical soundness for λ-term reduction in cases where relying on numerically evaluating intermediate terms fails. The powerful polygon set operation library GPC cannot be used to conclude the trivial equality (see Figure 1):

```
?- A = simple_polygon([(vertex 0.0 0.0), (vertex 3.0 0.0), (vertex 3.0 4.0)]),
   B = simple_polygon([(vertex 1.0 0.0), (vertex 4.0 1.1), (vertex 0.0 3.2)]),
   equal A (union (intersect A B) (difference A B)).
```

λProlog($\mathcal{QS}$) gives the query result *true*, which is correct. In contrast, when the intermediate results of (*intersect A B*) and (*difference A B*) are evaluated using GPC, and then combined with a GPC union, the result has two extra vertices that are not precisely on the boundary of A due to rounding errors: $((3.0, 0.0), (0.0, 0.0), (0.71, 0.94), (1.7, 2.3), (3.0, 4.0))$, thus leading to a logical inconsistency that $A \neq A$. The problem becomes more evident in the 3D case where PyMesh generates erroneous mesh artefacts from computing $((S_1 \setminus S_2) \cup (S_1 \cap S_2)) \cup S_2$ where S_1 and S_2 are two meshes that approximate spheres (Figure 1). The result *should* be equal to $(S_1 \cup S_2)$ but due to the artefacts this equality does not hold. Again, λProlog($\mathcal{QS}$) gives the correct result through reduction, and only evaluates the actual numerical (geometric) results using GPC and PyMesh when no further λ-term reductions can be made.
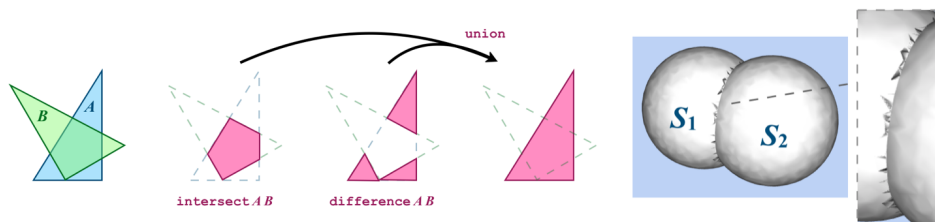
Figure 1 Cases where numerically evaluating intermediate functions using GPC and PyMesh results in logical inconsistencies. λProlog($\mathcal{QS}$) overcomes these limitations with λ-term reduction.

6   Conclusions

We have presented a framework and proof-of-concept implementation of λProlog($\mathcal{QS}$) that integrates functional spatial reasoning within logic programming. Our method facilitates efficient high-level reasoning about both spatial functions, domain-specific knowledge and spatial constraints in a seamless manner. In the broader AI research field, diverse frameworks have been developed that formalise notions of *space*, including: (a) geometric reasoning and constructive solid geometry [6]; (b) relational algebraic semantics of "qualitative spatial calculi" [8] (e.g., the SparQ spatial reasoning tool [16]); and (c) axiomatic frameworks of mereotopology and mereogeometry [1]. However, the distinction with our research here, and what we argue is lacking within the KR community, is a systematic formal account and computational characterisation of such spatial theories as a KR language – e.g., *suited for declarative modelling, commonsense inference and query*. In this paper we emphasise the power of such a research agenda, as our approach leverages from the strengths of both extensive research in functional logic programming and (declarative) spatial reasoning.

────── **References** ──────

1   Marco Aiello, Ian E. Pratt-Hartmann, and Johan F.A.K. van Benthem. *Handbook of Spatial Logics.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

2   Mehul Bhatt. Reasoning about Space, Actions and Change: A Paradigm for Applications of Spatial Reasoning. In *Qualitative Spatial Representation and Reasoning: Trends and Future Directions.* IGI Global, USA, 2012.

3   Mehul Bhatt, Jae Hee Lee, and Carl Schultz. CLP(QS): a declarative spatial reasoning framework. In *International Conference on Spatial Information Theory*, pages 210–230. Springer, 2011.

4   Cvetan Dunchev, Ferruccio Guidi, Claudio Sacerdoti Coen, and Enrico Tassi. ELPI: Fast, Embeddable, lambda-Prolog Interpreter. In *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 460–468. Springer, 2015.

5   Thom Frühwirth. Constraint handling rules. In *French School on Theoretical Computer Science*, pages 90–107. Springer, 1994.

6   Deepak Kapur and Joseph L. Mundy, editors. *Geometric Reasoning.* MIT Press, Cambridge, MA, USA, 1988.

7   Javier Leach, Susana Nieva, and Mario Rodríguez-Artalejo. Constraint logic programming with hereditary Harrop formulas. *Theory and Practice of Logic Programming*, 1(4):409–445, 2001.

8   Gérard Ligozat. *Qualitative Spatial and Temporal Reasoning.* Wiley-ISTE, 2011.

9   Gopalan Nadathur and Dale Miller. An overview of Lambda-PROLOG. *University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-88-40*, 1988.

10   Jakob Suchan and Mehul Bhatt. Semantic Question-Answering with Video and Eye-Tracking Data: AI Foundations for Human Visual Perception Driven Cognitive Film Studies. In *IJCAI*, pages 2633–2639, 2016.

11   Jakob Suchan, Mehul Bhatt, and Carl P. L. Schultz. Deeply Semantic Inductive Spatio-Temporal Learning. In James Cussens and Alessandra Russo, editors, *Proceedings of the 26th International Conference on Inductive Logic Programming (Short papers), London, UK, 2016.*, volume 1865 of *CEUR Workshop Proceedings*, pages 73–80. CEUR-WS.org, 2016. URL: `http://ceur-ws.org/Vol-1865/paper-12.pdf`.

12   Jakob Suchan, Mehul Bhatt, Przemysław Wałęga, and Carl Schultz. Visual explanation by high-level abduction: On answer-set programming driven reasoning about moving objects. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

13   Allen Van Gelder, Kenneth A Ross, and John S Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM (JACM)*, 38(3):619–649, 1991.

14   Przemysław Andrzej Wałęga, Mehul Bhatt, and Carl Schultz. ASPMT(QS): non-monotonic spatial reasoning with answer set programming modulo theories. In *International Conference on Logic Programming and Nonmonotonic Reasoning*, pages 488–501. Springer, 2015.

15   Przemysław Andrzej Wałęga, Carl Schultz, and Mehul Bhatt. Non-monotonic spatial reasoning with answer set programming modulo theories. *Theory and Practice of Logic Programming*, 17(2):205–225, 2017.

16   Diedrich Wolter and Jan Oliver Wallgrün. Qualitative spatial reasoning for applications: New challenges and the SparQ toolbox. In *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications*, pages 1639–1664. IGI Global, 2013.

Towards Modeling Geographical Processes with Generative Adversarial Networks (GANs)

David Jonietz[1]

HERE Technologies, Switzerland
david.jonietz@here.com

Michael Kopp

HERE Technologies, Switzerland
Institute of Advanced Research in Artificial Intelligence (IARAI), Austria
michael.kopp@here.com

───── **Abstract** ─────

Recently, Generative Adversarial Networks (GANs) have demonstrated great potential for a range of Machine Learning tasks, including synthetic video generation, but have so far not been applied to the domain of modeling geographical processes. In this study, we align these two problems and – motivated by the potential advantages of GANs compared to traditional geosimulation methods – test the capability of GANs to learn a set of underlying rules which determine a geographical process. For this purpose, we turn to Conway's well-known Game of Life (GoL) as a source for spatio-temporal training data, and further argue for its (and simple variants of it) usefulness as a potential standard training data set for benchmarking generative geographical process models.

1 Introduction

Contrary to the inherently rather space-focused perspective of Geographical Information Systems (GIS), spatial systems are in general highly dynamic. Thus, the involved geographical entities are susceptible to change with regards to their spatial (e.g., appearance, disappearance, expansion, contraction, movement) or thematic domain (changes of one or more attributes) [6, 4]. Modeling such dynamic behavior is of critical importance for a wide range of applications (e.g., transport planning and traffic prediction, weather forecasting, or disaster management), and involve both explanatory and predictive modeling approaches [20].

Explanatory models are typically targeted towards reaching a thorough understanding of the modeled domain. Relevant features are described in the form of a causal theoretical model, which is then either tested statistically or by hand-crafting a set of fundamental behavioral rules, running simulations and exploring different scenarios (e.g., traffic demand modeling, wildfire spread simulation). In geosimulation applications, in contrast to more aggregate statistical modeling approaches, the elementary system units are typically modeled in great detail as individual automata, using paradigms such as Cellular Automata (CA) or Agent-based Models (ABM) [3].

---

[1] corresponding author

Nowadays, however, due to an increased usage of geo-sensors, large-scale spatio-temporal data sets are widely available which describe various geographical processes with an unprecedented level of detail. Thus, for predictive models, where the focus is put less on understanding the functional principles of the system but rather on predicting its next state based on previous observations [20], supervised machine learning algorithms such as Artificial Neural Networks (ANN) have been increasingly used, e.g., for short-term traffic forecasting [14, 16] or precipitation prediction [26]. In contrast to explanatory approaches, these models are based on data rather than theory, and can usually be trained in an end-to-end manner. Abstract features of relevance for the predictive task are learned directly from the data, however, to what degree such models reach a true understanding of the problem domain remains largely unclear.

Recently, a new wave of generative models has had a disruptive effect on the Machine Learning community, which aim to generate realistic samples of a complex, real world distribution having only observed true samples of said distribution. Thus, being presented with data (e.g., images, text, music or videos), these models move beyond predictive models by learning a representation which particularly encodes important semantic features in order to generate new, hitherto unseen, 'realistic' samples, therefore potentially understanding the underlying data-generating process itself [5]. A particularly successful example for these models are Generative Adversarial Networks (GANs) [8], which require no prior assumptions or hypotheses about the function principles of the modeled system.

To the best of our knowledge, our work represents the first study which explores the potential of GANs for the simulation of geographical processes. This is motivated by the fact that in our view, GANs combine strengths of both explanatory and predictive modeling approaches. GANs, as used here, are explanatory with regards to a geographical process as they capture its underlying hidden rules and -on this basis- are able to not only generate novel sample states, but also provide a learned loss metric which describes how "realistic" a given sample is. In contrast to traditional explanatory models, however, GANs do not rely on hand-crafted parameters (as in expert systems), but directly learn them from observation alone while preserving the capability of capturing and applying complex rules (one could thus refer to them as "self-learned explanatory" models). If the generated samples describe future states of a process a GAN can be used as predictive model, thereby eliminating the need for descriptive rules or a set framework (cf. variational bayesian methods or deterministic methods such as SVMs) while preserving the apparent ability to sample highly complex naturally occurring distributions.

At this stage, we aim to demonstrate and quantify the performance of a GAN on a well-controlled test data set (as it is the only way of measuring the effectiveness of most other neural network architectures as well). For this, we choose a straight-forward example of a complex, non-trivial and non-deteriorating geo-spatial process that arises out of a simple set of deterministic rules: Conway's Game of Life (GoL) [7]. In general, in view of the multitude and diversity of potential use cases from different geo-spatial domains (and the according spatio-temporal data sets), we argue for the general need for a standard training data set for benchmarking generative geographical process models (comparable to MNIST [12] for image processing tasks), and propose to use the GoL - and selected adaptations - for this purpose. In our experiments, we demonstrate that a GAN can indeed learn the underlying rules of the data-generating process (and therefore play the GoL correctly), however, processes with different properties require different network architectures.

2 Generative Adversarial Networks (GANs)

GANs aim to capture the statistical distribution of training data and produce new, hitherto unseen, samples from that distribution. In its original form [8] each GAN model has two parts to it that compete against each other: a generator whose task it is to produce new "fake" samples from the underlying distribution of the observed data (forger) and a discriminator who, when faced with "real" and "fake" samples aims to tell them apart (policeman). The Generator G and Discriminator D can be defined as a functions

$$G_{\theta_g} : \{\text{random input}\} \longrightarrow \{\text{Samples}\}, z \longrightarrow x \tag{1}$$
$$D_{\theta_d} : \{\text{Samples}\} \longrightarrow [0,1], x \longrightarrow k \tag{2}$$

where z refers to an input random noise variable, which is mapped via G to a sample x in data space based on a set of parameters θ_g. Its counterpart D represents a function where any input sample x is mapped to a scalar k which expresses the probability that x is sampled from the original statistical distribution rather than created by G, based on a set of parameters θ_d. Typically, both functions G and D would be implemented as separate neural networks.

The training process outlined in [8] is defined by D and G playing a two-player minimax game with value function $V(G, D)$:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_r}[log D(x)] + \mathbb{E}_{z \sim p_z}[log(1 - D(G(z)))] \tag{3}$$

where p_r is the data distribution (in many cases unknowable) from which 'real' samples x_r are drawn and p_z is the data distribution over noise input z. Although GANs have been able to generate photo-realistic images, there is currently no known way of quantifying how well the generator in general approximates the original distribution. In particular, in some cases, GANs are known to experience mode collapse and a plethora of techniques are employed to mitigate this phenomena (see for instance [19]). So far, GANs have so far been rarely used in the geospatial domain, which is mainly due to their relative novelty and notoriety to be difficult to train. Exemplary applications have been set mostly in a remote sensing context (e.g., [27]), but also included e.g., the generation of static traffic [16] or urbanization patterns [1].

In this study, we use a conditional GAN [18], an extended concept where a conditional input value y is added to the random input z in the formula above so that the aim of the GAN is to produce samples from the corresponding conditional distribution:

$$G_{\theta_g} : \{ \text{ random input, conditional input } \} \longrightarrow \{\text{samples}\}, z, y \longrightarrow x \tag{4}$$
$$D_{\theta_d} : \{\text{samples , conditional input}\} \longrightarrow [0,1], x, y \longrightarrow k \tag{5}$$

In the past, GANs conditioned with past frames of videos have been successfully applied to next frame prediction tasks (see e.g., [15, 13]). In [10], a conditional GAN was used for augmenting the training set for a traffic prediction task. In this work, by modeling a geographical process using the traditional snapshot approach [2], where each 'frame' depicts a time-stamped map view of the current state of the spatial system, we conceptually align the tasks of spatio-temporal modeling and synthetic video generation.

3   Conway's Game of Life

Conway's Game of Life is a popular example of a CA-based game (see [7]). Formally, a CA can be defined as a discrete dynamic system consisting of a n-dimensional fixed lattice arrangement of cells C, each cell $c \in C$ being in a certain state $s_c(t) \in S$ during a discrete time step t where the value lies in some set S. We shall restrict our attention to $S = \{0,1\}$. At time step $t + 1$, it is succeeded by a state which can be described by a transition function φ taking only into account the previous state $s_c(t)$ as well as the previous state of neighbours of all direct neighbours of c in the lattice C. We shall restrict our attention to games where the dependence on the neighbours is indirect, given by a function on the neighbouring states, such as the sum of all 1s occurring. Formally, if N_c denotes the set of neighbours of c in the lattice C and $f_{N_c}(t)$ denotes the function value at time step t for the neighbors of c, then

$$s_c(t + 1) = \varphi\left(s_c(t), f_{N_c}(t)\right) \tag{6}$$

Moving the perspective of the state of an individual cell $c \in C$ to all states configurations of C at a given time step t, we define

$$X(t) = \{s_c(t) \,|\, c \in C\} \tag{7}$$

as the configuration of a CA at time t.

Conway's Game of Life is set on the two-dimensional square lattice $\mathbb{Z}^2$ (where each cell has precisely 8 neighbours) with only two states for each cell and simple rules given, with the notation above, by

$$\varphi(0, x) = \left\{ \begin{array}{ll} 0 & \text{if } x \neq 3 \\ 1 & \text{if } x = 3 \end{array} \right. \qquad \varphi(1, x) = \left\{ \begin{array}{ll} 0 & \text{if } x < 2 \text{ or } x > 3 \\ 1 & \text{if } x \in \{2, 3\} \end{array} \right. \tag{8}$$

If one calls cells with a value 1 and 0 "alive" and "dead", respectively, one can interpret this update rule in terms of survival (cells with 2 or 3 "alive" neighbours stay "alive"), death (through overpopulation or isolation) and birth, see [7]. In order to simulate the game on finite computer architectures, most implementations restrict their view of the lattice $\mathbb{Z}^2$ to $\{0, 1, \cdots, N - 1\}^2$ and decreeing that the state value of neighbours on the boundary of that lattice point square, but outside of it, have state zero.

Despite its simplicity, this game exhibits a surprising variety of oscillating, population increasing and self-replicating state patterns ([7], [22], [25]). In our view, it also represents a powerful abstraction of geographical processes in general, and is therefore a well-suited case study for benchmarking models. Thus, the GoL exhibits similarities to well-known attributes of geographical processes such as the conceptualization of objects, states, processes and events [21], properties related to process dynamics like initiation, cessation and constancy [4], or systemic attributes such as location, topology, spatial interaction [11], or emergence [3].

4   Method

In our experiments, we aim to test whether a GAN can learn the underlying rules of a geographical process, at this stage abstracted as a GoL simulation. For this, we train a GAN on the task of playing the game, i.e., generating the correct next cell configuration $\hat{X}(t)$ while being conditioned on the previous $X(t - n : t)$ configurations (here $n = 4$). Both the generator and the discriminator, therefore, have to internalize the game's transition rules, state space and neighborhood definition in order to successfully fool or expose their counterpart.

4.1 Adaptations of the Game of Life

As discussed previously, numerous properties of the GoL qualify it as a useful abstraction from real-world geographical processes, such as spatial and temporal locality (Moore neighborhood & Markov property) and spatio-temporal dependence. It is clear, however, that most real-world processes are guided by much more complex rules, which, as we argue, can be approximated by manipulating one or more parameters of the traditional game definition. Thus, for instance, more complex spatio-temporal dependencies could be achieved by abolishing the Markov property and introducing more complex, non-uniform neighborhood definitions. Other possible adaptations could include replacing the traditional deterministic with stochastic transition rules, among others.

As a first example for such adaptations, we test our GAN on two versions of the GoL, one following the traditional game definition (in the following: GoL I) and an adapted one (GoL II), where the neighborhood concept is re-defined as follows: If the cell at lattice point $(i,j) \in \{0, 1, \cdots, N-1\}^2$ is denoted by $c(i,j)$, then, if precisely one of $i < \frac{N}{2}$ and $j < \frac{N}{2}$ holds, we replace the neighbourhood $N_{c(i,j)}$ with

$$N^*_{c(i,j)} = \left\{ c(l,k) \,\middle|\, c(k,l) \in N_{c(i,j)} \right\} \tag{9}$$

It is not hard to see that, as sets, $N^*_{c(i,j)} = N_{c(j,i)}$ and since φ is only dependent on the sum of state values over those sets, the upshot of these operations is that we replaced neighbourhoods in the top right quadrant with those corresponding ones in the bottom left through transposition (if we stipulate that i and j, as is the case in matrices, grow from left to right and top to bottom, respectively, in order to define the quadrants above). Thus, with GoL II, the conditions of spatial proximity and homogeneity for defining the neighborhood of cells are dismissed.

From both GoL I and II, we sample 30 000 frame sequences of length 5 frames, each randomly initialized, and split them into training (90%) and test set (10%). For each of the samples in both sets, we use the first 4 frames as conditional input for both the generator and the discriminator, and generate - or discriminate, respectively- the subsequent final frame.

4.2 GAN Architecture

Our GAN architecture (see figure 1) is based on the convolutional long short term memory (convLSTM) approach which has proven successful for a similar spatio-temporal prediction task [26]. Concretely, in the generator the conditional input is encoded via three convLSTM layers with 128 (3 × 3), 128 (3 × 3), and 1 (3 × 3) filters with stride 1, and concatenated with the noise vector z, which has previously been encoded via 2 dense layers with 400 units and leaky ReLU activations [17]. Finally, the encoded features flow through two additional dense layers with each 400 units and leaky ReLU and a final sigmoid activation.

In the discriminator, the conditional input and the predicted $\hat{X}(t)$ or real frame $X(t)$ are concatenated and encoded by two convLSTM layers with 64 (3 × 3), and 64 (3 × 3) filters and stride 1. Instead of a noise vector, however, the encoded features are concatenated with the output of a minibatch discrimination layer [19] in order to prevent mode collapse, and then fed through a dense layer with 32 units and leaky ReLU before a final 1 unit sigmoid activation. To prevent the discriminator from completely dominating the generator, we apply drop-out with a rate of 0.6 to the former, thus randomly dropping 60 % of units during each training batch of both the last convLSTM layer and the first dense layer (see [23]).
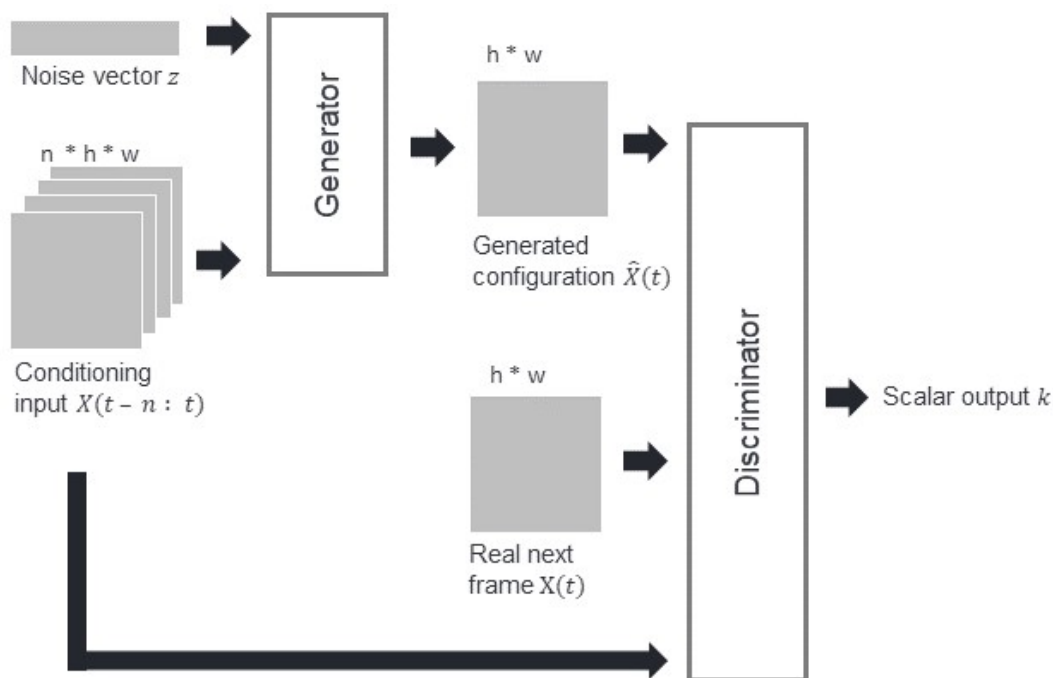
■ **Figure 1** GAN architecture.

5   Results

We implemented the GAN in Python, using the tensorflow library, and tested it on the data of GoL I and II for 50 epochs each with a batch size of 15, using the ADAM optimizer [9] for training both the generator and the discriminator. To track the learning progress, we additionally logged the cross entropy loss of real and generated frames, which is shown in figure 2 for both experiments. With regards to GoL I, the results for GoL I show an almost constant decrease towards 0 for both training and test set, and therefore clearly illustrate a successful learning progress (the high quality of the generated samples is illustrated in figure 3). Thus, the GAN has apparently internalized the underlying rules of the traditional game definition, and was able to generate correct predictions.

However, the results are different for the adapted GoL II. Here, apparently the convolutional layers were unable to successfully encode the adapted, non-proximity based neighborhood definition as defined for a subset of cells. This negative example demonstrates the need for developing and testing alternative network architectures on standardized training data sets to understand the relationships between properties of geographical processes and appropriate network structures. Thus, for instance, in case of GoL II (or other processes with non-proximity based neighborhoods), a deeper network of multiple stacked convolutional layers and larger filter sizes or an attention layer might lead to better results.

6   Conclusion

In this study, we have demonstrated the potential of GAN for understanding the underlying rules of a geographical process directly from its generated data. GANs do not rely on any expert knowledge or theoretical model of the study domain, can be trained end-to-end, and have the ability to generate indistinguishable samples from distributions of any complexity.
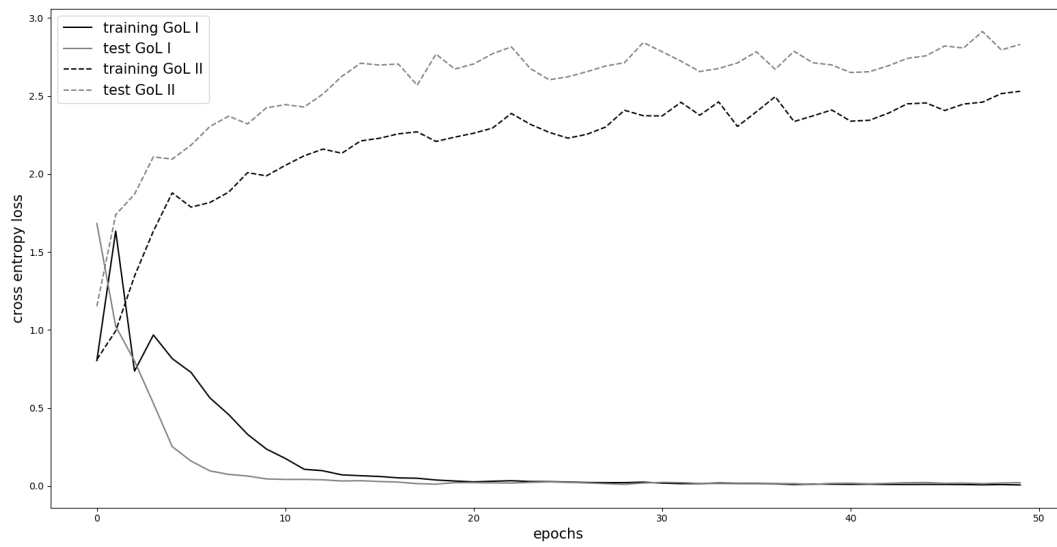
Figure 2 Cross entropy loss for GoL I and II.
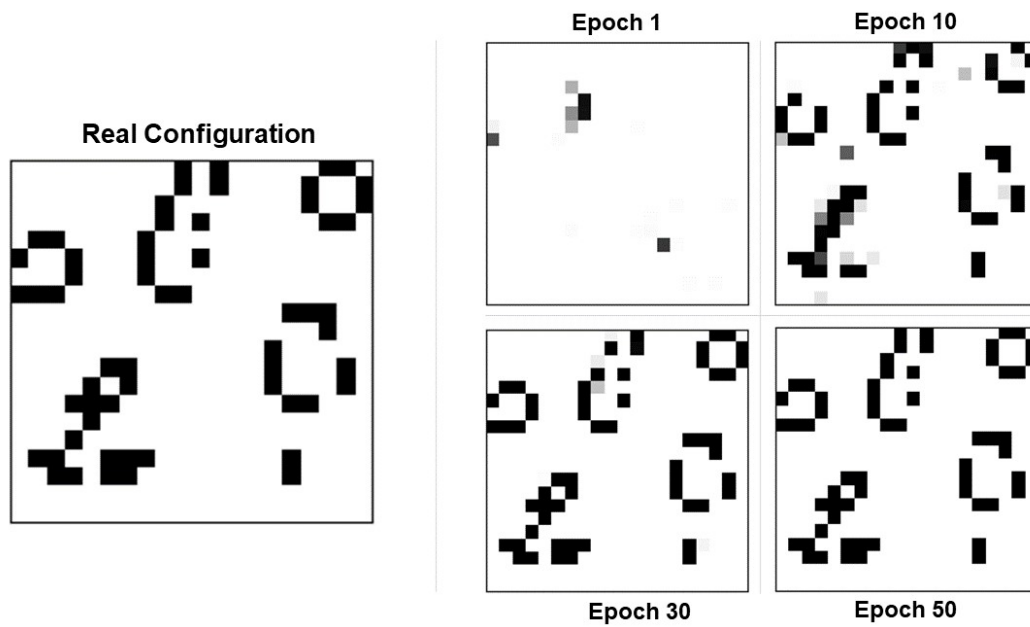
Figure 3 GoL I: Example of the progress of generated samples (right) approximating the real sample (left) until becoming indistinguishable.

Therefore, in our view they are highly promising candidates for simulating geographical processes in general, exploring different scenarios (by conditioning them with different inputs) as well as serve for predictive tasks.

In this preliminary study, we used the GoL as useful abstraction for geographical processes, still, it is clear that its restriction to (few and simple) purely local rules in terms of spatio-temporal interactions represent great simplification compared to real-world processes which are guided by much more complex rules and interactions. Thus, one could expect GAN architectures which were successful on the GoL to fail when presented with real-world spatio-temporal data. Still, however, by manipulating its rules (as we have demonstrated), one could define gradually more complex versions of the game while still maintaining comparable, standard data sets for benchmarking generative models of differing complexities. In general, GANs have no restriction with regards to the complexity of the modeled distribution, i.e., theoretically they can be applied to model any kind of geographical process. Still, more research is needed to evaluate their practical value as an alternative to traditional explanatory or predictive modeling approaches. Additionally, it can be expected that GANs with different architectures will be more or less appropriate to capturing the rules of processes with varying properties. To assess and drive the success of such architectures for general geo-spatial processes will require a set of well understood, plentiful benchmark processes created and utilized by the community.

A downside of GAN is their black box character. Thus, although the network itself has understood and internalized the internal workings of the process, it is challenging to translate that into human-understandable rule descriptions. Still, however, concepts such as transfer learning, where a learned model is transferred and applied to a different task, or attention mechanisms [24] where part of the network's internal reasoning can be made visible (identifying important features for the task), can help to either make explanatory models to a degree obsolete or visualize insights into the derived rules.

References

1 Adrian Albert, Emanuele Strano, Jasleen Kaur, and Marta González. Modeling urbanization patterns with generative adversarial networks. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 2095–2098. IEEE, 2018.

2 Marc P Armstrong. Temporality in spatial databases. In *GIS/LIS 88 Proceedings: Accessing the World*, pages 880–889, 1988.

3 Itzhak Benenson, Paul M Torrens, and Paul Torrens. *Geosimulation: Automata-based modeling of urban phenomena.* John Wiley & Sons, 2004.

4 Claudio EC Campelo, Brandon Bennett, and Vania Dimitrova. Identifying geographical processes from time-stamped data. In *International Conference on GeoSpatial Sematics*, pages 70–87. Springer, 2011.

5 Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in neural information processing systems*, pages 2172–2180, 2016.

6 Christophe Claramunt and Marius Theriault. Toward semantics for modelling spatio-temporal processes within GIS. *Advances in GIs Research I*, pages 27–43, 1996.

7 Martin Gardener. MATHEMATICAL GAMES: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123, 1970.

8 Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

9 Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

10 Arief Koesdwiady and Fakhri Karray. Improving Multi-Step Traffic Flow Prediction, 2018. `arXiv:1803.01365`.

11 Patrice Langlois. *Simulation of complex systems in GIS.* John Wiley & Sons, 2013.

12 Yann LeCun and Corinna Cortes. MNIST handwritten digit database, 2010. URL: `http://yann.lecun.com/exdb/mnist/`.

13 Alex X Lee, Richard Zhang, Frederik Ebert, Pieter Abbeel, Chelsea Finn, and Sergey Levine. Stochastic adversarial video prediction. *arXiv preprint arXiv:1804.01523*, 2018.

14 Linchao Li, Shanglu He, Fan Yang, Jian Zhang, and Bin. Ran. Space-mean-speed Prediction Based on the Fusion of Multi-source Data. In *TRB 96th Annual Meeting Compendium of Papers*, 2017.

15 Xiaodan Liang, Lisa Lee, Wei Dai, and Eric P Xing. Dual motion GAN for future-flow embedded video prediction. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1744–1752, 2017.

16 Yunyi Liang, Zhiyong Cui, Yu Tian, Huimiao Chen, and Yinhai Wang. A Deep Generative Adversarial Architecture for Network-Wide Spatial-Temporal Traffic-State Estimation. *Transportation Research Record*, page 0361198118798737, 2018.

17 Andrew L Maas, Awni Y Hannun, and Andrew Y Ng. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3, 2013.

18 Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

19 Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In *Advances in neural information processing systems*, pages 2234–2242, 2016.

20 Galit Shmueli et al. To explain or to predict? *Statistical science*, 25(3):289–310, 2010.

21 Willington Siabato, Christophe Claramunt, Sergio Ilarri, and Miguel Ángel Manso-Callejo. A Survey of Modelling Trends in Temporal GIS. *ACM Computing Surveys (CSUR)*, 51(2):30, 2018.

22 Michael Simkin. The hunting of the New Herschel Conduits, 2015. URL: `http://conwaylife.com/forums/viewtopic.php?f=2&t=1599&start=200#p19125`.

23 Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

24 Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

25 Andrew J. Wade. Gemini, 2010. URL: `http://conwaylife.com/w/index.php?title=Gemini`.

26 SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.

27 Yang Zhang, Xiang Li, and Qianyu Zhang. Road Topology Refinement via a Multi-Conditional Generative Adversarial Network. *Sensors*, 19(5):1162, 2019.

Granular Spatial Calculi of Relative Directions or Movements with Parallelism: Consistent Account

Reinhard Moratz
Institute for Geoinformatics, University of Münster, Germany
https://www.uni-muenster.de/Geoinformatics/en/institute/staff/index.php/334/
Reinhard_Moratz
moratz@uni-muenster.de

Leif Sabellek
Department of Mathematics and Computer Science, University of Bremen, Germany
https://user.informatik.uni-bremen.de/sabellek/
sabellek@uni-bremen.de

Thomas Schneider
Department of Mathematics and Computer Science, University of Bremen, Germany
http://www.informatik.uni-bremen.de/~schneidt/
thomas.schneider@uni-bremen.de

──── **Abstract** ────

The $\mathcal{OPRA}^*$ calculus family, originally suggested by Frank Dylla, adds parallelism to the $\mathcal{OPRA}$ calculus family which is very popular in Qualitative Spatio-temporal Reasoning (QSTR). Adding parallelism enables the direct representation of parallel moving objects, which is relevant in many applications like traffic monitoring. However, it turned out that it is hard to derive a sound geometric analysis. So far no sound spatial reasoning was supported. Our new generic analysis based on combining condensed semantics lower bounds with upper bounds from algebraic mappings of related calculi already leads to a close and sound approximization. This approximization can be easily augmented with a manual analysis of few geometrically underconstrained cases and then yields a complete analysis of possible configurations in this oriented point framework. This for the first time enables sound standard QSTR constraint reasoning for the $\mathcal{OPRA}^*$ calculus family.

1 Introduction

Qualitative spatial representations provide mechanisms which characterize essential properties of objects or configurations and only make relatively coarse distinctions between spatial relations and configurations, and typically rely on relative comparisons rather than measuring. The concept of *qualitative space* then can be characterized by the following quotation from Galton [7]: "The divisions of qualitative space correspond to salient discontinuities in our apprehension of quantitative space". Qualitative spatial and temporal calculi as formally defined and investigated in the research area of qualitative spatio-temporal reasoning (QSTR) aim at modeling this human commonsense reasoning about space and time using qualitative

relations for different spatial aspects such as topology (e.g., "included in"), direction ("to the left of"), and position as a combination of direction and distance, holding between elementary spatial entities such as points or regions. Coarse spatial knowledge can be used to represent incomplete and underdetermined knowledge systematically. This is especially useful if the task is to describe features of classes of configurations rather than features of individual configurations. For example, the observation that the goal keeper usually stands in front of the goal is true for a variety of ball games. A more specific expression about their position typically would have to refer to the corresponding configuration of a specific sport. Similarly, descriptions of allowed or desired spatial behavior are abstractions mapping infinite sets of possible quantitative configurations or trajectories to a single qualitative description. If qualitative spatial divisions serve as knowledge representation in a reasoning system, deductive inferences can be realized as constraint-based reasoning. Qualitative spatial calculi of relative directions are important for applications such as human-robot interaction, volunteered geographic information, scene understanding, outdoor robotic navigation [10].

A qualitative calculus consists of a set of base relations and a composition table; the latter enables spatial reasoning. There is a wide range of qualitative spatial/temporal calculi, and they are understood to varying levels of detail, see, e.g., the recent survey [3].

The calculi from the $\mathcal{DRA}$ [12, 11] and $\mathcal{OPRA}$ [9, 14] families are prominent examples of calculi of relative directions. They are available at varying granularities: $\mathcal{DRA}$ admits three granularities (variants $\mathcal{DRA}_c$, $\mathcal{DRA}_f$, $\mathcal{DRA}_{fp}$, i.e., "coarse-grained", "fine-grained" and "with parallelism and anti-parallelism"). In particular, $\mathcal{DRA}_{fp}$ extends $\mathcal{DRA}_f$ with the ability to capture parallelism, anti-parallelism, and positive and negative alignment. The $\mathcal{OPRA}$ family admits arbitrarily fine granularities, indicated by a subscript n. Already for small n, $\mathcal{OPRA}_n$ has a large number of base relations (72 for $n = 2$), which prohibits a manual computation of the composition table. For this reason, Moratz and Mossakowski [14] performed a systematic geometric analysis of oriented points in the 2D plane, resulting in a generic algorithm for computing the composition table in $\mathcal{OPRA}_n$ for any n.

Dylla and Lee [1, 2] extended $\mathcal{OPRA}$ in a way that is analogous to the way how $\mathcal{DRA}_{fp}$ extends $\mathcal{DRA}_f$. The resulting $\mathcal{OPRA}^*$ family refines $\mathcal{OPRA}$ with the ability to capture (anti-)parallelism and positive/negative alignment. It later turned out that the original algorithm for computing the composition table does not provide a sound geometric analysis, nor has an alternative algorithm been found yet. It is also far from obvious how to extend Moratz and Mossakowski's analysis to incorporate parallelism. For this reason, we develop an approach to compute the composition table of $\mathcal{OPRA}_n^*$ that relies on homomorphic embeddings into other calculi, geometric constraints on realizable triples of oriented points, and an enumeration of canonical configurations of triples of oriented points.

2   Qualitative Spatial and Temporal Reasoning

Objects and locations can be represented as simple, featureless points. In contrast, the $\mathcal{OPRA}_n$ calculus uses more complex basic entities: It is based on objects which are represented as oriented points. It is related to a calculus which is based on straight line segments (dipoles) [12]. Conceptually, the oriented points can be viewed as a transition from oriented line segments with concrete length to line segments with infinitely small length [11]. In this conceptualization the length of the objects no longer has any importance. Thus, only the orientation of the objects is modeled. *Opoints*, our term for oriented points, are specified as pair of a point and a orientation on the 2D-plane.

In a coarse representation, a single opoint induces the sectors depicted in Figure 1a. "Front", "Back", "Left", and "Right" are linear sectors. "Left-Front", "Right-Front", "Left-Back", and "Right-Back" are quadrants. The position of the point itself is denoted as "Same". This qualitative granularity corresponds to Freksa's double cross calculus [5].
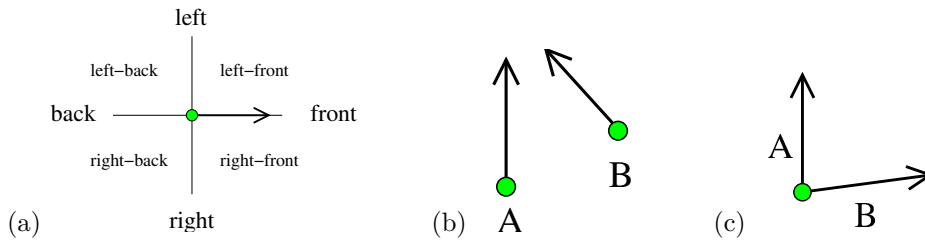
▪ **Figure 1** (a) An opoint and its qualitative spatial relative orientations. (b) and (c) Qualitative spatial relation between two opoints at (b) different positions, here $A \, {}^{\text{leFr}}_{\text{riFr}} \, B$, and (c) the same position, here $A \, {}^{\text{riFr}}_{\text{same}} \, B$.

A qualitative spatial relative direction relation between two opoints is represented by the sector in which the second opoint lies with respect to the first one and by the sector in which the first one lies with respect to the second one. For the general case of the two points having different positions we use the following relation symbols:

$$\frac{\text{front}}{\text{front}}, \frac{\text{leFr}}{\text{front}}, \frac{\text{left}}{\text{front}}, \frac{\text{leBa}}{\text{front}}, \frac{\text{back}}{\text{front}}, \frac{\text{riBa}}{\text{front}}, \frac{\text{right}}{\text{front}}, \frac{\text{riFr}}{\text{front}}, \frac{\text{front}}{\text{leFr}}, \frac{\text{leFr}}{\text{leFr}}, \dots, \frac{\text{riFr}}{\text{riFr}}.$$

The abbreviated sector name for the sector where the second opoint position is located from the perspective of the first opoint is the lower part of the relation symbol. Conversely, the sector name for the relative position of the first opoint location using the second opoint as a reference is put atop the other abbreviated sector name. We thus obtain 8×8 base relations for two opoints having different positions. The configuration in Figure 1b is expressed via the relation $A \, {}^{\text{leFr}}_{\text{riFr}} \, B$. If both opoints share the same position, the lower relation symbol part is the word "same" and the upper part denotes the orientation of the second opoint w.r.t. the first; see Figure 1c. Altogether we obtain 72 different atomic relations (8×8 general relations plus 8 with the opoints at the same position). These relations are jointly exhaustive and pairwise disjoint (JEPD). The relation $\frac{\text{front}}{\text{same}}$ is the identity relation. The granularity of the $\mathcal{OPRA}$ version we just described is $n = 2$, so this calculus version is called $\mathcal{OPRA}_2$. The general schema for arbitrary m is described below.

The $\mathcal{OPRA}_2^*$ calculus [1, 2] is similar to $\mathcal{OPRA}_2$. The important extension is a refinement of the relations by marking them with letters '+' or '−', 'P' or 'A', according to whether the two orientations of the oriented points are positive (e.g. turning the first opoint in the direction of the second opoint would need a mathematically positive turn), negative, parallel or anti-parallel.

A comprehensive simulation using the $\mathcal{OPRA}$ calculus for an important subtask was built by Dylla et. al. [16]. Their system SailAway simulates the behaviour of different vessels following declarative (written) navigation rules for collision avoidance. This system can be used to verify whether a given set of rules leads to stable avoidance between potentially colliding vessels. The different vessel categories that determine their right-of-way priorities are represented in an ontology. The vessel's movement is described by a method called conceptual neighborhood-based reasoning (CNH reasoning). CNH reasoning describes whether two spatial configurations of objects can be transformed into each other by small changes [6]. A CNH transformation can be an object movement in a short period of time.

Instead of using $\mathcal{OPRA}_4$, like in the original SailAway system, we use this domain to show how the $\mathcal{OPRA}_2^*$ calculus can model parallel movement like in a typical overtake (e.g. catch up with and pass while travelling in the same direction) event. Fig. 2 shows a CNH transition diagram which represents relative trajectories of two vessels during such an overtake event (for an earlier version of qualitative navigation simulation, see [4]). The depicted sequence between two vessels A and B is: $A \, {}^{\text{leBa}}_{\text{riFr}} \text{P} \, B \; \rightarrow \; A \, {}^{\text{left}}_{\text{right}} \text{P} \, B \; \rightarrow \; A \, {}^{\text{leFr}}_{\text{riBa}} \text{P} \, B$.
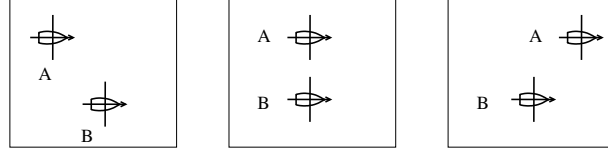
■ **Figure 2** Representation of vessel navigation with conceptual neighbourhood in $\mathcal{OPRA}_2^*$.

Preliminaries.   A *qualitative calculus* $\mathcal{A} = (U_\mathcal{A}, R_\mathcal{A})$ consists of a set $U_\mathcal{A}$ called the *universe* of $\mathcal{A}$ and a set $R_\mathcal{A}$ of binary relations on $U_\mathcal{A}$ called *base relations* that are *JEPD* (jointly exhaustive and pairwise disjoint), i.e. $r \cap s = \emptyset$ for $r, s \in R_\mathcal{A}$ with $r \neq s$ and $\bigcup_{r \in R_\mathcal{A}} r = U_\mathcal{A} \times U_\mathcal{A}$. Furthermore, if r is a base relation, then the *converse* $r^\smile = \{(a, b) \mid (b, a) \in r\}$ must be a base relation as well. A *general relation* is a union of base relations.

Every qualitative calculus $\mathcal{A} = (U_\mathcal{A}, R_\mathcal{A})$ gives rise to an algebraic structure via *weak composition* of relations from $R_\mathcal{A}$ in the following way. If $r, s \in 2^{R_\mathcal{A}}$ are general relations, then $r \diamond s = \{t \in R_\mathcal{A} \mid r \circ s \cap t \neq \emptyset\}$, where $r \circ s$ is the usual set-theoretic composition.

We define the $\mathcal{OPRA}_n$ and $\mathcal{OPRA}_n^*$ families of calculi as introduced in [9, 2]. Their universe is the set $\mathcal{O} = \mathbb{R} \times \mathbb{R} \times [0, 2\pi)$ of *opoints* in the 2D-plane. In the $\mathcal{OPRA}_n$, every opoint $p = (x, y, \phi)$ is associated with n lines, all intersecting at (x, y) and pointing to the directions $\{\phi + i \cdot \frac{\pi}{n} \mid i = 0, \ldots, n - 1\}$. These n lines partition $\mathbb{R} \times \mathbb{R} \setminus \{(x, y)\}$ into $2n$ sections which are numbered 0 to $2n - 1$: The ray which p points towards ϕ has number 0; the other sections are numbered counter-clockwise, so 1-dimensional (2-dimensional) rays are assigned even (odd) numbers. If $(u, v) \in \mathbb{R}^2$, we write $\mathsf{pos}(u, v, p) = i$ if (u, v) is in the i-th section of p, and $\mathsf{pos}(u, v, p) = s$ if $(u, v) = (x, y)$ ($s = same$).

The base relation between two opoints $p_1 = (x_1, y_1, \phi_1)$ and $p_2 = (x_2, y_2, \phi_2)$ is described by the location of p_2 relative to p_1 and the location of p_1 relative to p_2. For $i, j \in \{0, \ldots, 2n-1\}$ let $\angle_i^j$ be the set of all pairs (p_1, p_2) of opoints $p_1 = (x_1, y_1, \phi_1)$ and $p_2 = (x_2, y_2, \phi_2)$ such that $i = \mathsf{pos}(x_2, y_2, p_1)$ and $j = \mathsf{pos}(x_1, y_1, p_2)$. For $i \in \{0, \ldots, 2n - 1\}$ let $\angle_s^i$ be the set of all pairs (p_1, p_2) of opoints $p_1 = (x_1, y_1, \phi_1)$ and $p_2 = (x_2, y_2, \phi_2)$ such that $x_1 = x_2$ and $y_1 = y_2$ and ϕ_2 points into section i of p_1. Now $\mathcal{OPRA}_n$ is the qualitative calculus with universe $\mathcal{O}$ and base relations $\{\angle_i^j \mid 0 \leq i, j \leq 2n - 1\} \cup \{\angle_s^i \mid 0 \leq i \leq 2n - 1\}$.

The calculus $\mathcal{OPRA}_n^*$ refines $\mathcal{OPRA}_n$ by adding information about parallelism. Let $\alpha(p_1, p_2) = \phi_2 - \phi_1$ if $\phi_2 - \phi_1 \geq 0$ and $\alpha(p_1, p_2) = \phi_2 - \phi_1 + 2\pi$ otherwise. Every $\mathcal{OPRA}_n$ base relation $\angle_i^j$ can be partitioned into four relations, some of which will be empty:

$$\angle_i^j \mathsf{P} = \angle_i^j \cap \{(p_1, p_2) \mid \alpha(p_1, p_2) = 0\} \qquad \angle_i^j + = \angle_i^j \cap \{(p_1, p_2) \mid 0 < \alpha(p_1, p_2) < \pi\}$$

$$\angle_i^j \mathsf{A} = \angle_i^j \cap \{(p_1, p_2) \mid \alpha(p_1, p_2) = \pi\} \qquad \angle_i^j - = \angle_i^j \cap \{(p_1, p_2) \mid \pi < \alpha(p_1, p_2) < 2\pi\}$$

The base relations of $\mathcal{OPRA}_n^*$ are all non-empty relations of the form $\angle_i^j *$, where $0 \leq i, j \leq 2n - 1$ and $* \in \{\mathsf{P}, +, \mathsf{A}, -\}$ as well as all relations of the form $\angle_s^i$, where $0 \leq i \leq 2n - 1$.

Let r, s, t be base relations. We say that the triple (r, s, t) is *realizable*, if $r \diamond s \ni t$ and that the triple is *impossible* otherwise. For a realizable triple (r, s, t), we say that $(p_1, p_2, p_3) \in \mathcal{O}^3$ realizes (r, s, t), if $r(p_1, p_2)$, $s(p_2, p_3)$ and $t(p_1, p_3)$. Computing the composition table of a calculus is the same as computing the set of realizable triples.

3   Composition table of $\mathcal{OPRA}_2^*$

We compute the $\mathcal{OPRA}_2^*$ composition table twice, using two different algorithms which are performed independently of each other: (1) We enumerate realizable triples, using a condensed semantics approach in the spirit of [9, 11]. Since every realizable triple (r, s, t)

certifies that $r \diamond s \ni t$, the enumeration yields a *lower bound* for the composition table, that is, a subset of the set of all realizable triples. In contrast to [8], the condensed semantics approach does not generate realizable triples randomly but via a systematic enumeration that exploits the geometric properties of the underlying calculus. (2) Starting from the set of all triples, we eliminate impossible triples by computing homomorphisms $\mathcal{OPRA}_2^* \to \mathcal{OPRA}_2$ and $\mathcal{OPRA}_2^* \to \mathcal{OPRA}_1^*$, and by observing angular, location, and permutation constraints. This way we obtain an *upper bound* for the composition table.

The lower bound is obviously a subset of the upper bound. After computing both using the algorithms described, it will turn out that, for $\mathcal{OPRA}_2^*$, the upper and lower bound coincide. This implies that either of them computes the $\mathcal{OPRA}_2^*$ composition table.

3.1 Lower bound

Our aim in this section is to compute a lower bound for the composition table, i.e., a list (set) of configurations of opoint triples that are guaranteed to be contained in the table. Using the condensed semantics approach analogously to previous work on the Dipole calculus [11] and the $\mathcal{OPRA}$ calculus [9], we found a qualitative abstraction in a discrete geometry that has a mapping to the equivalence classes in the $\mathbb{R}^2$ plane of the original model domain.

We used a set of different qualitative triangles relevant for positions of three opoints. In the first triangle all three opoints are on the same location. In the second location triple two points are on the same position and the third point is at a different location. In a grid we constructed specific configurations of opoints as vertices of the following list of triangles. The first vertex is fixed at position $(0, 0)$ the second vertex is fixed at position $(8, 0)$. With a third vertex at the positions $(4, 0), (4, 2), (4, 3), (4, 4), (4, 8)$ we constructed five triangles. At each vertex there are only limited qualitatively different options for opoints in our $\mathcal{OPRA}_2^*$ domain. We used 32 orientations for opoints at each vertex. Then the exhaustive enumeration of all opoint options (e.g. including permutations of the three arguments) for all three vertices for all seven three location configurations generates a lower bound for the composition table. With this approach there is no guarantee that every possible opoint triple w.r.t. the $\mathcal{OPRA}_2^*$ domain is constructed. So our condensed semantics method provides only a lower bound without the guarantee that all entries in the composition table are complete. Therefore we augmented our approach with an upper bound using a method based on abstract algebra [15] presented next.

3.2 Upper bound

We describe the upper bound algorithm. We first introduce a homomorphism technique to derive information about $\mathcal{OPRA}_2^*$ from $\mathcal{OPRA}_2$ and $\mathcal{OPRA}_1^*$, making use of the fact that the composition tables for the latter calculi are known [14, 2]. Then we improve the upper bound using two methods which we call angular constraints and location constraints. The last two methods are then refined by considering permutations of relations in a triple.

Homomorphisms to $\mathcal{OPRA}_2$ and $\mathcal{OPRA}_1^*$. Let $(U_\mathcal{A}, R_\mathcal{A})$ and $(U_\mathcal{B}, R_\mathcal{B})$ be qualitative calculi. We observe that every map $f : R_\mathcal{A} \to R_\mathcal{B}$ with the condition $(*)$ $f(r \diamond s) \subseteq f(r) \diamond f(s)$ yields an upper bound for the composition table of $\mathcal{A}$ by $r \diamond s \subseteq f^{-1}(f(r \diamond s)) \subseteq f^{-1}(f(r) \diamond f(s))$, so for every cell $r \diamond s$ in the table, $f^{-1}(f(r) \diamond f(s))$ is an upper bound that can be computed using the composition in $\mathcal{B}$. We give a sufficient condition for a map f to have condition $(*)$.

A function $f : U_\mathcal{A} \to U_\mathcal{B}$ is said to *induce a map on base relations* if for every base relation $r \in R_\mathcal{A}$ there exists a base relation $s \in R_\mathcal{B}$ s.t. $f(r) \subseteq s$. In this case, we denote the induced map $R_\mathcal{A} \to R_\mathcal{B}$ by f as well. The following lemma is proved in [13].

▶ **Lemma 1.** *If $f : U_\mathcal{A} \to U_\mathcal{B}$ induces a map on base relations, then $f(r \diamond s) \subseteq f(r) \diamond f(s)$.*

Now we establish the first upper bound for the composition table of $\mathcal{OPRA}_2^*$. Since $\mathcal{OPRA}_2^*$ combines features from two calculi, we obtain upper bounds from two natural homomorphisms, namely the quotient homomorphisms $f : \mathcal{OPRA}_2^* \to \mathcal{OPRA}_2$ and $g : \mathcal{OPRA}_2^* \to \mathcal{OPRA}_1^*$, which are both induced by the identity on $\mathcal{O}$. Both f and g induce a map on base relations, so by Lemma 1, they yield two upper bounds for the composition table of $\mathcal{OPRA}_2^*$, which can be calculated from the known composition tables of $\mathcal{OPRA}_2$ and $\mathcal{OPRA}_1^*$. The homomorphism f forgets the information about parallelism, whereas g maps the regions of $\mathcal{OPRA}_2^*$ to the coarser ones of $\mathcal{OPRA}_1^*$. Formally, $f(\angle_i^j *) = \angle_i^j$ with $* \in \{\mathsf{P}, \mathsf{A}, +, -\}$, and $g(\angle_i^j *) = \angle_{\rho(i)}^{\rho(j)} *$ with $* \in \{\mathsf{P}, \mathsf{A}, +, -\}$, where ρ maps the number of a section in $\mathcal{OPRA}_2$ to that of the corresponding section in $\mathcal{OPRA}_1$, so $\rho(0) = 0$, $\rho(1) = \rho(2) = \rho(3) = 1$, $\rho(4) = 2$ and $\rho(5) = \rho(6) = \rho(7) = 3$.

Angular constraints. We describe the method of *angular constraints*, which excludes triples that are impossible due to contradictory information about the angle of the third point relative to the first point. Consider two opoints $p_i = (x_i, y_i, \phi_i) \in \mathcal{O}$, where $i \in \{1, 2\}$. We first describe how to obtain a constraint on $\alpha(p_1, p_2)$. Let the relative angle $\mathsf{a}(p_1, p_2)$ be the number of the section in which ϕ_2 points relative to p_1. Precisely, let $\alpha = \alpha(p_1, p_2)$, then

$$
\mathsf{a}(p_1, p_2) = \begin{cases}
0 & \text{if } \alpha = 0 \\
1 & \text{if } 0 < \alpha < \frac{\pi}{2} \\
2 & \text{if } \alpha = \frac{\pi}{2} \\
3 & \text{if } \frac{\pi}{2} < \alpha < \pi \\
4 & \text{if } \alpha = \pi \\
5 & \text{if } \pi < \alpha < \frac{3\pi}{2} \\
6 & \text{if } \alpha = \frac{3\pi}{2} \\
7 & \text{if } \frac{3\pi}{2} < \alpha < 2\pi
\end{cases}
$$

and if r is a base relation, we define $\mathsf{a}(r) = \{\mathsf{a}(p_1, p_2) \mid p_1 r p_2\}$.

Now assume we have a triple (r, s, t) of base relations and want to know if it is realizable. If the triple is realized by three opoints p_1, p_2, p_3, then $\mathsf{a}(p_1, p_2) \in \mathsf{a}(r)$, $\mathsf{a}(p_2, p_3) \in \mathsf{a}(s)$ and $\mathsf{a}(p_1, p_3) \in \mathsf{a}(t)$. At the same time, $\mathsf{a}(r)$ and $\mathsf{a}(s)$ impose another constraint on $\mathsf{a}(t)$ by composing the possible angles. If these two constraints on $\mathsf{a}(t)$ have an empty intersection, then (r, s, t) is an impossible triple. Figure 3 shows an example.

The subroutine `isAngleCombinationPossible` gets as an input a triple (r, s, t) of base relations and returns a Boolean indicating whether the triple is impossible due to contradictory information about the relative angle. If at least one of $\mathsf{ang}(r, s)$ and $\mathsf{ang}(s, t)$ is even, then the resulting constraint for $\mathsf{ang}(r, t)$ is the singleton set $S = \{(\mathsf{ang}(r, s) + \mathsf{ang}(s, t)) \bmod 8\}$; otherwise $S = \{(u-1) \bmod 8, u \bmod 8, (u+1) \bmod 8 \mid u = \mathsf{ang}(r, s) + \mathsf{ang}(s, t)\}$. If $\mathsf{ang}(r, t) \notin S$, then `false` is returned, otherwise `true`.

Location constraints. The next improvement is obtained by *location constraints*. Here we exclude triples that are impossible due to contradictory information about the location of the third point relative to the first. Figure 4 shows how to identify such impossible triples. The algorithm `isLocationCombinationPossible` gets as input a triple (r, s, t) of base relations and returns a Boolean indicating whether the triple is ruled out for the said reason. To achieve this, we assume there is a triple (p_1, p_2, p_3) realizing (r, s, t). From r and
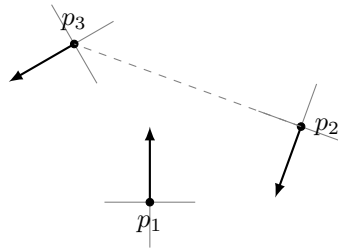
Figure 3 If $r = \angle_7^7+$, $s = \angle_6^3-$, and $t = \angle_7^3\mathsf{P}$, then $\mathsf{a}(r) = \{3\}$, $\mathsf{a}(s) = \{7\}$ and $\mathsf{a}(t) = \{0\}$. At the same time, $\mathsf{a}(r)$ and $\mathsf{a}(s)$ impose the constraint $\{1, 2, 3\}$ on $\mathsf{a}(t)$, and since $\{1, 2, 3\} \cap \{0\} = \emptyset$, the triple (r, s, t) is impossible. The image shows opoints p_1, p_2, p_3 such that $r(p_1, p_2)$ and $s(p_2, p_3)$. Under these circumstances, it is impossible that p_1 and p_3 are parallel, so $t(p_1, p_3)$ can be ruled out.

s, we compute a constraint on the location of p_3 relative to p_1, by systematically analyzing all possible cases and exploiting symmetry. See [13] for a complete list of all cases with visualizations.

Permutation constraints. Let r, s, t be base relations. It is easy to see that the triples (r, s, t), $(r^\smile, t, s)$, $(s, t^\smile, r^\smile)$, $(s^\smile, r^\smile, t^\smile)$, $(t, s^\smile, r)$ and $(t^\smile, r, s^\smile)$ are either all realizable or all impossible: if (x, y, z) realizes one of these triples, then its permutations $(x, z, y), \ldots, (z, y, x)$ realize the other triples. Hence the final step of computing the upper bound traverses all triples $(r, s, t) \in B^3$; whenever one such triple has been excluded by some homomorphism, angular constraint or location constraint, then its other five permutations are excluded, too.

Upper bound algorithm. Algorithm 1 is the final algorithm for the upper bound. Recall that f (g) is the homomorphism from $\mathcal{OPRA}_2^*$ to $\mathcal{OPRA}_2$ (to $\mathcal{OPRA}_1^*$).

3.3 Discussion

The lower bound from Section 3.1 is correct since it generates only realizable triples. The upper bound from Section 3.2 is correct since it eliminates only impossible triples. Our implementation shows that both bounds coincide for $\mathcal{OPRA}_2^*$, so our method computes the correct composition table for this calculus.

In principle, our method can be applied to other members of the $\mathcal{OPRA}_n^*$ family. However, the approaches to computing both the lower and upper bound rely on heuristics, and it is not reasonable to expect that the lower and upper bounds will always coincide. If they do not, then the method will only yield a "range" of possible composition tables and, in order to compute the table precisely, it would be necessary to find an appropriate refinement of the
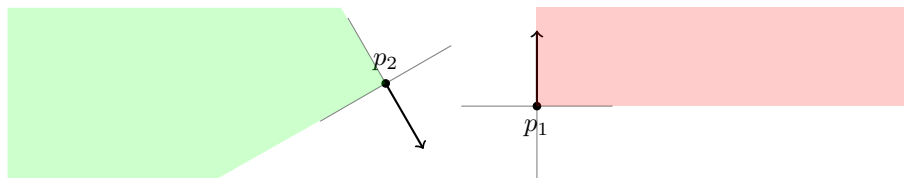
Figure 4 $(\angle_1^1-, \angle_5^1+, \angle_7^3-)$ is an impossible triple: the first 2 relations force the third opoint into the green area, which is contained in sections 1,2,3 of p_1, so p_3 cannot be in section 7 of p_1 (red).

■ **Algorithm 1** Upper bound for the $\mathcal{OPRA}_2^*$ composition table.

Result: an upper bound U on the set of realizable triples
$U \leftarrow$ all triples (r, s, t) of base relations
foreach *triple* (r, s, t) **do**
    if $t \notin f^{-1}(f(r) \diamond f(s))$ **then** remove (r, s, t) from U
    if $t \notin g^{-1}(g(r) \diamond g(s))$ **then** remove (r, s, t) from U
    if *not* `isAngleCombinationPossible(r,s,t)` **then** remove (r, s, t) from U
    if *not* `isLocationCombinationPossible(r,s,t)` **then** remove (r, s, t) from U
foreach *triple* (r, s, t) **do**
    if *(r,s,t) is not in* U **then**
        remove (r^-, t, s), (s, t^-, r^-), (s^-, r^-, t^-), (t, s^-, r) and (t^-, r, s^-) from U
return U

---

upper bound (e.g., by observing further constraints) and/or the lower bound (by extending the enumeration). An obvious candidate is $\mathcal{OPRA}_6^*$, in whose definition the quadrants from $\mathcal{OPRA}_2^*$ are replaced by twelfth-planes enclosing an angle of $30°$. Since that angle cannot be represented by integer ratios, our current enumeration, will no longer be complete, as it relies on integer arithmetics.

The success of our method on $\mathcal{OPRA}_2^*$ is largely due to two properties: (a) point-based calculi such as $\mathcal{OPRA}$ and $\mathcal{OPRA}^*$ exhibit a relatively simple and regular structure, which permits a complete geometric analysis such as to the one in [15]; (b) homomorphisms from $\mathcal{OPRA}_n^*$ to related calculi with established composition tables are easy to find. Whether our method yields useful results for calculi beyond the $\mathcal{OPRA}$ and $\mathcal{OPRA}^*$ families remains speculative and requires a thorough investigation of the previous two properties.

4 Conclusion

We presented our new generic analysis of the $\mathcal{OPRA}^*$ calculus family, which adds parallelism to the $\mathcal{OPRA}$ calculus family. Our analysis is based on combining condensed semantics lower bounds with upper bounds from algebraic mappings of related calculi. This for the first time enables sound standard QSTR constraint reasoning for $\mathcal{OPRA}^*$.

References

1 F. Dylla. *An agent control perspective on qualitative spatial reasoning: towards more intuitive spatial agent development.* PhD thesis, University of Bremen, 2008.

2 F. Dylla and J. H. Lee. A Combined Calculus on Orientation with Composition Based on Geometric Properties. In *Proc. of ECAI-10*, pages 1087–1088, 2010.

3 F. Dylla, J. H. Lee, T. Mossakowski, T. Schneider, A. van Delden, J. van de Ven, and D. Wolter. A Survey of Qualitative Spatial and Temporal Calculi: Algebraic and Computational Properties. *ACM Comput. Surv.*, 50(1):7:1–7:39, 2017.

4 F. Dylla and R. Moratz. Exploiting Qualitative Spatial Neighborhoods in the Situation Calculus. In *Spatial Cognition IV*, volume 3343 of *LNCS*, pages 304–322. Springer, 2004.

5 C. Freksa. Temporal Reasoning Based on Semi-Intervals. *Artif. Intell.*, 54(1):199–227, 1992.

6 C. Freksa. Using Orientation Information for Qualitative Spatial Reasoning. In *Proc. of GIS-92*, volume 639 of *LNCS*, pages 162–178. Springer, 1992.

7 A. Galton. *Qualitative spatial change.* Spatial information systems. Oxford University Press, 2000.

8 W. Liu and S. Li. On A Semi-Automatic Method for Generating Composition Tables. *CoRR*, abs/1105.4224, 2011. `arXiv:1105.4224`.

9 R. Moratz. Representing Relative Direction as a Binary Relation of Oriented Points. In *Proc. of ECAI-06*, pages 407–411, 2006.

10 R. Moratz. Qualitative Spatial Reasoning. In *Encyclopedia of GIS*, pages 1700–1707. Springer, 2017.

11 R. Moratz, D. Lücke, and T. Mossakowski. A condensed semantics for qualitative spatial reasoning about oriented straight line segments. *Artif. Intell.*, 175(16-17):2099–2127, 2011.

12 R. Moratz, J. Renz, and D. Wolter. Qualitative Spatial Reasoning about Line Segments. In *Proc. of ECAI-00*, pages 234–238. IOS Press, 2000.

13 Reinhard Moratz, Leif Sabellek, and Thomas Schneider. Granular Spatial Calculi of Relative Directions or Movements with Parallelism: Consistent Account. Technical report, University of Münster and University of Bremen, 2019. `http://publish.lycos.com/cosit2019/main/`.

14 T. Mossakowski and R. Moratz. Qualitative reasoning about relative direction of oriented points. *Artif. Intell.*, 180-181:34–45, 2012.

15 T. Mossakowski and R. Moratz. Relations Between Spatial Calculi About Directions and Orientations. *J. Artif. Intell. Res.*, 54:277–308, 2015.

16 D. Wolter, F. Dylla, S. Wölfl, J. O. Wallgrün, L. Frommberger, B. Nebel, and C. Freksa. SailAway: Spatial Cognition in Sea Navigation. *KI*, 22(1):28–30, 2008.