Report from Dagstuhl Seminar 21492

Representing and Solving Spatial Problems

Edited by

Pedro Cabalar[1], Zoe Falomir[2], Paulo E. Santos[3], and Thora Tenbrink[4]

1 **University of A Coruña, ES,** cabalar@udc.es
2 **Universitat Jaume I – Castello de la Plana, ES,** zfalomir@uji.es
3 **Flinders University – Adelaide, AU,** paulo.santos@flinders.edu.au
4 **Bangor University, UK,** t.tenbrink@bangor.ac.uk

Abstract

Everyday life takes place in space and time, and spatial experience lies at the heart of our existence. Understanding how we conceive spatial relations, and how we solve spatio-temporal problems, is therefore key to understanding human cognition. Spatial cognition research has advanced considerably over the past decades, with major successes particularly in computational implementations of knowledge representation and reasoning methods. Still, a range of key issues continue to pose major challenges. The goal of this report is to discuss the various options for the formalisation, implementation and automated solution of spatial problems including the following issues: the identification and specification of relevant concepts as expressed in human language; modules for automated understanding of domain descriptions; the use of spatial structures and affordances for direct spatial problem solving; and, the development of efficient planning systems capable of providing feasible solutions to spatial problems.

This report documents the program and the outcomes of Dagstuhl Seminar 21492 "Representing and Solving Spatial Problems".

Seminar December 5–10, 2021 – http://www.dagstuhl.de/21492
2012 ACM Subject Classification Applied computing; Human-centered computing; Information systems
Keywords and phrases knowledge representation, language analysis and cognitive processes, problem solving, spatial reasoning
Digital Object Identifier 10.4230/DagRep.11.11.1
Edited in cooperation with Homem, Thiago Pedro Donadon

We dedicate these discussions to the memory of Prof. Christian Freksa, who left his 'spatial' mark in everyone who knew him. His contribution to spatial cognition and spatial reasoning was key for our community. Also, his success in connecting researchers worldwide and his constant encouragement, mentorship, and support in our research endeavours provided us with endless inspiration.

Pedro, Zoe, Paulo, Thora

1 Executive Summary

Pedro Cabalar (University of A Coruña, ES)
Zoe Falomir (Universitat Jaume I – Castello de la Plana, ES)
Paulo E. Santos (Flinders University – Adelaide, AU)
Thora Tenbrink (Bangor University, UK)

Preamble: The idea of organising a Dagstuhl seminar on representing and solving spatial problems started with a discussion between Paulo Santos and the late Christian Freksa over a pair of Erdingers[1] during a previous Dagstuhl Seminar[2] in 2015. In this discussion we tried to solve, once and for all, the underlying reasoning strategies for spatial problem solving, whether it is logical-formal-symbolic, or it is totally/partially grounded on the real world (on what Christian called "the spatial substrates"). Considering the contributions in this volume, we believe the answer lies in the spectrum of approaches and that the right choice is dictated by the application domain. Indeed, one of the concrete contributions of this meeting was a discussion of possible problem scenarios that should constitute a Spatio-temporal Problem Repository to guide the future development of this field.

Summary of contributions: Spatio-temporal reasoning is a fundamental aspect of problem solving in general, and it occupies an ubiquitous place in AI and robotics, in the sense that almost all problem domains contain a spatio-temporal component. Paradoxically, despite its importance, the combination of spatial and temporal reasoning is one of the AI subfields with a wider gap between the current scientific progress and the human commonsense performance.

This seminar received contributions from researchers working in the various disciplines involved in investigating the problem of representing and reasoning about spatial problems (both in humans and artificial systems). The discussions were initially motivated by the following four main topics that are used below to organise the contributions in this collection:

1: **Representation**.
2: **Formalisation**.
3: **Description**.
4: **Problem solving**.

Representation

The discussions around the topic Representation were motivated by the following questions: How do humans conceptualise and mentally represent spatial problems? What is the role of high-level spatio-temporal structures for perceiving spatial problems, for manipulating spatial configurations, and for commonsense spatial problem solving?

In this volume, Cohn (Section 3.3) lists seven important problems that could be obstructing the development of automated commonsense spatial reasoning, including the lack of a proper definition of commonsense reasoning, the need for a foundational ontology of space, the

[1] One of which the former agent managed to spill all over the Schloss cellar floor!
[2] Dagstuhl Seminar 15192 – The Message in the Shadow: Noise or Knowledge?
`https://www.dagstuhl.de/en/program/calendar/semhp/?semnr=15192`

problem of how to tackle vagueness and implicit knowledge and the need for suitable default reasoning mechanisms for dealing with spatial information. The acquisition of commonsense knowledge and the role of embodiment in perception and spatial awareness were also pointed out as key open research questions. Complementarily, Kuhn (Section 3.9) advocates for the following three conceptualisations for spatio-temporal phenomena: (i) space-time varying fields of attributes, (ii) objects and object collections, (iii) events over fields and objects. Wörgötter (Section 3.22) goes one step up in the abstraction level suggesting the investigation of the very idea of concepts and how stimulus-driven experience drives the formation abstract thought. Starting from a distinct standpoint, Sioutis' arguments (Section 3.17) go in a similar direction with the proposal of a generic neurosymbolic framework for integrating qualitative spation-temporal reasoning with neural models from a probabilistic perspective. The idea for a "(symbolic) spatio-temporal knowledge base, naturally grounded on physics and human cognition" could be a good starting point for testing representation frameworks. Langley (Section 3.10) describes a cognitive architecture for embodied agents, whose recent extensions (towards more flexible representation of space and processes) could provide a fertile ground for the development of research into the core problems in the representation and reasoning about commonsense spatial information.

Formalisation

This topic was motivated by the investigation of what would be a suitable formalism for commonsense problem solving that allows an accurate, flexible and readable knowledge representation for spatio-temporal effects of actions performed by an intelligent agent.

Much work in spatio-temporal reasoning has focused on *Qualitative Spatial Reasoning* (QSR) [11], a field that attempts the formalisation of spatial knowledge based on primitive relations defined over elementary spatial entities. Although the combination of (qualitative) temporal and spatial reasoning is not infrequent in the literature in general (see for instance [1]), QSR approaches have traditionally overlooked a formal treatment of actions. Aiming to bridge the gap between commonsense reasoning, reasoning about actions and change and qualitative spatial representation and reasoning, some recent research has focused on spatial puzzles and games [2, 12, 3], as these domains offer a small number of objects requiring a minimum background knowledge about unrelated features, while they keep enough complexity to constitute a challenging problem of KR. This is in line with Cabalar's contribution in this volume (Section 3.2), where the challenging problem of evidence analysis of digital forensics is proposed as a challenging domain for the application of spatio-temporal reasoning strategies. Calabar also suggests a set of minimum requirements for a KRR formalism about common sense, which includes simplicity, natural understanding, clear semantics, computability and elaboration tolerance, where logical formalisms allowing non-monotonic reasoning is pointed as the best candidate to fulfil these requirements.

One of the most modern computational tools for non-monotonic reasoning is Answer Set Programming (ASP). Ludäscher (Section 3.12) describes two technical challenges in the ASP encoding of QSR reasoning systems, namely the difficulty in distinguishing spatial configurations using the natural ASP encoding of pairwise relations; and the exploration of the hierarchical structure of various QSR formalisms. While Ludäscher describes the latter issue from an implementation perspective, Stell (Section 3.18) suggests a research program to investigate the interaction with hierarchical models of discrete space and time representing changes at different levels of detail. The key idea here is to combine granular change with temporal change for supporting reasoning about effects of actions at different levels of detail.

From a more foundational standpoint, Borgo (Section 3.1) argues for the creation of a systematic program for the formal study of spatial notions, such as between, convex, simplex, parallel etc. This program should include the comparison of distinct perspectives for such modelling, facilitating the understanding of cognitive representation and reasoning. The notion of path is given as an example of how this systematic program could be developed.

Pease (Section 3.14) proposes the formalisation of everyday concepts and facts in a high-level classical logic engine called Suggested Upper Merged Ontology (SUMO). This formalisation would constitute a collection of real-world spatio-temporal problems that could be used to test the sufficiency of spatial knowledge representation strategies.

Description

This topic deals with the development of human readable descriptions of the inputs, reasoning steps and solutions of spatial problems. In particular, it addresses whether (and to what extent) it would be possible to develop high-level representations or interfaces for dealing with natural language and/or diagrammatic descriptions that allow specifying both the input knowledge and the output conclusions in terms of textual descriptions of spatial problems.

Research on spatio-temporal language highlights the range of meanings across contexts [13] as well as patterns of usage in relation to mental representations [14] and problem-solving [15] in various domains. Part of the insight gained in this research concerns the significance of what can (or will normally) *not* be represented in language. Often, non-verbalised concepts are those that are best understood through features of the spatial world itself. For instance, we may verbalise only the high-level goal of an everyday action (such as *dress up*), because every detailed action is represented through a combination of world knowledge with the affordances of the actual objects in question. There is no need to learn or conceptualise how to handle every possible instance of these objects, because the affordances of each exemplar are sufficiently clear to act upon as required. Some of these issues feature in the contributions of Dobnik (Section 3.4), Lopes (Section 3.11), Scheider (Section 3.16), Stock (Section 3.19) and Tenbrink (Section 3.20)

The following three semantic representational dimensions are listed by Dobnik (Section 3.4), as a summary of the core aspects of spatial descriptions studied in areas of linguistics, psychology, and computer science: (i) scene geometry, (ii) world knowledge and (iii) dynamic aspects of interaction. Dobnik's contribution also points out to a few open problems in this field, such as the lack of a unified computational language model that includes all of the cited dimensions, the limitation in the current research of grounding language in perception, that mostly deals with geometric perceptual context, and the need for experimental evaluation of models in real open domains.

Lopes (Section 3.11) suggests the use of the differences and interrelations between maps and navigation instructions to instantiate three cognitive core strategies of problem solving (attributed independently to McCarthy and Sloman), which are solving problems by following instructions, by descriptive knowledge representation and by analogical reasoning. Complementary, Stock (Section 3.19) argues that the current modelling of relative location descriptions is too restricted, not taking into account important contextual factors such as physical environment and objects in it; the observer's goals and expectations; the audience location and knowledge. Two main research advances were then proposed to push the boundaries of this field: (i) the incorporation of factors proved to be important in linguistic and cognitive science research and (ii) cognitive science investigation on some neglected factors that could be used in computational models, such as the influence of users' goals in the use of spatial relation terms. Tenbrink (Section 3.20) discusses the gap between human descriptions of routes and the state-of-the-art in the automatic generation of such descriptions to motivate three main challenges that deserve further attention: the combination of visual and verbal information, the consideration of change over time, and the flexibility in the use of various reference systems. Scheider (Section 3.16) contributes to this discussion suggesting the study of *transformation models* for concepts that are constitutive of spatial information, such as the concept of location, field, object, event and network. Arguing that transformations lie at the core of spatial reference systems, the author suggests that transformation models could be used to handle the variability of spatial information conceptualisations.

Problem Solving

The discussion about problem solving was initially motivated by questions about whether or not it would be possible (and desirable) to develop interfaces for dealing with spatial configurations including diagrammatic depictions and natural language descriptions to solve spatial puzzles in similar ways as humans do; and also, what are the commonsense problem solving capabilities involving spatio-temporal features including temporal explanation and planning under physical/geometric qualitative or semi-quantitative constraints. This issue also includes the investigation of appropriate problem solving algorithms and their potential applications to real-world domains that could be of interest to industry.

Calculi for qualitative temporal and spatial reasoning have a long history [4, 5]. More recently the relation between 2D and 3D structures of space (plus time) and their 1D formal descriptions have been investigated; the goal is a farther reaching exploitation of spatial structure for spatial problem solving. Cognitive architectures for robotic agents that make direct use of knowledge and spatial affordances in physical environments ('knowledge in the world') are currently being developed [10]. The knowledge in the mind of these agents controls their perception and action in ways that simplifies spatial configurations in order to reduce the difficulty of solving the spatial problem at hand, while knowledge about spatial relations is represented directly in perceivable and manipulable physical structures [9]. A principle underlying this approach is the use of mild abstraction [7] as employed in map navigation or in constructive geometry [8]. Much QSR and related reasoning research was originally motivated by the obvious discrepancy between logic-based metric computation (where spatial directions may be defined, for instance, by exact geometric angles and location specifications) and human concepts (such as "to the left"). Falomir's contribution (Section 3.5) starts from a discussion of the need to use cognitive heuristics for problem solving directly acting in the physical world, and argues for an effective combination of machine learning with automated reasoning for spatial problem solving. Spatial reasoning problems based on Perceptual/Differential Ability tests (PAT/DAT) are suggested as test domain for this development. Instead of assuming PAT/DAT as test domain, but still advocating the investigation of cognitive heuristics, Kroll (Section 3.8) describes recent results on the use of gaze heuristics for guiding an autonomous agent, and suggests the use of these ideas to dexterous assembly tasks; whereas Zachmann (Section 3.23) describes several issues pertaining the development of 3D geometric simulation methods for discovering affordances.

Hazarika (Section 3.6) discusses the use of mental maps defined over the space-time information structures underlying spatial problems (the *Spatial Substrate* [6]) for understanding how objects and their affordances affect the way humans reason about space. Following a similar line, Nath (Section 3.13) suggests the use of diagrammatic representations and reasoning to express the spatial problems and their physical substrate.

More explicit representation tools for problem solving were described by Homem (Section 3.7) and Santos (Section 3.15), in which they propose extending their previous work on applying case-based reasoning combined with QSR techniques towards complex real-world problems, such as real-time strategy games (Section 3.7) and collaborative mission planning for autonomous maritime vehicles (Section 3.15). Similarly, Wolter (Section 3.21) advocates the investigation of spatial problem solving techniques in dynamic domains (e.g. physical manipulation games) that allow for the progressive development of systems towards human-level capabilities.

Spatio-temporal Problem Repository

An issue that was of common agreement between the participants of the seminar is the need to create a Spatial Reasoning Problem repository. A few possible domain scenarios were discussed during the seminar, as listed below, but their full description is left for a future document.

1. Digital Forensics, cited in Section 3.2;
2. Puzzles such as Crazy Machines and Cut the Rope, cited in Section 3.2;
3. Angry birds, cited in Section 3.21
4. TPTP-style spatial problems, cited in Section 3.14
5. Perceptual/Differential Ability tests (PAT/DAT), cited in Section 3.5
6. Interpretation and generation of map descriptions (discussed during the seminar)
7. Multi step reasoning VQA (discussed during the seminar)
   - Ignore the images and concentrate on the scene graphs
   - Turn the scene graphs into logical formalism
8. From puzzle's descriptions to solutions (discussed during the seminar)
   - Input: Puzzles described as diagrams, or described with words or both
   - Output: A solution
   - it involves: Multiple step reasoning, combining diagrams, language and problem solving

References

1 Bennett, B., Cohn, A. G., Wolter, F., and Zakharyaschev, M. (2002). Multi-dimensional modal logic as a framework for spatio-temporal reasoning. *Applied Int.*, 17:239–251.

2 Cabalar, P. and Santos, P. E. (2011). Formalising the fisherman's folly puzzle. *Artificial Intelligence*, 175(1):346–377.

3 Cabalar, P. and Santos, P. E. (2020). Spatial reasoning about string loops and holes in temporal ASP. In Calvanese, D., Erdem, E., and Thielscher, M., editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 182–192.

4 Freksa, C. (1991). *Cognitive and linguistic aspects of geographic space*, chapter Qualitative spatial reasoning. Kluwer, Dordrecht.

5 Freksa, C. (1992). Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54(1):199 – 227.

6 Freksa, C. (2015). Computational problem solving in spatial substrates – A cognitive systems engineering approach. *Int. J. Softw. Informatics*, 9(2):279–288.

7 Freksa, C., Barkowsky, T., Dylla, F., Falomir, Z., Olteteanu, A.-M., and van de Ven, J. (2018a). Spatial problem solving and cognition. In Zacks J, T. H., editor, *Representations in Mind and World*, pages 156–183. New York: Routledge.

8 Freksa, C., Barkowsky, T., Falomir, Z., and van de Ven, J. (2018b). Geometric problem solving with strings and pins. *Spatial Cognition* & *Computation*. (in press).

9 Freksa, C., Olteteanu, A.-M., Ali, A. L., Barkowsky, T., van de Ven, J., Dylla, F., and Falomir, Z. (2016). Towards spatial reasoning with strings and pins. In *Fourth Annual Conference on Advances in Cognitive Systems*, volume 4.

10 Freksa, C., Vasardani, M., and Kroll, F. (2020). Dynamic problem solving in space. In Šķilters, J., Newcombe, N. S., and Uttal, D., editors, *Spatial Cognition XII*, pages 18–32, Cham. Springer International Publishing.

11 Ligozat, G. (2013). *Qualitative Spatial and Temporal Reasoning*. John Wiley & Sons.

12 Santos, P. E. and Cabalar, P. (2016). Framing holes within a loop hierarchy. *Spatial Cognition & Computation*, 16(1):54–95.

13 Tenbrink, T. (2007). *Space, time, and the use of language: An investigation of relationships*. Mouton de Gruyter, Berlin.

14 Tenbrink, T., Andonova, E., Schole, G., and Coventry, K. (2017). Communicative success in spatial dialogue: The impact of functional features and dialogic strategies. *Language and Speech*, 60:318–329.

15 Tenbrink, T. and Taylor, H. A. (2015). Conceptual transformation and cognitive processes in origami paper folding. *Journal of Problem Solving*, 8:2–22.

2 Table of Contents

3 Contributions

3.1 Path in mereology

Stefano Borgo (National Research Council – Povo (Trento), IT)

Main reference Stefano Borgo, Claudio Masolo: "Full Mereogeometries", Rev. Symb. Log., Vol. 3(4), pp. 521–567, 2010.
URL http://dx.doi.org/10.1017/S1755020310000110

The formal study of spatial notions (like between, convex, simplex, parallel and so on) whose expressivity lies on the landscape with mereology on the one hand and mereogeometry on the other, has brought interesting results in the last 30 years. Research shows that the mereogeometrical realm offers a rich variety of formal systems with respect to the systems existing within the Euclidean geometry approach. Notwithstanding, today our knowledge of these notions is still partial, the study of these systems lacks systematicity, and an overall research program has not been proposed. Spatial information within the mereological framework can be modeled from different perspectives. (I wrote "spatial information" but this observation applies more broadly, e.g., it can be stated for temporal and cognitive information.) Not surprisingly, many spatial notions that are exploited in common-sense reasoning have rich conceptual flexibility and rich expressivity power. The formal study of the alternative interpretations they allow and the systematic comparison of these alternatives would give us a wide range of formal systems which are, I believe, effective tools in exploring (human and beyond) cognitive representation and reasoning.

Recently, I have concentrated on one of these notions which I find very promising, namely, the notion of path. This notion has not received much attention in mereology even though it seems to be as old as our study of space. My interest on this notion has two motivations. First, it is well developed in mathematical topology and characterises almost all spaces encountered in analysis. Furthermore, all metric spaces are Hausdorff spaces. The rich structure that the notion of path allows to build in mathematics pushes for more attention. In comparison, we know almost nothing about how to understand path as a mereological notion. Second, it seems to me that path is a precursor (if not a precondition) to understand other notions that have been studied in mereogeometry starting from the very notion of congruence. My suggestion is to investigate our formal understanding of path moving away from the simple (and simplistic) mathematical intuition of path as "there is a continuity of positions from x to y". There are different ways to make sense of this intuition of "path" in mereology and they give rise to a variety of path-concepts.

Generally speaking, our cognitive understanding of space relies on a series of concepts which we use to build qualitative models. We know that each of these models relies on several, often implicit, modeling choices. One problem that we face is the mutual incompatibility of most of these models. Some claim that this incompatibility is in the nature of these representation systems. I would not be so sure. A deeper study of cognitive primitives, and of the possible readings that their informal understanding allows, would give us more flexibility in building and understanding these very models, and this is a first step toward the interconnection, if not integration, of what today we consider isolated and context-dependent systems.

3.2 Formalising Qualitative Spatial Knowledge: position paper

Pedro Cabalar (University of A Coruña, ES)

This position paper summarises the debate and drawn conclusions on the topic of Formalisation inside the Dagstuhl Seminar on Representing and Solving Spatial Problems (21492).

Introduction

One of the long-term goals of Artificial Intelligence (AI) is to endow computers with *common-sense reasoning*, a topic that dates back to the very beginning [10] of the area of *Knowledge Representation and Reasoning* (KRR). A device with common sense should be capable of making similar assumptions as those made by humans in ordinary situations of their daily life. These assumptions involve common knowledge about their physical environment, the behaviour of other agents, the existing interactions and their possible effects, all of them, crucial capabilities for the development of intelligent robotic systems. Our group in the University of A Coruña is specialised in KRR mostly under the paradigm of *Answer Set Programming* [2] (ASP), a popular logic programming formalism based on the *answer set* (or *stable model*) semantics [9] and well-suited for practical KRR and problem solving. Although great part of our research has been focused on theoretical foundations of ASP, we are known in the area for our results in extensions of this paradigm, especially with the addition of modal temporal operators [1], epistemic reasoning [4] and causal reasoning [5]. One of the goals for these fundamental results is the formalisation of dynamic or action domains, with the definition and interpretation of action languages and the study of different kinds of temporal problems such as prediction, explanation, planning or diagnosis. An application domain in which we became recently interested is the use of spatio-temporal reasoning for evidence analysis in *Digital Forensics*, being actively involved in an European COST action[3] on that topic. In collaboration with Paulo E. Santos, one line of research where we have also been particularly active is the formalisation of common sense reasoning about actions with *Qualitative Spatial Reasoning*. In particular, in a series of papers, [3, 6, 7, 8] plus [11, 12, 13], we have studied domains involving flexible objects like strings and make them interact with holed objects. Following a bottom-up methodology well-established in KRR, we have considered different families of puzzles involving strings, starting from simpler cases and gradually increasing the complexity of the operations required for solving the puzzle.

Discussion on Formalisation during the Seminar

One of the main four topics in the seminar was *formalisation* of Qualitative Spatial Reasoning (QSR). During the first round of presentations, I presented an introduction to KRR explaining some of the basic concepts and the usual scientific methodology followed in the area. For instance, among the desirable goals of a KRR formalism, we may include:

[3] *DigForASP: Digital Forensics – evidence Analysis via intelligent Systems and Practices*, European COST action CA-17124. `https://digforasp.uca.es/`

1. *simplicity*: the formalism should deal with a minimum amount of expressions and should allow compact descriptions
2. *natural understanding*: that is, some kind of correspondence with natural language or human methods of communication
3. *clear semantics*: an unambiguous correspondence between syntax and semantics, so that humans have a way of accurate communication with the machine
4. *computability*: the reasoning tasks associated to our language should be computable in an efficitent way, or at least, their computational complexity (or even decidability) should be known.
5. *elaboration tolerance* or *flexibility*: small changes in a problem should mean small changes in its representation in our KRR language.

The most usual candidate languages used in KRR that satisfy these goals up to a reasonable extent are *logical formalisms*, especially those allowing non-monotonic or default reasoning, if we are concerned about elaboration tolerance. Curiously, the goal of natural understanding (number 2) seems to require a variation in the case of QSR, since humans usually combine natural language with *diagrams* when they wish to communicate spatial knowledge. For this reason, the seminar included a thorough debate about diagrams as a representation tool. One of the conclusions was that diagrams may act as a perfectly valid KRR formalism containing both qualitative and sometimes quantitative information. Diagrams usually contain some features that provide accurate spatial information, whereas others are just illustrative: for instance, in a spreadsheet chart using bars, the height of the bar faithfully corresponds to some measured quantity, whereas its colour or width is usually selected for illustration. Similarly, a subway map just illustrates the existing connections, but distances have no direct correspondence with the real geographic distribution. Automated reasoning and knowledge representation can be performed on the diagram features that are actually representative.

Another part of the discussion about formalisation had to do with the usual methodology in KRR based on so-called *drosophila* examples. The name comes from an analogy proposed by Alexander Kronrod but made popular by John McCarthy: geneticists used the drosophila (or fruit fly) to study the genome because it is a simple organism but it keeps enough complexity to perform experimentation on the topic. In the same way, KRR in Artificial Intelligence has frequently used small examples (games, puzzles, small scenarios, etc) focused on some given feature under study. In many cases, the KRR formalisms have evolved by a successive incorporation of more complex features, covering new challenging examples unsolved before. An example of that methodology is our aforementioned work on puzzles with strings and holed objects. After some discussion about the adequacy of counting with challenging scenarios, an interesting proposal promoted by Sabine Timpf is preparing a *repository* with several groups of problems. This repository has been already started as an online board covering the categories: (1) problems with no representation or reasoning required; (2) problems with implicit representation required; (3) problems that require an explicit representation of knowledge; (4) problems that require representation and reasoning or problem solving algorithms.

One interesting feature that attracted much attention among the participants from different disciplines was the treatment of *affordances*, that is, finding a goal-driven utility for a given object or spatial configuration, possibly different from its normal or intended use. For instance, we know from movies that a high-heeled shoe can be used to drink champagne. Several possible drosophila have been discussed as challenging scenarios for experimentation on affordances, especially video games that require imaginative solutions (like *Crazy Machines* or *Cut the Rope*). As ongoing work, we expect to collect these examples in a larger repository and make it publicly available soon.

References

1 F. Aguado, P. Cabalar, M. Diéguez, G. Pérez, A. Schuhmann, and T. Schaub. Linear-time answer set programming. *Theory and Practice of Logic Programming*, 2021. to appear.

2 G. Brewka, T. Eiter, and M. Truszczyński. Answer set programming at a glance. *Communications of the ACM*, 54(12):92–103, 2011.

3 P. Cabalar and P. Santos. Strings and holes: an exercise on spatial reasoning. In *Proc. of IBERAMIA*, volume 4140 of *LNAI*, pages 419–429. Springer, 2006.

4 Pedro Cabalar, Jorge Fandinno, and Luis Fariñas del Cerro. Autoepistemic answer set programming. *Artificial Intelligence*, 289:103382, 2020.

5 Pedro Cabalar, Jorge Fandinno, and Michael Fink. Causal graph justifications of logic programs. *Theory and Practice of Logic Programming*, 14(4-5):603–618, 2014.

6 Pedro Cabalar and Paulo E. Santos. Formalising the Fisherman's Folly puzzle. *Artificial Intelligence*, 175(1):346–377, January 2011.

7 Pedro Cabalar and Paulo E. Santos. A qualitative spatial representation of string loops as holes. *Artificial Intelligence*, 238:1 – 10, 2016.

8 Pedro Cabalar and Paulo E. Santos. Spatial reasoning about string loops and holes in temporal ASP. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 182–192, 2020.

9 M. Gelfond and V. Lifschitz. The Stable Model Semantics For Logic Programming. In *Proc. of the* 5^{th} *International Conference on Logic Programming (ICLP'88)*, page 1070–1080, Seattle, Washington, 1988.

10 J. McCarthy. Programs with common sense. In *Proc. of the Teddington Conference on Mechanization of Thought Processes*, pages 75–91, 1959.

11 Paulo E. Santos and Pedro Cabalar. The space within Fisherman's Folly: Playing with a puzzle in mereotopology. *Spatial Cognition & Computation*, 8(1-2):47–64, 2008.

12 Paulo E. Santos and Pedro Cabalar. An investigation of actions, change, space within a hole-loop dichotomy. In *Proc. of the 11th Intl. Symp. on Logical Formalizations of Commonsense Reasoning (Commonsense'13)*, Ayia Napa, Cyprus, May 2013.

13 Paulo E. Santos and Pedro Cabalar. Framing holes within a loop hierarchy. *Spatial Cognition & Computation*, 16(1):54–95, 2016.

3.3 The Challenge of Automated Commonsense Spatial Reasoning

Anthony Cohn (University of Leeds, UK)

Joint work of Brandon Bennett, Anthony G Cohn
Main reference Brandon Bennett, Anthony G. Cohn: "Automated common-sense spatial reasoning: still a huge challenge". In Human-Like Machine Intelligence (ed.S. Muggleton and N. Chater), p. 405. Oxford University Press, 2021.
URL https://doi.org/10.1093/oso/9780198862536.003.0020

Achieving commonsense reasoning capabilities in a computational system has been one of the goals of Artificial Intelligence since its inception in the 1960s [32, 31, 43]. However, as Marcus and Davis have recently argued [30]: "Common sense is not just the hardest problem for AI; in the long run, it's also the most important problem". Moreover, it is generally accepted that space (and time) underlie much of what we regard as commonsense reasoning. For example, in the list of commonsense reasoning challenges given at `www-formal.stanford.edu/leora/commonsense/`, most of these rely crucially on spatial information.

From the 1990s onwards, considerable attention has been given to developing theories of spatial information and reasoning, where the vocabulary of the theory was intended to correspond closely with properties and relationships expressed in natural language but the structure of the representation and its inference rules were formulated in terms of computational data and algorithms [17, 14, 19, 18, 29, 27, 20, 46, 16, 15, 24, 34, 35] or in a precise logical language, such as classical first-order logic [38, 25, 6, 5, 7, 21, 37, 36, 9, 22].

However, despite a great number of successes in dealing with particular restricted types of spatial information, the development of a system capable of carrying out automated spatial reasoning involving a variety of spatial properties, of similar diversity to what one finds in ordinary natural language descriptions, seems to be a long way off. The lack of progress in providing general automated commonsense spatial reasoning capabilities suggests that this is a very difficult problem.

As with most unsolved problems, there are a variety of opinions about why commonsense spatial reasoning is so difficult to achieve and what might be the best approach to take. A point of particular contention, which is explored in detail in the full chapter [3], is: what is the role of natural language in relation to commonsense spatial reasoning?

The main purpose of the chapter is to help researchers orient and focus their investigations within the context of a highly complex multi-facetted area of research. We believe that research into computational commonsense spatial reasoning is sometimes misdirected for one or both of the following reasons: a) the goal of the research may incorporate several sub-problems that would be better tackled separately; b) the methodology of the research may assume that other related problems can be solved much more easily than is actually the case.

The chapter gives a fairly general (though not comprehensive) overview of the goal of automating commonsense reasoning by means of symbolic representations and computational algorithms. Previous work in the area will be surveyed, the nature of the goal will be clarified and the problem will be analysed into a number of interacting sub-problems. Key difficulties faced in tackling these problems will be highlighted and some possibilities for solving them will be proposed.

The chapter is structured in terms of the following list of what we consider to be the most important problems that are obstructing the development of automated commonsense spatial reasoning systems:

1. Lack of a precise meaning of "commonsense reasoning".
2. Difficulty of establishing a general foundational ontology of spatial entities and relationships.
3. Identification and organisation of a suitable vocabulary of formalised spatial properties and relations.
4. How to take account of polysemy, ambiguity and vagueness of natural language.
5. Difficulty of modelling the role of various forms of implicit knowledge (context, background knowledge, tacit knowledge).
6. Lack of a default reasoning mechanism suited to reasoning with spatial information.
7. Intrinsic complexity of reasoning with spatial information.

The main body of the chapter is largely negative in tone: pointing out the challenges in endowing machines with commonsense spatial reasoning and the problems listed above. Of course there has been progress towards this goal, and indeed we mention some of this in the chapter. Foremost in this direction is the work on qualitative spatial representation and reasoning (QSR). There are now a large number of QSR calculi capable of representing spatial information about (mereo)topology, direction, shape, distance, size among other aspects of spatial information. The computational complexity of reasoning with many of these calculi, at least the constraint languages associated with them, has been investigated thoroughly, and tractable subclasses identified (e.g. [39]). There are toolkits for reasoning with many of these, such as SparQ [45] and for extracting QSR representations from video data, e.g. QSRlib [23]. Moreover there are many implemented systems, particularly in the domain of activity understanding which exploit QSR (e.g. [13]) or which learn about spatial relations (e.g. [2]) from real world data. There is still though a disconnect between much of this work on QSR and the real problems of commonsense reasoning, as noted by [11]. Davis has contributed much to the field of commonsense reasoning, and spatial reasoning in particular e.g. his work on liquids [10] and containers [12].

There has also been work addressing the problem of how to acquire symbolic knowledge from perceptual sensors which are typically noisy and only incompletely observe the world, e.g. because of occlusion. Approaches in the literature which try to address these issues, include the use of formalisms which explicitly represent spatial vagueness such as [8], or ways of smoothing noisy detections (e.g [42]), building probabilistic models of QSR, e.g. [28], or by explicitly reasoning about occlusion, e.g. [4].

As is the case for AI in general, the more task/domain is constrained and well specified, the easier it is to come up with a (spatial) theory that is sufficient for appropriate reasoning and inference. The real challenge is to achieve general commonsense (spatial) reasoning.

In the chapter we decompose the problem of achieving automated commonsense spatial reasoning into a number of sub-problems (seven to be precise), which we consider to be key to solving the general problem, and are sufficiently independent from each other as to be addressed separately. Possibly, we have missed out further important problems, or conflated issues that would be best treated separately.

For example, one issue that we do not discuss much is how a commonsense knowledge could be acquired by an automated reasoning system, and in particular spatially related knowledge[4]. One approach, adopted by the CYC system already mentioned above is to manually specify such knowledge; the challenge here is the enormity of the knowledge and it is clear that despite several decades of research and development this remains an unfinished enterprise. The alternative is to try to acquire such knowledge via a process of learning. The NELL project [33] aims to learn such knowledge from text. An alternative is to learn from multimodal data, which has the advantage in simultaneously learning a semantic grounding in the perceptual world. For example [1] show how the meaning of object properties, spatial relations and actions, as well as a grammar, can be learned from paired video-text clips, while [40] demonstrate how the different senses of spatial prepositions such as in, above, against, and under can be acquired from human annotations in a virtual reality setting.

Another issue we hardly discuss is how *embodiment* affects perception and spatial awareness[5]. Tversky, among others, has discussed at length how embodiment affects the human reasoning: “Spatial thinking comes from and is shaped by perceiving the world and acting in it, be it through learning or through evolution” [44]. There is work in AI which takes an embodied approach to spatial cognition and spatial commonsense (e.g. [2, 41]) but more

[4] Use of automated knowledge extraction is potentially a way to circumvent problems that arise in several of the categories of difficulty that we identified. In particular problems 3 and 5 above and to some extent also 2 and 4.

[5] Embodiment could be regarded as being the source of a particular (perhaps especially significant) form of implicit/tacit knowledge (see problem 5 above). Though it does seem it also affects reasoning, perhaps by providing an unconscious way in which the constraints of space and time become apparent to us.

research on this is certainly needed. Possibly it is a way that humans and other intelligent animals can mitigate the problem of computational complexity (6) – their "embodiment" circumvents the need for spatial reasoning in many cases because we can directly experience the consequences of spatio-temporal properties and relationships.

Most of the problems we discuss actually apply to commonsense reasoning in general, rather than exclusively to spatial reasoning; and yet in the examples we have consider, it is primarily in the spatial aspects of semantics and reasoning where the difficulties lie. This is because the spatial domain is extremely rich and manifests huge variety and complexity. Issues relating to ambiguity vagueness are particularly apparent for spatial relationships because, although we have well-developed mathematical theories within which geometrical constraints can be precisely defined, there is no direct mapping from natural language terms to these precise constraints. And, even if these interpretative problems are circumvented, reasoning about space involves many highly intractable computations (though perhaps these go beyond the realm of commonsense).

Our analysis is not intended to be prescriptive of a particular research direction or methodology[6]. As well as exposing a large number of problems, we indicate a variety of different approaches that might lead to their solution. Our aim is primarily to provide an overview that would help researchers progress effectively by focusing their attention on some particular aspect of the highly complex problem of achieving automated commonsense spatial reasoning.

Acknowledgements

This research was partly inspired by prior collaboration by the first author with Vinay K. Chaudhri and Nikhil Dinesh at SRI International, while working on Project HALO. The second author was partially supported by the EPSRC under grant EP/R031193/1, and the EU under H2020 project 824619, and by a Turing Fellowship from the Alan Turing Institute. We also thank two anonymous referees for their helpful comments.

References

1 Alomari, Muhannad, Duckworth, Paul, Hogg, David C, and Cohn, Anthony G (2017). Natural language acquisition and grounding for embodied robotic systems. In *Proc. AAAI*.
2 Alomari, Muhannad, Li, Fangjun, Hogg, David C, and Cohn, Anthony G (2022). Online perceptual learning and natural language acquisition for autonomous robots. *Artificial Intelligence*, **303**, 103637.
3 Bennett, Brandon and Cohn, Anthony (2021). Automated common-sense spatial reasoning: still a huge challenge. In *Human-Like Machine Intelligence* (ed. S. Muggleton and N. Chater), p. 405. Oxford University Press.
4 Bennett, Brandon, Magee, Derek R., Cohn, Anthony G., and Hogg, David C. (2008). Enhanced tracking and recognition of moving objects by reasoning about spatio-temporal continuity. *Image and Vision Computing*, **26**(1), 67–81.
5 Borgo, Stefano, Guarino, Nicola, and Masolo, Claudio (1996). A pointless theory of space based on strong congruence and connection. In *Proc. KR-96*. Morgan Kaufmann.
6 Cohn, Anthony G. (1995). A hierarchical representation of qualitative shape based on connection and convexity. In *Proc. COSIT-95*, pp. 311–326.

[6] [11] suggest some research directions including the development of benchmarks and evaluation metrics, integration of different AI methodologies which have complementary strengths (e.g. facts gathered from web mining with mechanisms for formal reasoning), and a better understanding of human commonsense reasoning (as the second author has been attempting in robotic manipulation [26]

7 Cohn, A G, Bennett, B, Gooday, J, and Gotts, N (1997). RCC: a calculus for region based qualitative spatial reasoning. *GeoInformatica*, **1**, 275–316.

8 Cohn, A G and Gotts, N M (1996). The "egg-yolk" representation of regions with indeterminate boundaries. In *Proceedings, GISDATA Specialist Meeting on Geographical Objects with Undetermined Boundaries* (ed. P. Burrough and A. M. Frank), pp. 171–187. Francis Taylor.

9 Cohn, A G and Hazarika, S M (2001). Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, **46**(1-2), 1–29.

10 Davis, Ernest (2008). Pouring liquids: A study in commonsense physical reasoning. *Artificial Intelligence*, **172**(12), 1540 – 1578.

11 Davis, Ernest and Marcus, Gary (2015). Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM*, **58**(9), 92–103.

12 Davis, Ernest, Marcus, Gary, and Frazier-Logue, Noah (2017). Commonsense reasoning about containers using radically incomplete information. *Art. Int.*, **248**, 46 – 84.

13 Duckworth, Paul, Hogg, David C, and Cohn, Anthony G (2019). Unsupervised human activity analysis for intelligent mobile robots. *Artificial Intelligence*, **270**, 67–92.

14 Egenhofer, M (1991). Reasoning about binary topological relations. In *Proc. SSD'91*, pp. 143–160. Springer.

15 Escrig, M T and Toledo, F (1996). Qualitative spatial orientation with constraint handling rules. In *Proceedings ECAI*, pp. 486–490.

16 Faltings, B (1995). Qualitative spatial reaoning using algebraic topology. In *Spatial Information Theory: a theoretical basis for GIS* (ed. A. U. Frank and W. Kuhn), Volume 988, Lecture Notes in Computer Science, Berlin, pp. 17–30. Springer-Verlag.

17 Forbus, K. D., Nielsen, P., and Faltings, B. (1991). Qualitative spatial reasoning: The clock project. *Artificial Intelligence*, **51**, 417–471.

18 Frank, A (1992). Qualitative spatial reasoning with cardinal directions. *Journal of Visual Languages and Computing*, **3**, 343–371.

19 Freksa, C (1992). Using orientation information for qualitative spatial reasoning. In *Proc. Int. Conf. on Theories & Methods of Spatio-Temporal Reasoning in Geographic Space*. Springer.

20 Gahegan, M (1995). Proximity operators for qualitative spatial reasoning. In *Proc. COSIT*, Berlin, pp. 31–44. Springer Verlag.

21 Galton, A.P. (1998). Modes of overlap. *J. of Vis. Lang. & Computing*, **9**, 61–79.

22 Galton, Antony (2004). Multidimensional mereotopology. In *Proc. KR-04.*, pp. 45–54.

23 Gatsoulis, UY, Alomari, M, Burbridge, C, Dondrup, C, Duckworth, P, Lightbody, P, Hanheide, M, Hawes, N, Hogg, D C, and Cohn, A G (2016). QSRlib: a software library for online acquisition of qualitative spatial relations from video. In *29th International Workshop on Qualitative Reasoning*.

24 Gerevini, A. and Renz, J. (1998). Combining topological and qualitative size constraints for spatial reasoning. In *Proc. CP'98*. Springer.

25 Gotts, N M (1994). How far can we C? defining a "doughnut" using connection alone. In *Proc. KR9* (ed. E. S. J Doyle and P. Torasso). Morgan Kaufmann.

26 Hasan, Mohamed, Warburton, Matthew, Agboh, Wisdom C, Dogar, Mehmet R, Leonetti, Matteo, Wang, He, Mushtaq, Faisal, Mon-Williams, Mark, and Cohn, Anthony G (2020). Human-like planning for reaching in cluttered environments. In *2020 International Conference on Robotics and Automation (ICRA)*.

27 Hernández, D (1993). Reasoning with qualitative representations: Exploiting the structure of space. In *Proc. of the III IMACS Int. Workshop on Qualitative Reasoning and Decision Technology*, CIMNE, Barcelona, pp. 493–502. QUARDET '93.

28 Kunze, Lars, Burbridge, Chris, and Hawes, Nick (2014). Bootstrapping probabilistic models of qualitative spatial relations for active visual object search. In *2014 AAAI spring symposium series*.

29 Ligozat, G F (1993). Qualitative triangulation for spatial reasoning. In *Spatial Information Theory: A Theoretical Basis for GIS (Proceedings of COSIT'93)* (ed. A. U. Frank and I. Campari), Berlin, pp. 54–68. Springer-Verlag.

30 Marcus, Gary and Davis, Ernest (2019). *Rebooting AI*. MIT Press.

31 McCarthy, J (1989). Artificial intelligence, logic and formalizing common sense. In *Philosophical Logic and Artificial Intelligence* (ed. R. Tomason). Kluwer.

32 McCarthy, John and Hayes, Patrick J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligenge* (ed. B. Meltzer and D. Mitchie), Volume 4, pp. 463–502. Edinburgh University Press.

33 Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., Carlson, A., Dalvi, B., Gardner, M., Kisiel, B., and et al. (2018, April). Never-ending learning. *Commun. ACM*, **61**(5), 103 – 115.

34 Moratz, Reinhard, Lücke, Dominik, and Mossakowski, Till (2011). A condensed semantics for qualitative spatial reasoning about oriented straight line segments. *Artificial Intelligence*, **175**(16-17), 2099–2127.

35 Mossakowski, Till and Moratz, Reinhard (2012). Qualitative reasoning about relative direction of oriented points. *Artificial Intelligence*, **180-181**, 34–45.

36 Pratt, I. (1999). First-order qualitative spatial representation languages with convexity. *Spatial Cognition and Computation*, **1**, 181–204.

37 Pratt, Ian and Schoop, Dominik (1998). A complete axiom system for polygonal mereotopology of the real plane. *Journal of Philosophical Logic*, **27**, 621–658.

38 Randell, David A., Cui, Zhan, and Cohn, Anthony G. (1992). A spatial logic based on regions and connection. In *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, San Mateo, pp. 165–176. Morgan Kaufmann.

39 Renz, Jochen and Nebel, Bernhard (1999). On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the Region Connection Calculus. *Artificial Intelligence*, **108**(1–2), 69–123.

40 Richard-Bollans, Adam, Bennett, Brandon, and Cohn, Anthony G (2020). Automatic generation of typicality measures for spatial language in grounded settings. In *Proc. ECAI*.

41 Spranger, Michael, Suchan, Jakob, Bhatt, Mehul, and Eppe, Manfred (2014). Grounding dynamic spatial relations for embodied (robot) interaction. In *Pacific Rim International Conference on Artificial Intelligence*, pp. 958–971. Springer.

42 Sridhar, Muralikrishna, Cohn, Anthony G, and Hogg, David C (2011). From video to RCC8: exploiting a distance based semantics to stabilise the interpretation of mereotopological relations. In *Proc. COSIT*, pp. 110–125. Springer.

43 Thomason, R H (1991). Logicism, AI, and common sense: John McCarthy's program in philosophical perspective. In *Artificial Intelligence and Mathematical Theory of Computation: Essays in Honor of John McCarthy* (ed. V. Lifschitz), pp. 449–466. Academic Press.

44 Tversky, Barbara (2009). Chapter 12: Spatial cognition: Embodied and situated. In *Cambridge Handbook of Situated Cognition*. Cambridge University Press.

45 Wolter, Diedrich and Wallgrün, Jan Oliver (2013). Qualitative spatial reasoning for applications: New challenges and the sparq toolbox. In *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications*, pp. 1639–1664. IGI Global.

46 Zimmermann, K (1993). Enhancing qualitative spatial reasoning – combining orientation and distance. In *Proc. COSIT*, pp. 69–76.

3.4 Computational Generation and Interpretation of Spatial Descriptions

Simon Dobnik (University of Gothenburg, SE)

Spatial descriptions have been studied extensively in linguistics, computational linguistics, psychology, computer science, robotics and geo-information science. Theoretical work suggests that the following are the relevant semantic representational dimensions: (i) *scene geometry:* a two-dimensional or three-dimensional coordinate frame in which we can represent geometric objects and angles and distances between them; (ii) *world knowledge:* object conceptualisation and dynamic kinematic routines between objects; (iii) *dynamic aspects of interaction* such as *assignment of perspective* or *frame of reference*: directionals such as "to the left of" require a specification of a viewpoint, e.g. "from where I stand".

Physical sciences have developed ways in which space can be described with a high degree of accuracy, for example by measuring distances and angles. Such measures can be represented on a continuous scale of real numbers. However, humans refer to space quite differently: descriptions such as "the chair is to the left of the table" or "turn right at the next crossroad" refer to discrete units such as points, regions and volumes and require knowledge about how the objects related by a preposition interact with each other. Since spatial descriptions connect human conceptual and perceptual domains and distributions of perceptual features are associated with conceptual labels they are notoriously *vague*. Generation and understanding relies on mechanisms of *attention* determined from both the perceptual and linguistic communicative contexts and *contextual distractors*. Although the challenges for computational modelling of spatial language are sufficiently well-defined, to date there has been no unified computational language model that would include all aspects of their meaning.

Associating linguistic descriptions with perceptual representations is known as *grounding language in perception*. The majority of the current computational models used in situated agents only consider geometric perceptual context as a meaning component of spatial descriptions (see [3] for an example and references). Our earlier work focuses on individual components that could be applied in spatial language interpretation and generation. In [7, 8, 5, 13] we estimate the effect a world knowledge (i.e. functional vs geometric knowledge) on the use of spatial preposition from a large corpus of image descriptions. We also find the strength of association between different object pairs and prepositions and generalise categories of objects that prepositions occur with. Such model can be used in addition to the geometric model to identify the most natural preposition to be used with a particular pair of objects. In [5] we test the functional-geometric hypothesis as a function of perplexity of a neural language model which we train on a much larger corpus of spatial descriptions from image captions. In [4] we also examine if the functional-geometric bias of spatial relations is also expressed in the geometric arrangement of objects. In [6, 9] we examine how the spatial perspective is assigned in dialogue interaction between human conversational partners.

The deep learning approach using artificial neural networks has shown that they can successfully learn multimodal representations, in particular grounded language models in images in an unsupervised fashion in the domain of image captioning. Consequently, deep learning has the potential to be a useful approach to the problem of learning multimodal representations of the semantics of spatial descriptions [11, 12, 14]. Furthermore, deep learning describes a family of neural network models and there are some early examples in the literature of the application of neural networks to the problem of learning spatial representations [17, 2] which, however, have only been tested and implemented within a constrained experimental environment and participant responses rather than in real open domains.

The discussions in this seminar addressed the issues related to identifying the components of spatial semantics and how their representations can modelled both top-down and bottom-up. A central issue that has been identified is that of the role of *affordances* in spatial cognition which reminds us of the distinction made in [16]. It has been concluded that while there has been an important work done related to modelling spatial geometry (*where*) the effect of interacting objects and detection of affordances (*what*) is still one of the most challenging open questions. Interestingly, this direction leads spatial cognition away from its core geometric modelling to modelling of common sense world knowledge which can be expressed in *neural language models* and *ontologies*. Ontologies encode knowledge of their designer(s) top-down and this raises an interdisciplinary research question to what degree such information can be integrated in the bottom up data-driven systems (neural language models, visual embeddings), a problem that is commonly known as *information fusion*. Different spatial problems appear to rely on different sources of knowledge and therefore a need was identified to construct collections of spatial *problem sets* on which different systems can be studied and evaluated, similarly to the work on natural language inference [1]. Data-driven resources collected by the machine learning community such as the Visual Genome dataset [15] are also useful for constructors of rule-based systems and cognitive researchers as they provide examples of spatial language use and cognition. Here, the central questions are what aspects should be modelled, what are the basic modelling building blocks and what level of granularity of representation should be considered. An interesting connected discussion point was the relation between representations used to model spatial problems and descriptions in natural language. While formal spatial representations aim to be consistent and unambiguous, descriptions in natural language are underspecified (ambiguous) and frequently inconsistent. However, they are used in interaction between several participants who can exploit interactive mechanisms such as joint attention and clarification to resolve inconsistencies when they arise, for example to resolve reference [10].

References

1. Robin Cooper, Dick Crouch, Jan Van Eijck, Chris Fox, Johan Van Genabith, Jan Jaspars, Hans Kamp, David Milward, Manfred Pinkal, Massimo Poesio, et al. 1996. Using the framework. Technical report, Technical Report LRE 62-051 D-16, The FraCaS Consortium. Ftp://ftp.cogsci.ed.ac.uk/pub/FRACAS/del16.ps.gz.
2. Kenny R. Coventry, Angelo Cangelosi, Rohanna Rajapakse, Alison Bacon, Stephen Newstead, Dan Joyce, and Lynn V. Richards. 2005. Spatial prepositions and vague quantifiers: Implementing the functional geometric framework. In Christian Freksa, Markus Knauff, Bernd Krieg-Brückner, Bernhard Nebel, and Thomas Barkowsky, editors, *Spatial Cognition IV. Reasoning, Action, Interaction*, volume 3343 of *Lecture Notes in Computer Science*, pages 98–110. Springer Berlin Heidelberg.
3. Simon Dobnik. 2009. *Teaching mobile robots to use spatial words*. Ph.D. thesis, University of Oxford: Faculty of Linguistics, Philology and Phonetics and The Queen's College, Oxford, United Kingdom.
4. Simon Dobnik and Mehdi Ghanimifard. 2020. Spatial descriptions on a functional-geometric spectrum: the location of objects. In *Spatial Cognition XII: Proceedings of the 12th International Conference, Spatial Cognition 2020, Riga, Latvia*, pages 219–234, Cham, Switzerland. Springer International Publishing.
5. Simon Dobnik, Mehdi Ghanimifard, and John D. Kelleher. 2018. Exploring the functional and geometric bias of spatial relations using neural language models. In *Proceedings of the*

First International Workshop on Spatial Language Understanding (SpLU 2018) at NAACL-HLT 2018, pages 1–11, New Orleans, Louisiana, USA. Association for Computational Linguistics.

6 Simon Dobnik, Christine Howes, and John D. Kelleher. 2015. Changing perspective: Local alignment of reference frames in dialogue. In *Proceedings of goDIAL – Semdial 2015: The 19th Workshop on the Semantics and Pragmatics of Dialogue*, pages 24–32, Gothenburg, Sweden.

7 Simon Dobnik and John D. Kelleher. 2013. Towards an automatic identification of functional and geometric spatial prepositions. In *Proceedings of PRE-CogSci 2013 Production of referring expressions – bridging the gap between cognitive and computational approaches to reference at CogSci*, pages 1–6, Berlin, Germany.

8 Simon Dobnik and John D. Kelleher. 2014. Exploration of functional semantics of prepositions from corpora of descriptions of visual scenes. In *Proceedings of the Third V&L Net Workshop on Vision and Language at COLING*, pages 33–37, Dublin, Ireland. Dublin City University and the Association for Computational Linguistics.

9 Simon Dobnik, John D. Kelleher, and Christine Howes. 2020. Local alignment of frame of reference assignment in English and Swedish dialogue. In *Spatial Cognition XII: Proceedings of the 12th International Conference, Spatial Cognition 2020, Riga, Latvia*, pages 251–267, Cham, Switzerland. Springer International Publishing.

10 Simon Dobnik and Vera Silfversparre. 2021. The red cup on the left: Reference, coreference and attention in visual dialogue. In *Proceedings of PotsDial – Semdial 2021: The 25th Workshop on the Semantics and Pragmatics of Dialogue*, Proceedings (SemDial), pages 50–60, Potsdam, Germany.

11 Mehdi Ghanimifard and Simon Dobnik. 2017. Learning to compose spatial relations with grounded neural language models. In *Proceedings of IWCS 2017: 12th International Conference on Computational Semantics*, pages 1–12, Montpellier, France. Association for Computational Linguistics.

12 Mehdi Ghanimifard and Simon Dobnik. 2018. Knowing when to look for what and where: Evaluating generation of spatial descriptions with adaptive attention. In *Computer Vision – ECCV 2018 Workshops. ECCV 2018*, volume 11132 of *Lecture Notes in Computer Science (LNCS)*, pages 1–9, Proceedings of the Workshop on Shortcomings in Vision and Language (SiVL), ECCV 2018, Munich, Germany. Springer, Cham.

13 Mehdi Ghanimifard and Simon Dobnik. 2019a. What a neural language model tells us about spatial relations. In *Proceedings of the Combined Workshop on Spatial Language Understanding (SpLU) and Grounded Communication for Robotics (RoboNLP)*, pages 71–81, Minneapolis, Minnesota, USA. North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019), Association for Computational Linguistics.

14 Mehdi Ghanimifard and Simon Dobnik. 2019b. What goes into a word: generating image descriptions with top-down spatial knowledge. In *Proceedings of the 12th International Conference on Natural Language Generation (INLG-2019)*, pages 1–15, Tokyo, Japan. Association for Computational Linguistics.

15 Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, Michael Bernstein, and Li Fei-Fei. 2017. Visual Genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, 123(1):32–73.

16 Barbara Landau and Ray Jackendoff. 1993. "what" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16(2):217–238, 255–265.

17 Terry Regier. 1996. *The human semantic potential: spatial language and constrained connectionism.* MIT Press, Cambridge, Massachusetts, London, England.

3.5 Towards Bridging the Gap between Spatial Reasoning and Machine Learning for Solving Spatial Problems

Zoe Falomir (Universitat Jaume I – Castello de la Plana, ES)

Everyday we are solving spatial problems (i.e. finding a new address in a map, folding clothes, solving puzzles with your kids, etc.). These problems appear in our real physical space and we are solving them there. According to [5] a spatial problem is: (1) A question about a given spatial configuration (of arbitrary physical entities) that needs to be answered. For example: is there water in a glass? (2) The challenge to construct a spatial configuration with certain properties from a given spatial configuration. For example, playing tic-tac-toe.

How do we humans solve problems? As Helie and Pizlo mentions [7]: *"The travelling salesperson (TSP) problem with only 20 cities has 19!/2 tours, which is about 1017 possible states. The number of neurons in the brain is one million times smaller than this number. And our working memory (WM) can only hold and manipulate 4-7 items at a time. But a human subject produces a TSP tour in 60 seconds."* So, if we humans have less memory than a computer and our biological neurons are slower than the digital switches in computers, how do we humans solve problems on a daily basis in a so fast and efficient way? Helie and Pizlo [7] concludes that humans do not typically engage in exploring even a small fraction of the problem space, that *"we humans build a problem representation and solve the represented problem, not the problem that is out there"*, that is, in the real physical space. So humans may use heuristics to solve spatial problems. For example, the gaze heuristic [6] allowed an effective behaviour in agents catching a falling ball without using any dynamic model of the flying object (more details by F. Kroll [11] in this volume).

Strategies. Two main types of problem solving strategies were distinguished by [5]:

- Direct problem solving: operating on the real space. Others [3] may name this direct problem solving as zero-shot learning,
- Abstract problem solving: representing the problem, using this representation for reasoning and instating the solution to the real space.

Sometimes solving the problem by directly acting in the physical space is quicker than representing the problem and solving it by computing on the representations. According to [3], we can find examples of zero-shot learning observing animals solving spatial problems. That is because they have the ability to break down complex problems into simpler, previously learned sub-problems and this hierarchical approach allow humans and some intelligent animals to solve previously unseen problems in a zero-shot manner, that is, without any trial and error.

Proposal. Knowledge based techniques allow us to represent/abstract and reason to solve spatial problems and machine learning techniques allow agents to learn directly from space, that is, from the data gathered while interacting with it through their actions. How to combine machine learning with automated reasoning for validating and improving strategies for spatial problem solving is the idea behind this position paper.

Methods. In cognitive psychology, human abilities to solve spatial problems are measured by perceptual/differential aptitud tests (PAT/DAT), some of them are: paper-folding-and-punching tests, mental rotation tests, 3D object perspective tests, etc. In the literature, there are studies that have shown that people with better spatial reasoning abilities are more successful in Science, Technology, Engineering and Math (STEM) disciplines [13], they are also more creative and produce more patents and scientific innovations [10]. Moreover, some studies by Sorby [14] demonstrated that spatial abilities can be trained at any age.

So, when developing tools for people to train their spatial abilities we need to present them with spatial problems and: (i) show them clues about how to solve them or/and (ii) provide feedback to them when they got wrong solutions.

As an example of spatial problem, we took the paper-folding-and-punching tests and we developed a representation for the problem and an automated reasoning algorithm which later we tested it in a virtual space (videogame) [4]. The master mode in this videogame can create automatically instances of questions/problems and answers/solutions using the representation and the reasoning method embedded in it. So the question-answered produced are random and can be used to train human spatial skills.

The main aim would be to define an artificial agent that learns the reasoning behind paper folding by interacting with the game. In the learning process, it would be interesting to study if there are differences between defining the starting point for the agent to learn as the virtual space or the representation. May the representation act as a cognitive heuristic? If this expert representation takes the form of a knowledge graph (KG) or an ontology, does it allow comparison with the knowledge learnt from the data-driven system to allow validation? In most cases, the knowledge extracted from machine learning algorithms are not explicable or understandable by users. May a comparison with the expert knowledge graph provide a possible explanation of the learning process? If so, how do we evaluate the adequacy/rationality of these explanations?

Seminar Discussion. As Stefano Borjo [1] stated in this volume, one problem that we face in spatio-temporal reasoning is the mutual incompatibility of some qualitative models which have been build on different representation systems choosing different modelling choices. There is a need to integrate these models, which are nowadays very context-dependent systems. Our proposal is to study algorithms of machine learning in a context and extract the knowledge-graph (KG) that the system has learnt and study if that KG is solving a specific problem or an approximate generalisation of all the problems which can provide a hint about how to integrate the models that produce their solutions.

As Simon Dobnik [2] discussed in this volume, spatial representations used to solve spatial problems can be modelled top-down (by experts using ontologies) and bottom up by data-driven systems (neural language models learn the spatial representations). The challenge here is to which degree top-down models can be integrated with data-driven models. This is also known by the *information fusion* problem.

In this seminar, T. Homem and P. Santos [8, 9] also discussed their algorithm for Qualitative Case-Based Reasoning and Learning (QCBRL) which is a case-based reasoning system that uses EOPRA model (cardinal directions with local distances [12]) to retrieve and reuse cases combined with reinforcement learning to allow the agent to learn new qualitative cases at runtime, without assuming a pre-processing step. This work is an example about how to integrate a spatial reasoning model with reinforcement learning technique and it will inspire us towards solving the challenge presented in this position paper.

D. Wolter [15] also presented AI Birds in this seminar: a physical manipulation game for conducting research in spatial problem solving and for evaluating the progress made towards human capabilities. The agents deal with a large action space, have incomplete knowledge about the physical parameters of objects and thus the consequences of possible actions can only be estimated. Machine learning approaches have been largely unsuccessful maybe due to the limited number of Angry Birds levels for training. This game is another example that provide evidence of how games have been used as a microcosm for conducting AI research and it will inspire us in our challenge presented in this position paper.

Seminar Conclusion. A repository of spatial problems (including insight problems) is needed to study procedures that lead to solutions that arise either in a direct way (applying zero-shot learning/reasoning) or exploiting a representation that allows effective reasoning/learning mechanisms. This repository will allow to compare computational procedures and methods among them and to human performance.

Acknowledgements. I want to thank the Schloss Dagstuhl – Leibniz Center for Informatics for hosting our Seminar 21492 on Representing and Solving Spatial Problems. And I also acknowledge the funding provided by the Ramon y Cajal fellowship (RYC2019-027177-I / AEI / 10.13039/501100011033) awarded by the Spanish Ministry of Science, Innovation and Universities and the research project PDC2021-121097-I00.

References

1 Stefano Borjo. Path in mereology. In Pedro Cabalar, Zoe Falomir, P. E. Santos, and Thora Tenbrink, editors, *Report from Dagstuhl Seminar 21492: Representing and Solving Spatial Problems*, this volume. Dagstuhl Reports, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, 2022. `doi:10.4230/DagRep.11.11.1`.

2 Simon Dobnik. Computational generation and interpretation of spatial descriptions. In Pedro Cabalar, Zoe Falomir, P. E. Santos, and Thora Tenbrink, editors, *Report from Dagstuhl Seminar 21492: Representing and Solving Spatial Problems*, this volume. Dagstuhl Reports, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, 2022. `doi:10.4230/DagRep.11.11.1`.

3 M. Eppe, C. Gumbsch, M Kerzel, P. Nguyen, Martin Butz, and S. Wermter. Intelligent problem-solving as integrated hierarchical reinforcement learning. *Nat Mach Intell*, 4:11–20, 2022. URL: `https://doi.org/10.1038/s42256-021-00433-9`.

4 Zoe Falomir, Ruben Tarin, Aurelio Puerta, and Pablo Garcia-Segarra. An interactive game for training reasoning about paper folding. *Multim. Tools Appl.*, 80(5):6535–6566, 2021. `doi:10.1007/s11042-020-09830-5`.

5 C. Freksa, T. Barkowsky, F. Dylla, Z. Falomir, A-M. Olteţeanu, and J. van de Ven. Spatial problem solving and cognition. In H. Taylor J. Zacks, editor, *Representations in Mind and World*, pages 156–183. Routledge, Taylor and Francis, New York, 2018.

6 Gerd Gigerenzer and Henry Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1:107 – 143, 01 2009. `doi:10.1111/j.1756-8765.2008.01006.x`.

7 Sebastien Helie and Zygmunt Pizlo. When is psychology research useful in artificial intelligence? a case for reducing computational complexity in problem solving. *Topics in Cognitive Science*, n/a(n/a), 2021. `doi:https://doi.org/10.1111/tops.12572`.

8 Thiago Homem. Transfer learning through qualitate spatial reasoning. In Pedro Cabalar, Zoe Falomir, P. E. Santos, and Thora Tenbrink, editors, *Report from Dagstuhl Seminar 21492: Representing and Solving Spatial Problems*, this volume. Dagstuhl Reports, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, 2022. `doi:10.4230/DagRep.11.11.1`.

9 Thiago P. D. Homem, Paulo E. Santos, Anna Helena Reali Costa, Reinaldo Augusto da Costa Bianchi, and Ramon Lopez de Mantaras. Qualitative case-based reasoning and learning. *Artificial Intelligence*, 283:103258, 2020. `doi:https://doi.org/10.1016/j.artint.2020.103258`.

10 HJ Kell, D Lubinski, CP Benbow, and JH Steiger. Creativity and technical innovation: spatial ability's unique role. *Psychol Sci*, 24(9):1831–6, 2013. doi:10.1177/0956797613478615.
11 Felix Kroll. Strong spatial strategy retrieval in the context of an assembly task. In Pedro Cabalar, Zoe Falomir, P. E. Santos, and Thora Tenbrink, editors, *Report from Dagstuhl Seminar 21492: Representing and Solving Spatial Problems*, this volume. Dagstuhl Reports, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, 2022. doi:10.4230/DagRep.11.11.1.
12 R. Moratz and J.O. Wallgruen. Spatial reasoning with augmented points: Extending cardinal directions with local distances. *Journal of Spatial Information Science*, 5(2012):1–30, 2012. doi:10.5311/JOSIS.2012.5.84.
13 N. Newcombe. Picture this: Increasing math and science learning by improving spatial thinking. *American Educator*, 34(2):29–35, 2010.
14 Sheryl A. Sorby. Educational research in developing 3D spatial skills for engineering students. *International Journal of Science Education*, 31(3):459–480, 2009. doi:10.1080/09500690802595839.
15 Diedrich Wolter. Spatial problem solving in physical manipulation games. In Pedro Cabalar, Zoe Falomir, P. E. Santos, and Thora Tenbrink, editors, *Report from Dagstuhl Seminar 21492: Representing and Solving Spatial Problems*, this volume. Dagstuhl Reports, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, 2022. doi:10.4230/DagRep.11.11.1.

3.6 Spatio-Temporal Reasoning Over Diagrams Exploiting Object Affordance

Shyamanta Moni Hazarika (Indian Institute of Technology Guwahati – Assam, IN)

Main reference Chayanika Deka Nath, Shyamanta M. Hazarika: "Activity recognition in video sequences over qualitative abstracts of a diagram-based representation schema", J. Vis. Commun. Image Represent., Vol. 76, p. 103061, 2021.
URL http://dx.doi.org/10.1016/j.jvcir.2021.103061
Main reference Rupam Bhattacharyya, Shyamanta M. Hazarika: "Object Affordance Driven Inverse Reinforcement Learning Through Conceptual Abstraction and Advice", Paladyn J. Behav. Robotics, Vol. 9(1), pp. 277–294, 2018.
URL http://dx.doi.org/10.1515/pjbr-2018-0021

My primary interest is in Robotic Neurorehabilitation. This translates into interest in qualitative spatio-temporal reasoning; particularly reasoning about actions for daily living activities. How do objects and their affordances govern the way people reason and interact with their environments? I am interested in exploring the use of mental maps or diagrams directly over the spatial substrate for solving such problems, particularly exploiting object affordance. Work I am conducting on understanding the neural correlates of affordance reasoning may be relevant.

For human beings, spatial reasoning is a particularly powerful and accessible mode of cognition. In our everyday interaction with the physical world, spatial reasoning appears to be driven by qualitative abstractions rather than complete quantitative knowledge a priori. Space and time are inextricably linked. Connection between time and space has been a recurring topic, initially in geography, and more recently within AI. Reasoning about space often involves reasoning about change in spatial configurations. There is a requirement to build into the spatial representations which changes respect the underlying continuous nature of change. Formalizing notions of continuity for a theory of spatial change / motion holds promise.

Human basically convey effective solutions based on mental maps or spatial organization of problems. Mental Maps exploit the spatial substrate, i.e., space − time information structures underlying a spatial problem. Diagrams for representation and reasoning narrows the option of ambiguity in relational composition. Freksa introduced the need of comparison between formal and Diagrammatic Reasoning processes for the same underlying problems. Diagrammatic Reasoning and Qualitative Spatio-temporal Reasoning needs to be revisited from the perspective of not only Cognitive Vision but other areas involving space and action including Language Understanding. There remains immense potential for research advances in the use of diagrams as a representation and reasoning paradigm. Object affordance, a characterization of the different functionalities of an object, refers to the numerous possibilities of interaction with the object. This forms an implicit background of our everyday spatio-temporal reasoning. An object affordance driven novel pipeline have allowed us to demonstrate that symbolic conceptual abstraction helps to curb curse of dimensionality present in high dimensional demonstration for learning. Object affordance can directly link to higher level behaviours such as intents. It has a significant part in driving understanding of action verbs and priming motoric actions. Investigating the influence of affordance driven motor priming holds promise for design of Intelligent Assistive Devices and is part of ongoing research within the Biomimetic Robotics and Artificial Intelligence Lab at IIT Guwahati.

3.7 Transfer Learning Through Qualitate Spatial Reasoning

Thiago Pedro Donadon Homem (IFSP – Sao Paulo, BR)

Main reference Thiago Pedro Donadon Homem, Paulo Eduardo Santos, Anna Helena Reali Costa, Reinaldo Augusto da Costa Bianchi, Ramón López de Mántaras: "Qualitative case-based reasoning and learning", Artif. Intell., Vol. 283, p. 103258, 2020.
URL http://dx.doi.org/10.1016/j.artint.2020.103258

The ability of an agent to learn, adapt and solve new problems is one of the big challenge of research in Artificial Intelligence. In some cases, the domains are too complex or the input data provided by the sensors is purely quantitative, which creates a complex problem and the agent is not able to learn or, in the best case, it learns too slowly.

Especially in problems that contain spatial information, Qualitative Spatial and Temporal Reasoning abstracts the metric information and transforms the problem into a kind of human-like representation and reasoning. This could create generalizations and solutions to the learning and to the transfer learning problems.

The purpose of this work consists of two phases: 1) abstracting the metric information into qualitative spatial relationships and creating a qualitative machine learning model in a source domain; 2) transfer the qualitative model to a target domain.

The proposal evaluation includes model metric evaluation as follows: in phase (1) the qualitative learning model must perform on the target domain as well as when trained with quantitative data and in phase (2) the qualitative learning model must perform in the source domain as well as running in the target domain or by running the learning phase directly on the target domain or by running the learning phase directly on the target domain with quantitative data.

This proposal extends the work presented in [1] but with the new focus on qualitative learning and transfer learning. In terms of qualitative representation and reasoning, Elevated Oriented Point Relation Algebra (EOPRA) [2] or Region Connection Calculus (RCC8) [3] are considered to be used. For the task of qualitative learning and transfer learning, Deep Learning is considered, such as the works of [4] and [5].

The application domain is still uncertain, but probably robot soccer or real-time strategy games will be chosen, allowing the proposal to be expanded in real-world problems.

References

1 Thiago Pedro Donadon Homem, Paulo Eduardo Santos, Anna Helena Reali Costa, Reinaldo Augusto da Costa Bianchi, and Ramón López de Mántaras. Qualitative case-based reasoning and learning. *Artif. Intell.*, 283:103258, 2020.

2 Reinhard Moratz and Jan Oliver Wallgrün. Spatial reasoning with augmented points: Extending cardinal directions with local distances. *Journal of Spatial Information Science (JOSIS)*. 5. 1-30. 10.5311/JOSIS.2012.5.84.

3 David A. Randell and Zhan Cui and Anthony G. Cohn. A Spatial Logic based on Regions and Connection. *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*. Morgan Kaufmann. 1992.

4 Nikhil Krishnaswamy, Scott Friedman, and James Pustejovsky. Combining Deep Learning and Qualitative Spatial Reasoning to Learn Complex Structures from Sparse Examples with Noise. *AAAI'19/IAAI'19/EAAI'19: Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*. 2911 – 2918. 2019.

5 Vinyals, O., Babuschkin, I., Czarnecki, W.M. et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575, 350 – 354. 2019.

3.8 Strong Spatial Strategy Retrieval in the Context of an Assembly Task

Felix Kroll (Technische Hochschule Deggendorf, DE)

Strong Spatial Strategy

The gaze heuristic described by Gerd Gigerenzer [5] has proven to be a very effective strategy to catch a flying object [3] in terms of generalization of the object trajectories that are catchable. Common analytical approaches are bound to their modeled dynamics and thus have a big generalization gap due to the vast differences in object properties like air drag or Magnus forces and external factors like side winds. In contrast to these, the gaze heuristic is not bound to any dynamic model of the flying object. It just requires the agent to establish an geometric alignment based on the angle of observation towards the object and constant applied correcting movements.

The gaze heuristic can be categorised as a *strong spatial strategy* in the sense of the strong spatial cognition paradigm [2], because it offloads cognition into the environment rather than reason with an internal representation. A reduced set of immediate observations is taken to control the agents actions. The continuous observations are directly mapped to a corresponding action. The work of Christian Freksa [2] describes other examples that fall into this category, for example to distinguish the distance between objects by concentrating on their parallax effect and making a lateral movement.

We believe that these kinds of strategies are the key to enable time and energy efficient spatio-temporal problem solving for all kinds of organisms. Furthermore does it allow them to react robustly on unseen situations because the strategies are not dependent on a (full) model of the problem dynamics. In the case of the gaze heuristic the only assumption that needs to hold is the continuity of the flying object. The work of Robert P. Hamlin [6] showcases that in nature this assumption is key to intercept prey or – by utilising it – trick the hunters and escape their grasp. It holds for high enough sampling rates and therefore is decoupled from the objects dynamics.

Assembly Task

We think that a field where those strong spatial strategies could be particularly beneficial and natural is the one of dextrous assembly task. These are tasks where a bigger construct or product is built from a number of smaller parts by grasping and combining them. In this context, a typical and seemingly simple sub-task is the problem of re-orienting an object (e.g. bolts, plugs, packages or screws) "in-hand" while it is held by the manipulator, i.e., one has to change the orientation of the object relative to the manipulator.

While there are many different solutions to this problem, e.g. make use of a second manipulator or perform in-hand-manipulation with a multi-finger-gripper, a common approach in robotics is to utilise fixtures in the environment, e.g., a table to perform the re-orientation. Most current robot-manipulators are two-finger-grippers. Hence in-hand-manipulation is not applicable and we will focus in this discussion on a two finger scenario which requires fixtures to solve the task.

From a human perspective a strategy can be described as follows: If the goal is to re-orient a almost vertical bolt held between two fingers to be in a horizontal orientation, the human would likely push the bolt onto a table until the bolt's orientation is the same as the table horizontal surface. The human would perform a parabola like downwards movement until the the lower end of the bolt touches the table and then continue the movement until finally the fingers touch the table. During this movement the bolt would be kept between the fingers with the minimal required pressure so it can rotate but not fall down. The touch input is a strong indication about the state of the strategy. It indicates when to adjust the grasp pressure and when the strategy is completed. Since the bolt is rigid and the table too, the bolt can not penetrate the table or get stuck assuming both have a low friction coefficient. The table affords a re-orientation by its physical properties if the bolt is pushed against it with the given constraints of the two finger grasp.

In robotics this strategy can be replicated with the common analytical modelling approach: It would require the bolt's, table's and gripper's shape in form of meshes. Also their physical properties need to be known; the centre of mass and their friction coefficients. The aim is to use a planner that searches continuously for a movement trajectory of the arm to perform the strategy. In order to use the planner one would use the given properties of the environment to replicate it in form of a physics simulation which must be capable of solving multi-contact situations between the table, the bolt and the gripper. This simulation would be utilised in the motion planner. In order to keep the simulation aligned with reality one would have to track the bolt's and the gripper's orientation-state with a camera or object tracking system. This approach does not scale well to objects with varying geometries and physical properties, and in terms of computational costs.

When looking at this problem from a strong spatial perspective, these limitations can be avoided. The bolts state would be an angle between the table and itself. The torque sensors of a torque controlled robotic arm could be used to register the contact with the fixture surface. From there, the robot could perform the same movement algorithm as the human until it would also register a force feedback or the bolt reaches the desired orientation angle. This strategy could be applied to many bolts with varying size, weight and friction coefficients because the angular state is independent of the object's physical properties and the force feedback signals the full alignment of the bolt with the surface's orientation. Further, the mapping of measurements to actions would be far simpler to compute because it avoids to create a replica of the world that must be kept up to date.

There are still many challenges to solve and this strategy is not as well understood as the gaze heuristic, but it displays how well it circumvents the modeling-challenges of classical model-based strategies. There is no need for a model and the object's representation can be reduced to a 2D visual orientation of the object because it is rigid. We assume, that the measurement of a torque controlled robotic arm and one external mounted camera is enough to perform the strategy.

Further, this re-orientation problem would also be of interest to do research on the question of if and how humans use strong spatial strategies in the field of dextrous assembly tasks. This would be similar to the gaze heuristic, where observing football players [1, 7] or dogs [8] helped in deriving and understanding it.

Seminar Discussion

We utilise the strong spatial cognition paradigm as a framework to find strategies for manipulation tasks in robotics. Human behaviour might lead to new strong spatial strategies. However, keeping the idea of affordances in mind (in the sense of James J Gibson [4]), we can also engineer new strong spatial strategies from the ground up.

While we believe that we can construct autonomous agents which act purely based on strong spatial strategies, we face often problems where the most useful observation for applying it is not available or disturbed, e.g. in the case of of the gaze heuristic if no geometric relation to the ball can be established due to the optical sensor's excessive exposure to the sun. In this scenario we can fall back to spatial reasoning and constructing a (minimal) spatial representation to get back on track. Furthermore, there are scenarios where the agent can not act inside the environment and thus no strong spatial strategy can be applied. In this case there can be still observations which the agent can utilise to plan a head or come up with a solution by reasoning. Such observations are retrieved form external sources or own experiences, e.g. street maps or from other agents in form of speech or the agents memories. Here (analytical) reasoning is the only option since there are only indirect observations of the spatio-temporal environment accessible.

A key result of the seminar is that there is no free lunch for spatial cognition. Sometimes it is hard to find a spatial representation to apply reasoning. In those cases we could look for a strong spatial strategy. But other times the spatial representation is easy to construct and the observations are not available in rapid intervals like many strong spatial strategies require them to be. Here, reasoning provides a more robust solution. We propose, that the spatial cognition research should shift its attention to combine strong spatial strategies with analytical predictions to circumvent representational bottlenecks without sacrificing computational and representational efficiency and a sense for the outcome (knowledge of uncertainty).

This ambivalence leads the seminar's participants to conclude that a (spatio-temporal) problem repository is necessary to find robust unified representations, reasoning methods or strong spatial strategies by testing and evaluating them in the problem domain of the repository. The robotic assembly task, like the here mentioned bolt re-orientation, could be part of such a problem repository. However, it is not yet clear in which description format the problems can be defined formally.

References

1 Seville Chapman. Catching a baseball. *American Journal of Physics*, 36:868–870, 10 1968.
2 Christian Freksa. Strong spatial cognition. In Sara Irina Fabrikant, Martin Raubal, Michela Bertolotto, Clare Davies, Scott Freundschuh, and Scott Bell, editors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9368, pages 65–86. Springer International Publishing, 2015.
3 Christian Freksa, Maria Vasardani, and Felix Kroll. Dynamic problem solving in space. In *Spatial Cognition 2020*, pages 18–32. Springer International Publishing, 2020.
4 James J Gibson. *The theory of affordances.* Lawrence Erlbaum Associates, 1977.
5 Gerd Gigerenzer and Henry Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1:107–143, 2009.
6 Robert P. Hamlin. "the gaze heuristic:" biography of an adaptively rational decision process. *Topics in Cognitive Science*, 9:264–288, 4 2017.
7 Peter McLeod and Zoltan Dienes. Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance*, 22:531–543, 1996.
8 Dennis M. Shaffer, Scott M. Krauchunas, Marianna Eddy, and Michael K. McBeath. How dogs navigate to catch frisbees. *Psychological Science*, 15:437–441, 7 2004.

3.9 Core Conceptualizations of Space and Spatial Information

Werner Kuhn (ETH Zürich, CH)

How should one conceptualize space in order to effectively and efficiently solve a spatial problem? This question has driven my research on ontologies of geographic space for the last three decades. While I have primarily addressed it to improve the usability of Geographic Information Systems (GIS), my longer-term goal is to design spatial computations that are cognitively more adequate. My arguments will be limited to geographic space, which I define as space in time at the scales of human activities. Some common examples of problems in geographic space are those arising in wayfinding, assessing environmental impacts of a construction project, or understanding what drives climate change.

The preliminary answer to the stated open question that I will advocate and explore at this Dagstuhl Seminar is that three conceptualizations of geographic space are necessary and sufficient. In these three conceptualizations, one can view spatio-temporal phenomena as

- fields of attributes varying across space and time
- objects and object collections with identity
- events with fields and objects as participants.

None of these concepts are, of course, novel to modelers and practitioners. My point is that they are necessary and sufficient to deal with geographic phenomena, if suitably (re)defined:

- fields as continuous functions from space and time to attribute
- objects as individuals bounded in space (though not necessarily having an explicit boundary)

- events as temporally chunked instances of processes, bounded in time, and located by the fields and objects they change.

Each of these three conceptualization highlights a key aspect of spatial phenomena, namely

- the spatio-temporal variation of attributes (such as temperature, gravity, or population density);
- the properties, relations, and behavior of individuals (such as animals or vehicles) and groups or aggregates of these;
- the attribution of changes in fields and objects to events (such as the passing of a cold front, an earthquake, or a pandemic).

The three conceptualizations serve as conceptual lenses that one chooses and combines when solving spatial problems [1]. They precede and are independent of the data models used to represent them (such as the various vector and raster models used in GIS), in the sense that their choice does not imply or necessitate a particular data model, and that data in a data model (say, a raster data set) do not necessarily represent a particular conceptualization (such as a field view). While some data models apply more easily to some conceptualizations than others (and vice versa), it is essential for the usability and interoperability of systems to understand the difference between conceptualizing the world and representing the result in a convenient computational model.

In addition to the three conceptualizations of geographic space, I suggest that spatial information itself, in particular its quality, can also be viewed in only a few ways, currently in terms of the three notions of

- granularity or level of detail;
- accuracy with respect to a given or assumed reference;
- provenance from an agent that used some data and methods to produce it.

Non-geographic spaces, such as those of human organs, chemical molecules, or galaxies also appear to be best conceptualized through fields, objects, or events, and information about them is also usefully assessed in terms of granularity, accuracy, and provenance. Furthermore, spatializations of abstract data are beneficially designed around these concepts. Given their presumed general validity, I refer to the concepts of field, object, event, granularity, accuracy, and provenance as Core Concepts of Spatial Information.

To formalize the core concepts, I use parameterized type classes in Haskell. Such a functional approach, where spatial questions are functions returning spatial information, allows for an explicit account of how conceptualizations constrain representations, which in turn are data types instantiating the type classes. It also allows for combining multiple views of space, and translating between representations (though not between conceptualizations, which is by definition not achievable).

References

1 Allen, Christopher, Thomas Hervey, Sara Lafia, Daniel W. Phillips, Behzad Vahedi, and Werner Kuhn. 2016. Exploring the Notion of Spatial Lenses. In *International Conference on Geographic Information Science*, 259 – 74. Lecture Notes in Computer Science 9927, Springer-Verlag.

2 Kollasch, Fenja, and Werner Kuhn. 2017. The Concept of Location in Astronomic Spaces. In *Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory*, 67 – 71. Lecture Notes in Geoinformation and Cartography, Springer-Verlag.

3 Kuhn, Werner. 2007. An Image-Schematic Account of Spatial Categories. In *8th International Conference on Spatial Information Theory (COSIT)*, edited by Stephan Winter, Matt Duckham, Lars Kulik, and Ben Kuipers, 152 – 68. Lecture Notes in Computer Science 4736, Springer-Verlag.

4 Kuhn, Werner. 2012. Core Concepts of Spatial Information for Transdisciplinary Research. *International Journal of Geographical Information Science: IJGIS* 26 (12): 2267 – 76.

5 Lafia, Sara, Werner Kuhn, Kelly Caylor, and Libby Hemphill. 2021. Mapping Research Topics at Multiple Levels of Detail. *Patterns 2* (3): 100210.

3.10 Incorporating Spatial Cognition into an Embodied Cognitive Architecture

Pat Langley (Stanford University, US)

Main reference Pat Langley, Mike Barley, Ben Leon Meadows, Dongkyu Choi, Edward P. Katz: "Goals , Utilities , and Mental Simulation in Continuous Planning", in Proc. of the Fourth Annual Conference on Cognitive Systems, June 23-26, 2016, Evanston, IL, USA
URL http://www.cogsys.org/proceedings/2016/paper-2016-15.pdf

Research on cognitive architectures aims to develop unified theories of intelligent systems [6, 2]. This paradigm incorporates many ideas from cognitive psychology and focuses on agents that operate over time. However, an emphasis on generality has led the architecture community to avoid making commitments about the representation and processing of spatial content. The fields needs additional work that addresses this critical omission and, in this position piece, we report some initial progress in this direction.

In previous research, we have developed PUG, a cognitive architecture for embodied agents [3]. The framework incorporates ideas from classical architectures, such as distinguishing long-term from short-term memory and encoding their content as modular mental structures. However, it combines symbolic and numeric representations and processing to support a unified account of discrete planning and continuous control. Key theoretical commitments of the architecture include:

- Symbolic *concepts* are *grounded* in constituent physical objects and their associated numeric attributes.
- Symbolic concepts are *graded*, matching to different degrees based on constituent objects' attributes.
- Symbolic *skills* include equations for control attributes that are functions of *mismatch* to target concepts.
- *Execution* involves the iterative use of skills to compute control values *based on these mismatches.*
- To determine values for control attributes, execution takes the *vector sum* of results from active skills.
- *Processes* specify equations that predict changes to state attributes based on state and control attributes.
- *Mental simulation* uses skills and processes to predict trajectories in the agent's state space over time.
- *Motion planning* involves search through a space of skill sets to achieve the agent's desired beliefs.

- *Task planning* searches through a space of motion plan sequences to achieve the agent's desired beliefs.
- *Trajectories* produced by mental simulation are used to evaluate candidate motion plans and task plans.

PUG provides a high-level programming language we have used to create simple robot agents that operate in simulated two-dimensional environments. Demonstrations have included tasks in which robots must move to static objects and must pursue mobile ones, in both cases avoiding obstacles along their paths. We have also developed an extended framework, PUG/X, that integrates generation, execution, and monitoring of task and motion plans to let agents detect and respond to unexpected events [2].

The PUG architecture supports agents that operate in dynamic physical environments and it unifies symbolic approaches to problem solving with continuous approaches to control, but it does not yet incorporate any strong theoretical commitments about how to encode or reason over spatial content. In response, we have devised some promising extensions to the framework that address its representational and processing limitations. These include distinctive claims that:

- Agents encode spatial situations using *continuous numeric relations* to objects in the environment.
- Agents represent their spatial relations to these objects in terms of *egocentric polar coordinates*.
- A *place* P is a concept that specifies a set of visible objects, along with distances and angles to them.
- An instance I of place P includes the perceived distances and angles to P's associated objects.
- Inference estimates I's distance and angle to P from distances and angles to P's associated objects.
- The degree of I's match to place P is a function of its estimated distance and angle to P.
- Skill execution uses I's match to P as a guide for moving to a situation that matches P acceptably.
- A *map* is a collection of place concepts that include distances and angles to *other places*.
- Skill execution can use a place definition P to move from P to another place Q that it references.

These theoretical postulates differ markedly from those adopted by most research efforts on mobile robotics, which encode space as discretized, world-centric grids. However, they share elements with a few frameworks, such as those reported by [1, 7, 8]. We have implemented a number of these ideas and demonstrated them in a simulated robotic environment [5]. Future work will test the augmented architecture on agent localization, inter-place navigation, and other tasks that require spatial reasoning.

References

1 Kuipers, B. J., & Byun, Y.-T. (1991). A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems*, *8*, 47–63.

2 Langley, P. (2017). Progress and challenges in research on cognitive architectures. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (pp. 4870–4876). San Francisco: AAAI Press.

3 Langley, P., Barley, M., Meadows, B., Choi, D., & Katz, E. P. (2016). Goals, utilities, and mental simulation in continuous planning. *Proceedings of the Fourth Annual Conference on Cognitive Systems*. Evanston, IL.
4 Langley, P., Choi, D., Barley, M., Meadows, B., & Katz, E. P. (2017). Generating, executing, and monitoring plans with goal-based utilities in continuous domains. *Proceedings of the Fifth Annual Conference on Cognitive Systems*. Troy, NY.
5 Langley, P., & Katz, E. P. (2021). *A cognitive systems approach to continuous control*. Unpublished manuscript, Center for Design Research, Stanford University, Stanford, CA.
6 Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*, 141–160.
7 Yamauchi, B., & Langley, P. (1997). Place recognition in dynamic environments. *Journal of Robotic Systems*, *14*, 107–120.
8 Yeap, W. K. (2011). How Albot0 finds its way home: A novel approach to cognitive mapping using robots. *Topics in Cognitive Science*, *3*, 707–721.

3.11 On Maps and Vision-and-Language Navigation Tasks

Marcos Lopes (Universidade de São Paulo, BR)

In one of his seminal papers on Artificial Intelligence, John McCarthy compared humans and machines receiving and following instructions for problem solving. Machines normally take sequences of imperative sentences whereas humans are instructed "mainly in declarative sentences describing situations in which action is required" [3]. In other words, humans would know what to do directly by understanding the states of affairs. A few years later, A. Sloman would add yet another human strategy to problem solving: "analogical representations such as maps models", which could be employed as means for enhancing reasoning capabilities of artificial systems (or describing human intelligence) with non-linguistic representations [4]. These discussions from the beginnings of AI seem to fully apply to nowadays attempts of formalizing spatial relationships understanding by machines and humans. A comparison of cartographic representations in maps and route instructions that could be given both to humans and to robotic agents may shed a light on this matter.

How do maps and navigation instructions overlap and interrelate? One of the main uses for maps is to allow people to find paths between cities or points of interest outdoors or indoors. Nevertheless, the mere depiction of roads or tracks on the map does not automatically translate into instructions on how to go from one location to another. One still has to algorithmically convert route knowledge from the map into action steps for reaching the target. Easy as this may sound, maps do not stand as necessary nor sufficient conditions for navigation. On the other hand, the inverse frequently applies: new paths often engender new maps. Yet not only the information needed for navigation is different from pure cartographic information, it also relies on different cognitive capabilities.

Starting off by what maps and navigation instructions have in common, they both deal with ways of conceiving spatial relationships, and they both serve specific purposes. Purpose is explicit for navigation tasks – namely, to reach a certain target location. Regarding maps, purposes may vary but are always there: to provide information on political borders, or on topographic features etc. And, as [1] put it, purposes are often made explicit by the map title.

Generation and understanding of navigation instructions are part of a specific Artificial Intelligence multimodal task called Vision-and-Language Navigation (VLN). Despite of the fact that maps and VLN both deal with visual and natural language data, it is arguable that knowledge is represented in quite different ways on each of them. As diagrams, maps render all the relevant information about the depicted objects simultaneously available. It is important to mention that the merotopological relations on the map do not reflect the merotopology of the "real world" directly, but rather the merotopological *knowledge* about the world. Medieval *mapa mundi* will prove this point. Objects not exhibited in the map are either irrelevant for its purpose or unknown. In VLN, on the contrary, it is not essential to know all the necessary steps towards the target beforehand. Instead, new objects and moves are added at each step as navigation goes on. Any assertions about future states of affairs related to the path are at least partly uncertain (i.e., logically incomplete).

Navigation instructions can rely exclusively on topological directives such as "turn right", "go forth" and so on, whereas maps exhibit *places* instead of mere locations or spatial relationships. Places convey meanings [6] and can be often identified by Named Entities in maps or in speech [8].

In papers dealing with navigation instructions uttered by humans for humans, objects, buildings, and landscape sights are typically used as landmarks [5, 7, 2]. Addressees of the instructions have their attention called upon those landmarks for situating themselves before starting to walk or drive. Landmarks are systematically chosen employing familiarity criteria. They must be easy to recognize (matching the description) by addressees. Navigation instructions are thus grounded on familiar objects (indoors) or familiar landmarks (outdoors). They explicitly rely on common ground information.

Instruction utterances are ordered by two main factors: the order of presentation, which is roughly analogous to the move steps (what comes first on the path is uttered first), and the order of familiarity in which supposedly recognizable objects help to find out which way to go throughout potentially unfamiliar paths. It is also important to notice that path descriptions usually rely not on geometrical objects like lines and curves, but on body orientation and movement. This too marks a difference from maps, for which there are no such embodied nor dynamic features.

Finally, while maps are not capable of representing negation at all, propositions in navigation instructions are often canceled ("There is a door to your right – no, I mean, to your left"). In such cases, the two propositions combined are never understood as a persisting contradiction but instead the latter always cancels the former. As for maps, they have no means to express any negative propositions whatsoever. Whenever the knowledge depicted in the map changes for any reason (say, for instance, when national borders change), a new map must be drawn – epistemically speaking every new piece of knowledge should hence bring about a new map (or, as [1] call it, a new *layer* in an existing map).

Summing up, I would argue that differences and interrelations between maps and navigation instructions are capable of fully instantiating the three cognitive models for problem solving in general as put forth by McCarthy and Sloman (among others) early in the dawn of AI, namely: by explicitly following instructions; by descriptively representing knowledge, and by diagrammatically (i.e., analogically) apprehending relationships.

References

1 Casati, Roberto and Achille Varzi (2018). *Map semantics 20 years on: Basics and open problems.* Workshop. EHESS, Paris.

2 Linde, Charlotte and William Labov (1975). "Spatial networks as a site for the study of language and thought". In: *Language*, pp. 924 – 939.

3 McCarthy, John (1960). *Programs with common sense*. RLE and MIT computation center Cambridge, MA, USA.
4 Sloman, Aaron (1971). "Interactions between philosophy and artificial intelligence: The role of intuition and non-logical reasoning in intelligence". In: *Artificial Intelligence* 2.3-4, pp. 209 – 225.
5 Taylor, Holly A and Barbara Tversky (1992). "Descriptions and depictions of environments". In: *Memory &: Cognition*, 20.5, pp. 483 – 496.
6 Tenbrink, Thora (2011). "Reference frames of space and time in language". In: *Journal of Pragmatics*, 43.3, pp. 704 – 722.
7 Tversky, Barbara, Paul Lee, and Scott Mainwaring (1999). "Why do speakers mix perspectives?" In: *Spatial Cognition and Computation*, 1.4, pp. 399 – 412.
8 Yousaf, Madiha and Diedrich Wolter (2021). "A reasoning model for georeferencing named and unnamed spatial entities in natural language place descriptions". In: *Spatial Cognition & Computation*, pp. 1 – 39.

3.12 Customizable Qualitative Spatial Reasoning: Is Answer Set Programming the Answer?

Bertram Ludäscher (University of Illinois at Urbana-Champaign, US)

The composition table of region connection calculus can be easily and naturally specified via ASP (answer set programming) rules. For example, the partial knowledge represented by disjunctions in RCC-5 or RCC-8 [4] is naturally expressed by disjunctions in ASP rule heads or by choice rules. We have successfully employed different ASP encodings to RCC reasoning problems [7, 6, 5, 3, 1, 2], but encountered two main technical challenges. First, the natural ASP encoding via pairwise relations of regions does not allow to distinguish configurations that should be distinguishable. For example, if we know that three regions A, B, and C are pairwise overlapping, we do not know whether or not the intersection of all three regions is empty or not. A fine-grained case-analysis should enumerate the alternative solutions. To overcome this limitation, a simple approach is to extend the underlying vocabulary from n region names to 2^n "combined region" names, corresponding to the minterms in a Karnaugh-Veitch diagram. Unfortunately, the gain in expressive power is paid for with an exponentially larger search space, which seems to render the approach infeasible for all but the smallest problems. A second challenge is how to exploit the hierarchical structure of many reasoning problems, e.g., when aligning and merging two taxonomies using RCC relations [7] in order to obtain a more scalable approach. Based on our experience and initial experiments, we believe that answer set programming is an ideal paradigm to specify and implement user-customizable qualitative reasoning approaches, provided we can solve the two technical challenges just outlined. We invite the qualitative reasoning and ASP community to join forces on this endeavor.

References

1 Chen, M., Yu, S., Franz, N., Bowers, S., Ludäscher, B.: Euler/X: A Toolkit for Logic-Based Taxonomy Integration. In: Hanus, M., Rocha, R. (eds.) Joint Session: 22nd Intl. Workshop on Functional and (Constraint) Logic Programming (WFLP) and 27th Workshop on Logic Programming (WLP), at Kiel Declarative Programming Days. Institut für Informatik, Christian-Albrechts-Universität zu Kiel (Sep 2013)

2 Chen, M., Yu, S., Franz, N., Bowers, S., Ludäscher, B.: A Hybrid Diagnosis Approach Combining Black-Box and White-Box Reasoning. In: RuleML. Lecture Notes in Computer Science, vol. 8620, pp. 127–141. Springer (2014)
3 Cheng, Y.Y., Franz, N., Schneider, J., Yu, S., Rodenhausen, T., Ludäscher, B.: Agreeing to disagree: Reconciling conflicting taxonomic views using a logic-based approach. Proceedings of the Association for Information Science and Technology **54**(1), 46–56 (Jan 2017). https://doi.org/10.1002/pra2.2017.14505401006, `http://onlinelibrary.wiley.com/doi/10.1002/pra2.2017.14505401006/abstract`
4 Cohn, A.G., Renz, J.: Chapter 13 Qualitative Spatial Representation and Reasoning. In: van Harmelen, F., Lifschitz, V., Porter, B. (eds.) Foundations of Artificial Intelligence, Handbook of Knowledge Representation, vol. 3, pp. 551–596. Elsevier (Jan 2008). https://doi.org/10.1016/S1574-6526(07)03013-1, `https://www.sciencedirect.com/science/article/pii/S1574652607030131`
5 Franz, N.M., Chen, M., Kianmajd, P., Yu, S., Bowers, S., Weakley, A.S., Ludäscher, B.: Names are not good enough: Reasoning over taxonomic change in the Andropogon complex 1. Semantic Web **7**(6), 645–667 (Jan 2016). https://doi.org/10.3233/SW-160220, `https://content.iospress.com/articles/semantic-web/sw220`
6 Franz, N.M., Chen, M., Yu, S., Kianmajd, P., Bowers, S., Ludäscher, B.: Reasoning over Taxonomic Change: Exploring Alignments for the Perelleschus Use Case. PLOS ONE **10**(2), e0118247 (Feb 2015). https://doi.org/10.1371/journal.pone.0118247, `http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118247`
7 Thau, D., Bowers, S., Ludäscher, B.: Merging Taxonomies under RCC-5 Algebraic Articulations. Journal of Computing Science and Engineering **3**(2), 109–126 (2009)

3.13 Problem Representation and Analysis over Spatial Substrate

Chayanika Deka Nath (Tezpur University – Napaam, IN)

An explicit representation of visual input seems promising for efficient low-level perceptions during visual information processing. This abstract discusses an approach for mapping spatial objects and relations, exploring the power of diagrams directly influencing an information perception and visualization paradigm over the problem's physical substrate. Our objective is a comprehensive representation and reasoning of visual information combining Qualitative Spatial and Temporal Reasoning (QSTR) and diagram-based reasoning techniques within Diagrammatic Reasoning (DR). Qualitative Spatial and Temporal Reasoning (QSTR) [3] is a foremost mechanism with numerous formalisms to deal with spatial entities for space-time relational abstractions. However, existing spatial problem representation techniques are often inefficient to perceive exact spatial relations [6]. Under such scenarios, explicit representation through diagrams holds promise. Diagrammatic Reasoning (DR) allows reasoning over diagrams through a set of actions like diagram modification [12], manipulation [4], and re-organization [2]. In line with Freksa's idea [5, 4] on spatial problem solving directly over the spatial substrate as a computationally efficient technique within cognitive processing, we intend to address the above shortfall and strengthen existing formal representation and reasoning techniques through a QSTR-DR framework.

3.14 Position Paper: Represention and Reasoning with Spatial Knowledge in SUMO

Adam Pease (Articulate Software – San Jose, US)

Joint work of Adam Pease, Paulo Santos, Alexandre Rademaker
Main reference Adam Pease, Paulo Santos, Alexandre Rademaker: "A Corpus of Spatial Reasoning Problems", in Proc. of the 6th Conference on Artificial Intelligence and Theorem Proving, 2021.
URL http://www.adampease.org/Pease-AITP2021.pdf

Spatial reasoning is very common in common sense problems. Where could I have put my shoes? How can I get home from work and pick up milk on the way? Can this shelf hold my suitcase? We deal with problems like this throughout the day and solving them requires extensive knowledge about the world. We discussed in our sessions at Dagstuhl that much work to date has been towards solving isolated problems like route planning, or solving well structured mathematical problems [3], or representing general abstractions [13]. But comparatively little work has been towards capturing the wealth of often mundane concepts and facts that are needed to encode in a machine the broad knowledge that we use about the world in order to reason every day. Until now there has not been a collection of real-world reasoning problems available to test the sufficiency of spatial knowledge representation.

In our work we utilize the Suggested Upper Merged Ontology (SUMO)[7, 9][7], a comprehensive ontology of around 20,000 concepts and 80,000 hand-authored logical statements in a higher-order logic, that has an associated integrated development environment[12] including leading theorem provers such as Eprover [15] Vampire [6] and LEO-II [1], and manually-created links[8] to all 117,000 word senses in the WordNet lexico-semantic database[4]. SUMO has automatic translations [12] to the strictly first order language of TPTP [18], as well as the TF0 language (first order logic with typed arithmetic) [10] and THF[2]. The recent development of our TF0 translator means that we can reason with numerical measures in a truth-preserving representation [14]. Numerical reasoning is very common in practical spatial problems.

In order to demonstrate progress and a method for solving common-sense spatial problems, we developed a set of problems stated in English, formalized in SUMO, and for some of those problems, solved with the Vampire theorem prover. The current set of problems with solutions is available on GitHub[8] and described briefly in [11].

A challenge for representation is the large numbers of possible objects, processes and relationships that one may need in order to capture a set of non-trivial, real-world problems. In particular, one of our Dagstuhl sessions discussed the need for *affordances*[5] that can designate the tasks that some kinds of objects can participate in. SUMO is a library of just such a large set of formalized concepts. SUMO has 4801 subclasses of `Object` many of which are extensively formalized with rules that define their real world relationships and behavior. There are 1321 kinds of `Process` that govern what the use and purpose of many of those objects may be. Physical things exist in time, so there are 39 `TemporalRelation`s that include facilities for representing particular metric times, including days of the week and holidays. Central to this effort are the instances of `SpatialRelation`, of which there are 83.

In order to use these formalizations in theorem proving, it is essential that they be free of contradiction. However, in such a large system, it is not possible to ensure consistency by hand. We employ E and Vampire to test the consistency of SUMO. While issues can often

[7] `https://www.ontologyportal.org`
[8] `https://github.com/ontologyportal/sumo/blob/master/tests/SpatialQs.txt`

be found simply by asking the provers to prove "false", we can also employ an automated means of focusing on different parts of the ontology. The E prover has a mechanism [16] for generating thousands of test problems that focus on different relation symbols, which are then subjected to theorem proving. We also employ the StarExec [17] server cluster to make such large scale testing practical.

References

1. Christoph Benzmüller, Laurence Paulson, Frank Theiss, and A. Fietzke. (2008). LEO-II – A Cooperative Automatic Theorem Prover for Higher-Order Logic. In Proceedings of the Fourth International Joint Conference on Automated Reasoning (IJCAR'08), LNAI volume, 5195:162 – 170, 2008.
2. Christoph Benzmüller and Adam Pease. Progress in automating higher-order ontology reasoning. In Boris Konev, Renate Schmidt, and Stephan Schulz, editors, Workshop on Practical Aspects of Automated Reasoning (PAAR-2010). CEUR Workshop Proceedings, Edinburgh, UK, 2010.
3. E. W. Dijkstra. A note on two problems in connexion with graphs. Numer. Math., 1(1):269 – 271, dec 1959.
4. Christiane Fellbaum, editor. WordNet: An Electronic Lexical Database. MIT Press, 1998.
5. J. J. Gibson. The senses considered as perceptual systems. Houghton Mifflin, Boston, 1966.
6. Laura Kovács and Andrei Voronkov. First-order theorem proving and vampire. In Proceedings of the 25th International Conference on Computer Aided Verification, volume 8044 of CAV 2013, pages 1 – 35, New York, NY, USA, 2013. Springer-Verlag New York, Inc.
7. Ian Niles and Adam Pease. Toward a Standard Upper Ontology. In Chris Welty and Barry Smith, editors, Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), pages 2 – 9, 2001.
8. Ian Niles and Adam Pease. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In Proceedings of the IEEE International Conference on Information and Knowledge Engineering, pages 412 – 416, 2003.
9. Adam Pease. Ontology: A Practical Guide. Articulate Software Press, Angwin, CA, 2011.
10. Adam Pease. Arithmetic and inference in a large theory. In AI in Theorem Proving, 2019.
11. Adam Pease, Paulo Santos, and Alexandre Rademaker. A corpus of spatial reasoning problems (abstract) Proceedings of AI in Theorem Proving (AITP2021).. 2021.
12. Adam Pease and Stephan Schulz. Knowledge Engineering for Large Ontologies with Sigma KEE 3.0. In The International Joint Conference on Automated Reasoning, 2014.
13. David A Randell, Zhan Cui, and Anthony G Cohn. A spatial logic based on regions and connection. KR, 92:165 – 176, 1992.
14. Stephen Read. Logical consequence as truth-preservation. Logique and Analyse, 183(4):479 – 493, 2003.
15. S. Schulz. E. A Brainiac Theorem Prover. Journal of AI Commun., 15(2/3):111 – 126, 2002.
16. Stephan Schulz, Geoff Sutcliffe, Josef Urban, and Adam Pease. Detecting inconsistencies in large first-order knowledge bases. In Proceedings of CADE 26, pages 310 – 325. Springer, 2017.
17. Aaron Stump, Geoff Sutcliffe, and Cesare Tinelli. Starexec: A cross-community infrastructure for logic solving. In Stéphane Demri, Deepak Kapur, and Christoph Weidenbach, editors, Automated Reasoning, pages 367 – 373, Cham, 2014. Springer International Publishing.
18. Steven Trac, Geoff Sutcliffe, and Adam Pease. Integration of the TPTPWorld into SigmaKEE. In Proceedings of IJCAR '08 Workshop on Practical Aspects of Automated Reasoning (PAAR-2008). CEUR Workshop Proceedings, 2008.

3.15 Collaborative Mission Planning for Autonomous Maritime Vehicles

Paulo E. Santos (Flinders University – Adelaide, AU)

Joint work of Paulo E. Santos, Karl Sammut
Main reference Thiago Pedro Donadon Homem, Paulo Eduardo Santos, Anna Helena Reali Costa, Reinaldo Augusto da Costa Bianchi, Ramón López de Mántaras: "Qualitative case-based reasoning and learning", Artif. Intell., Vol. 283, p. 103258, 2020.
URL http://dx.doi.org/10.1016/j.artint.2020.103258

The exploration of unknown environments, often presents unforeseen challenges and inherent risks due to the uncertainties involved. While single autonomous systems are capable of completing complex missions, the introduction of multi-robotic teams can permit increased level of efficiency to a given task, especially when these tasks cover large areas and mission completion time is a critical constraint. The use of heterogeneous robotic teams of autonomous underwater vehicles (AUVs), autonomous surface vehicles (ASVs) and seabed crawler vehicles can further facilitate a higher degree of flexibility and redundancy when analysing complex environments under diverse environmental conditions. The challenge of traditional autonomous-vehicle team-based localisation and control techniques are, however, considerably magnified in the underwater domain by the lower reliability and potential asynchronicity of underwater acoustic communications as compared to RF based communication in the above water domain. This complicates possible mission tasking approaches for hybrid teams of autonomous marine vehicles in terms of obtaining a shared understanding of the environment and the team status, thus requiring new solutions for control, coordination, collaboration and communication to overcome these complications.

This brings the need to investigate efficient reasoning processes, communication strategies and underlying low-level control mechanisms necessary to coordinate heterogeneous teams of autonomous marine vehicles, in dynamic and uncertain environments. An underlying assumption of this work is that the autonomous agents have to achieve a common agreement, via a negotiation procedure, in order to solve complex problems collaboratively. Our aim is the development of a mixed-initiative system [1], where the interaction and negotiation between the agents will maximise their resources in order to optimise the successful execution of a common task. The negotiation procedure between vehicles will be conducted in game-theoretic terms [4], where the agent interaction is modelled as a cooperative game and the Nash equilibrium (representing the agents' agreement) will be obtained by online distributed algorithms [3]. This provides efficient task allocation solutions that can be easily extended to consider outside threats along with team collaboration, where the interactions with additional agents are modelled as a non-cooperative game [5]. Negotiation, however, occurs only when there is some level of conflict in the perceived states, assigned actions across the team and the availability of resources within the heterogeneous team. To obtain an efficient problem solving policy for any given problem we propose to use a novel algorithm, Qualitative Case-Based Reasoning and Learning (QCBRL) [2]. In QCBRL, cases are predetermined solutions for groups of autonomous agents (represented as QSR formulae) that could be adapted to similar situations. A reinforcement learning (RL) module enables the team of agents to learn new solutions to unforeseen situations at runtime, without assuming a pre-processing step. Extending QCBRL with the game-theoretic delegation model to a team of underwater vehicles is one of the main tasks to be executed. The project should bridge the gap between current QSR laboratory experiments to large-scale robotics application.

References

1 Jean Bouchard, Jonathan Gaudreault, Claude-Guy Quimper, Philippe Marier, Edith Brotherton, and Nathaniel Simard. Mixed-initiative system for tactical planning allowing real-time constraint insertions. *IFAC-PapersOnLine*, 50(1):15233 – 15240, 2017. 20th IFAC World Congress.

2 Thiago Homem, Paulo E. Santos, Anna H. Reali Costa, Reinaldo A. C. Bianchi, and Ramon Lopez de Mantaras. Qualitative case-based reasoning and learning. *Artificial Intelligence*, 283:103258, 2020.

3 K. Lu, G. Li, and L. Wang. Online distributed algorithms for seeking generalized nash equilibria in dynamic environments. *IEEE Transactions on Automatic Control*, pages 1–1, 2020.

4 Simon Parsons and Michael Wooldridge. Game theory and decision theory in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 5:243–254, 09 2002.

5 Cheng-Kuang Wu and Xingwei Hu. A game theory approach for multi-agent system resources allocation against outside threats. *Journal of Risk Analysis and Crisis Response*, 9, 11 2019.

3.16 Conceptual Transformations of Spatial Information

Simon Scheider (Utrecht University, NL)

Main reference Simon Scheider, Jürgen Hahn, Paul Weiser, Werner Kuhn: "Computing with cognitive spatial frames of reference in GIS", Trans. GIS, Vol. 22(5), pp. 1083–1104, 2018.
URL http://dx.doi.org/10.1111/tgis.12318

Conceptual models of space are required to model spatial reference in natural language processing, orientation in spatial cognition, as well as to make effective use of spatial information for answering questions. Yet, such models often seem limited in accounting for and distinguishing the different ways how space can be conceptualized. One reason seems to lie in our inability to assess the possible transformations of the concepts underlying spatial information. Therefore, future research should study *transformation models* for those concepts that are constitutive of spatial information, including the core concepts of location, field, object, event and network.

Transformations lie at the core of spatial reference systems. For example, to understand a location of a particular *coordinate reference system*, we can express it as a series of transformations starting from a common frame, e.g. from a geocentric Cartesian frame using ellipsoidal angles and projection equations. This fundamental observation applies equally to *cognitive reference frames*. Frames enable orientations in the wild, as well as interpreting spatial references in natural language texts. For example, to move towards the specified location, an intrinsic spatial reference such as "in front of the church" may need to be transformed into an allocentric overview map, and finally into an egocentric frame. Likewise, "mapping" spatial references in texts onto a geographic map requires suitable cognitive transformations into a coordinate reference system. To this end, we need to account for the *diversity of linguistic frames*. Frames frequently go beyond the classical Euclidean strategies (as used by Levinson), including zonal strategies (neighbourhoods around and within objects), topological strategies (inside, outside, etc.), as well as linear referencing strategies (distance along a path) (see Table 1). Transformation models could therefore help us uncover possible frame variants in texts, and at the same time give us a way to transform locative expressions into a map.

Table 1 Transformation types for core concepts.

transformation type		core concept (from)	parameters	core concept (to)
coordinate reference transformations	projection ellipsoidal angle	location (point)	ellipsoid datum	location (point)
cognitive reference transformations	euclidean	location (fuzzy)	object direction distance	location (fuzzy)
	zonal	location (fuzzy)	object distance	location (fuzzy)
	topological	location (fuzzy)	object topological relation	location (fuzzy)
	linear	location (fuzzy)	object path	location (fuzzy)
core concept transformations	closest object distance	objects	distance	field
	field capacity	field	location	amount
	field coverage	field	amount	location
	object capacity	objects	location	amount
	object coverage	objects	amount	location

Transformations also provide a key towards handling the variability of conceptualizations of spatial information. Location (as defined by spatial reference systems) is only one out of a range of *core concepts* needed to deal with the content of spatial information (i.e., what geo-data represents). The latter includes also objects, fields, events and networks. Similar to spatial reference systems, core concepts often remain implicit in the geodata representing them. Still, conceptualizations vary with the way the underlying geodata was generated. That is, geodata transformations reflect also a *transformation on the conceptual level.* For example, a choropleth map of administrative units, produced by averaging a spatial field generated by measuring the Euclidean distance to the closest park, corresponds to a conceptual transformation from objects (parks) to some (distance) field, and from a field to amounts (average distance) defined as object qualities (unit statistics). In Table 1, the former transformation is called *closest object distance* transformation, and the latter *field-capacity* transformation. In field-capacity transformations, *amounts* play a central role in quantifying other core concepts. Amounts can be used to measure the *coverage of fields or objects* in terms of the locations they occupy ("the area further than one kilometre from a park"). It turns out that such transformations are at the core of Geographic Information Systems (GIS), and not only constrain the meaning of the resulting map ("average distance to the closest park") but also its purpose (e.g., for accessibility assessments).

In short, an important future research task lies in studying conceptual transformation models which account for the transformation possibilities of spatial information. This would not only enable us to deal with the inherent variability of concepts of space, it would also give us a way to model the ways spatial information can be transformed according to some purpose.

3.17 Hybrid AI Systems Grounded on Qualitative Spatio-Temporal Reasoning

Michael Sioutis (Universität Bamberg, DE)

We argue for the importance of exploiting the area of Qualitative Spatio-Temporal Reasoning (QSTR), as a means to develop hybrid AI systems involving computations with spatio-temporal information. In short, QSTR is a Symbolic AI framework for representing and reasoning about spatial and temporal information via the use of disjunctive natural relations, e.g., "Task **A** is scheduled *after* or *during* Task **C**". We give some examples of how QSTR can be the backbone of hybrid AI systems, emphasizing on abductive learning and AI planning in particular.

One of the main challenges in Artificial Intelligence (AI) today, and where the most progress towards the AI dream is expected to be seen over the next decade, is the seamless integration of statistical learning and symbolic reasoning [9].[9] Such an integration is the area of study of Neuro-Symbolic AI [2].[10] Specifically, over the past years, statistical learning and symbolic reasoning have been developed mostly separately in AI – with few exceptions [2]. What is more, the representations (models) that are learned in statistical learning are generally low-level (sub-symbolic), whereas the representations used in symbolic reasoning are high-level (symbolic); thus, integrating the former with the latter representations is challenging, and it is exactly this task that Neuro-Symbolic AI addresses. In sum, Neuro-Symbolic AI seeks to combine principles from neural-networks learning and logical reasoning, by leveraging the strengths of both worlds to the extent possible. In a sense, this framework closely resembles how humans perform problem-solving, as, from a psychological viewpoint, the perception, which is a data-driven process, and the logical reasoning, which is a knowledge-driven process, are entangled rather than separated in humans.

Drawing inspiration from the work in [1], where a qualitative spatial reasoner is used to act as a referee upon the output of a classifier, we argue for working towards a *generic* neuro-symbolic framework that will integrate qualitative spatio-temporal reasoning (QSTR) [4] and neural methods from a probabilistic perspective. Such an integration is currently identified as an open challenge in the AI community [5, Section 9]. In our context, qualitative spatial or temporal variables can be annotated with the probability-infused output of a neural network, and spatial or temporal relations themselves may carry a probability too. A simplified example of such a neuro-symbolic formula would look as follows (%s denote confidence):

$$X^{(95\%\text{ yolk})} \text{ is } \textit{contained in}^{(45\%\text{ true})} \text{ or } \textit{overlaps}^{(55\%\text{ true})}\ Y^{(90\%\text{ egg})}.$$

In that sense, QSTR becomes neurally-enhanced, and probabilities are used to encode a *bidirectional feedback loop* between the symbolic framework and the Machine Learning (ML) model, much like as in *Abductive Learning* [10, Figure 1]. We propose to use logic to compose concepts learned by ML methods, and also allow learned concepts by ML methods to influence that logical composition. We argue that a (symbolic) spatio-temporal knowledge base, naturally grounded on physics and human cognition, could provide a dependable causal seed upon which machine learning models could generalize [6]. Moreover, in relation to

[9] http://ai100.stanford.edu/2021-report
[10] http://www.neural-symbolic.org/

obtaining the most likely interpretation of an uncertain environment through neuro-symbolic reasoning, we argue for using this framework in order to integrate the two largely separate viewpoints of environment mapping and planning, and learn qualitative spatio-temporal representations of the environment to facilitate motion planning and control under complex spatio-temporal tasks. As an example, a service robot moves through a swarm of humans and other obstacles in a café to deliver a coffee; the uncertainty here pertains to noisy sensor readings and unpredictable moving objects. The qualitative approach of QSTR could discretize numerical data online and form the most likely qualitative spatio-temporal representation of the environment at each point of time, encoding the qualitative relations among the various actors (e.g., in terms of orientation, topology, or anything else needed). Such a qualitative approach can indeed drastically reduce the huge search space that is typical when constructing a plan in dynamic contexts of heterogeneous knowledge; this has been shown to some extent in [3], where a qualitative orientation calculus is used.
Excerpts of the above text have appeared in [8] and [7].

References

1. Marjan Alirezaie, Martin Längkvist, Michael Sioutis, and Amy Loutfi. Semantic referee: A neural-symbolic framework for enhancing geospatial semantic segmentation. *Semantic Web*, 10:863–880, 2019.
2. Artur S. d'Avila Garcez, Tarek R. Besold, Luc De Raedt, Peter Földiak, Pascal Hitzler, Thomas Icard, Kai-Uwe Kühnberger, Luís C. Lamb, Riisto Miikkulainen, and Daniel L. Silver. Neural-Symbolic Learning and Reasoning: Contributions and Challenges. In *AAAI Spring Symposium on Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches.*, 2015.
3. Thiago Pedro Donadon Homem, Paulo Eduardo Santos, Anna Helena Reali Costa, Reinaldo Augusto da Costa Bianchi, and Ramón López de Mántaras. Qualitative case-based reasoning and learning. *Artif. Intell.*, 283:103258, 2020.
4. Gérard Ligozat. *Qualitative Spatial and Temporal Reasoning.* ISTE Series. Wiley, 2011.
5. Luc De Raedt, Sebastijan Dumancic, Robin Manhaeve, and Giuseppe Marra. From Statistical Relational to Neuro-Symbolic Artificial Intelligence. In *IJCAI*, 2020.
6. Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward Causal Representation Learning. *Proc. IEEE*, 109:612–634, 2021.
7. Michael Sioutis and Hua Meng. Towards Robust Qualitative Spatio-Temporal Reasoning for Hybrid AI Systems. In *ISKE*, 2021. To appear.
8. Michael Sioutis and Diedrich Wolter. Qualitative Spatial and Temporal Reasoning: Current Status and Future Challenges. In *IJCAI*, 2021.
9. Leslie G. Valiant. Knowledge Infusion: In Pursuit of Robustness in Artificial Intelligence. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, 2008.
10. Zhi-Hua Zhou. Abductive learning: towards bridging machine learning and logical reasoning. *Sci. China Inf. Sci.*, 62:76101:1–76101:3, 2019.

3.18 Spatio-Temporal Granularity

John Stell (University of Leeds, UK)

Joint work of John Stell, Giulia Sindoni, Katsuhiko Sano
Main reference Giulia Sindoni, Katsuhiko Sano, John G. Stell: "Expressing discrete spatial relations under granularity", J. Log. Algebraic Methods Program., Vol. 122, p. 100682, 2021.
URL http://dx.doi.org/10.1016/j.jlamp.2021.100682

The topic of *Formalism* at this seminar asks what would be "suitable formalism for commonsense problem solving that allows an accurate, flexible, and readable knowledge representation for spatio-temporal effects of actions performed by an intelligent agent". Such a means of knowledge representation needs to have a rigorous semantics if the requirement to be accurate is to be justified. The need for semantics is also essential if the formalism is to support reasoning. The activity of reasoning is itself important if intelligent agents are to be capable of explaining their actions, and arguing for one course of action rather than another.

The need for commonsense problem solving, and the desire for a representation that is readable by humans, indicates a role for qualitative approaches to spatio-temporal descriptions underlying any formalism. Qualitative representations are well established [1], but the interaction with hierarchical models of discrete space and time representing changes at different levels of detail should be important here. In [3] a foundation for a theory of hierarchical change was set out, but this is restricted by the qualitative descriptions used. This theory incorporates both change over time and over level of detail and should be capable of integration with more recent work [2] which models qualitative variation across two levels of detail, but in a static setting.

In [3] change of level of detail is modelled in a discrete space through the operations of opening and closing as used in the image-processing techniques of mathematical morphology. This allows the modelling the way level of detail changes by capturing readily understandable phenomena. Examples include that way distinct regions separated by a narrow channel appear to fuse into a single region from a more "zoomed out" perspective. This need not be directly visual, but could describe the way qualitative closeness of objects can be affected by distance from an agent. The qualitative changes in [3] can be seen as changes to a single spatial entity, made up multiple parts that can split and merge both over time and over granularity.

Spatial relationships, such as two distinct entities overlapping, can also change in time and according to granularity, but are more complex than the existing model. Qualitative relationships changing over time at a single level of detail is well-known through ideas such as conceptual neighbourhoods. Granular change, as in [2], needs to be combined with temporal change to produce a a richer description that can support agents reasoning about effects of actions at different levels of detail.

References

1 A. G. Cohn and J. Renz. Qualitative spatial representation and reasoning. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of knowledge representation*, pages 551–596. Elsevier, 2008.
2 G. Sindoni, K. Sano, and J. G. Stell. Expressing discrete spatial relations under granularity. *Journal of Logical and Algebraic Methods in Programming*, 122, 2021, doi.org/10.1016/j.jlamp.2021.100682.
3 J. G. Stell. Granular description of qualitative change. In *23rd International Joint Conference on Artificial Intelligence, IJCAI*, pages 1111–1117, 2013.

3.19 Where Are You Really? Creating a More Complete Automated Model of the Factors that Influence the Interpretation of Spatial Relation Terms

Kristin Stock (Massey University – Auckland, NZ)

Main reference Kristin Stock, Christopher B. Jones, Shaun Russell, Mansi Radke, Prarthana Das, Niloofar Aflaki: "Detecting geospatial location descriptions in natural language text", International Journal of Geographical Information Science, Vol. 36(3), pp. 547–584, Taylor & Francis, 2022.
URL http://dx.doi.org/10.1080/13658816.2021.1987441

Current methods for computational modelling of the locations referred to in natural language relative location descriptions mostly rely on either geometric or linguistic information. Geometric models assume that specific relation terms reflect specific geometric configurations in space. For example, near (e.g. the church near the Aotea Centre) implies a specific distance range relative to a reference object; west of implies a particular azimuth (angle relative to north); and left (I live on the left side of Hillcrest Road) implies a region relative to some axis of a reference system which may depend on the position of the observer and/or other objects in the environment. It has long been recognised that these geometric models provide only part of the picture, and that humans take into account a multitude of contextual factors when interpretating spatial relation terms, including the physical environment and objects in it; the observer's goal and expectations of the audience knowledge; the audience location and knowledge and the characteristics of the object.

Some research has acknowledged the tendency for spatial relation terms to have degrees of "goodness of fit" in particular locations – areas where the relation definitely applies (e.g. 270 degree azimuth for west of), and areas where they may still fit, but less well, and probably depending on the context (e.g. 250 degree azimuth for west of) and developed probabilistic of density field based models to accommodate this. In addition, machine learning models have attempted to incorporate context to address the problem, by learning the interpretation of a spatial relation term within an expression from similar expressions, using word embeddings and other new technologies (e.g neural networks).

However, the list of contextual factors that have been incorporated in automated spatial relation term models (either explicitly or implicitly through language modelling) is still very small, mostly incorporating characteristics of the located and reference objects (e.g. type, size), and there is much potential for research advances in this field, in three directions:

1. Incorporating factors that have been shown to be important in linguistic and cognitive science research into computational models, which involves determining how to:
   a. Automatically extract data from natural language text or related geographic data sets that would be informative in regard to factors identified in linguistic and cognitive science research. For example, the importance of reference frame has been recognised, but this is not easy to automatically extract either from language (to encode expressions in machine learning models), or from the geographic environment (e.g. to decide in a given scene which objects count as being on the left side of Hillcrest Road).
   b. Encode extracted data in a form that would suit state of the art machine learning models (e.g. transfer learning models that work with large text corpora). This encoding must ensure that machine learning models can recognise the similarity in a particular aspect for pairs of expressions, and the data must be encoded in a way that highlights this. This may be easy for numerical factors (e.g. area), but much more difficult for

qualitative factors. Another example is "knowledge of the world" that is used to aid in interpretation – human understanding of how possible or likely some particular configuration might be (e.g. if we refer to the café beside the Thames, we assume that this is most likely on land, but not necessarily, as in the case of a café barge moored in the river).

2. Directing linguistic and cognitive science research towards factors that may be important, but have not yet been explored in the research to sufficient degree, or in a form to facilitate their incorporation in computational models. For example, user goals may influence the use of spatial relation terms, and to incorporate this in computational models, specific and concrete models of the influence of such factors on spatial relation terms (e.g. that can be encoded as numerical features) are useful. There are examples of interdisciplinary research that has used linguistic and cognitive methods to develop models for computational purposes, and more of this would benefit this research area.

3.20 Challenges for the implementation of spatiotemporal concepts

Thora Tenbrink (Bangor University, UK)

The formalisation and implementation of spatial/spatiotemporal concepts posits a range of hard challenges that are not likely to be overcome by standard automated machine-learning approaches. Human-readable description of complex spatiotemporal problems and relationships requires a clear understanding of human thoughts and thought processes, to an extent sufficiently detailed and fine-grained to allow for modelling, implementation and eventually automatic generation in language that humans will find easy to follow, even if it exhibits systematic differences to how humans would generate such a description themselves. For example, although automatically generated route descriptions have massively improved over the past decades, they still don't follow some of the most basic conceptual principles of wayfinding reflected in human route descriptions, such as references to landmarks and qualitative rather than metric information. Nevertheless, they are sufficiently aligned with human needs to be very useful.

Wayfinding and navigation have been extensively investigated over the past decades, but other areas of spatiotemporal problem-solving pose different challenges that haven't received quite as much attention. Here are some examples:

- Combination of visual and verbal information: Even considering the high quality of standard automatically generated route descriptions, it is notable that the visual information is detached from the verbal. In human communication, in contrast, all types of information are simultaneously integrated and frequently cross-referred, for instance by deictic pointers. This allows for systematic use of overspecification (providing more information than strictly needed, particularly for complex challenges), underspecification (providing only selected necessary information, allowing humans to make straightforward inferences as appropriate and adopting heuristics such as moving in a vague, visually approximated direction), and complementary presentation (where each bit of information is presented in the visual or verbal mode that is most suitable for it, e.g., shapes and angles are clearer in visual form, whereas temporal information is best conveyed verbally).

- Changes over time: Wayfinding takes time as a process, but as it happens within a short time frame it can be assumed that the world itself doesn't change. However, the world does change continuously, and humans adapt flexibly to such changes as a normal part of everyday life, in ways that are not well understood from a computational perspective. Some changes are simply nuisances (such as road blocks on the way to work – a minor problem that even most automatic navigation systems can address by calculating detours), whereas other changes are more significant, such as climate related changes (floods, sinkholes, extended fires) that are threatening to livelihoods, affect personal relations to places, and change habits more permanently. Crucially, humans have the extraordinary ability (not available to automatic systems) to use whatever method available at any given time and place to work around a challenge – especially if motivated, for instance by place attachment. To support this, automatic systems would need access to more flexible reasoning abilities that allow for re-purposing available resources in the face of local changes.
- Flexibility of reference systems: Within spatial cognition and computation research, the research focus on standard activities such as navigation and object localisation seems to have distracted from other actions in which humans engage around the world, and which require different (and often more flexible) ways of thinking about space or spatiotemporal relationships. The use of spatial reference systems not only differs across cultures (as shown for Levinson's intrinsic, relative and absolute reference systems) but also across activities in the everyday life of a single culture – such as using wind-based concepts for establishing directionality during sailing.

With ideas about "Artificial Intelligence" (though often represented by machine-learning based "intelligence") increasingly dominating agendas and debates both within and outside academia, it is time to address such fundamentally human ways of thinking, and consider how computational methods can be adapted to support them more flexibly and efficiently.

3.21 Spatial Problem Solving in Physical Manipulation Games

Diedrich Wolter (Universität Bamberg, DE)

Main reference Diedrich Wolter, Alexandra Kirsch: "Leveraging Qualitative Reasoning to Learning Manipulation Tasks", Robotics, Vol. 4(3), pp. 253–283, 2015.
URL http://dx.doi.org/10.3390/robotics4030253

One aim of AI research is to develop systems capable of solving real-world problems. To this end, processing of spatial, temporal and physical knowledge is key, in particular when envisaging future service robots that assist humans in manifold ways. A challenge faced in everyday situations human excel at is to come up with effective strategies in the light of uncertain knowledge about the environment. Humans may be poor at, or even unable to estimate physical parameters like friction, density, geometry, etc. but they can reliably stack dishes in the kitchen sink, arrange groceries in the shopping bag, or use tools in new but effective ways. By contrast, physical manipulation skills in robotics and AI largely hinge on constructing (respectively learning) precise physical models that are employed in forward models for planning. This limits applicability to problem instances sufficiently similar to what the agent has experienced so far or has been explicitly prepared for.

Unfortunately, sensor and motor capabilities available to robotic systems are both expensive, economically as well as with respect to the time required for maintenance and preparation. Above all, they are still no match to human sensorimotor skills. As a result, we may not see service robots in our homes capable of human-like object manipulation anytime soon. An economic way to make progress towards such skills can be seen in the realm of physical manipulation games. In these games, humans (or computers) are confronted with physical puzzles that need to be solved in a controlled yet not fully observable environment. Similar to how games have been regarded as a microcosm for conducting AI research in classic fields such as search [1], physical manipulation games can present a suitable basis for conducting research in spatial problem solving and conveniently evaluating the progress made towards human capabilities. One of such games is the game of Angry Birds which is used as basis for the AI competition "AI Birds" [2]. The objective of this game is to catapult objects at structures in a 2D world to destroy targets by physical impact. Although the task appears very simple, instances can be crafted that require a great variety of physical knowledge in order to be solved. Put differently, difficulty can easily be scaled. In particular, since the game is interfaced only via the graphical representation of a scene, physical parameters remain hidden like in real-world tasks.

By putting a limit on the time an agent is allowed to try solving a previously unseen level, abstract physical problem solving skills become imperative. A unique and important feature of games and real-world tasks is that evaluation based on performance punishes over-simplified approaches that make assumptions that cannot be met by technical systems. Our work on creating an agent for AI Birds has forced us to stress the interaction between abstract problem solving, perception, and task execution. We must not consider wrong estimates about the environment to be an exception (e.g., the shape of objects and whether they would fit into a container or not), but consider it to be too typical to ignore. We need ways that allow the agent to reflect its decision in the light of the effects caused, and to adapt quickly by drawing the right conclusions. In essence, physical manipulation games force us to take a sufficiently broad perspective on spatial problem solving that fosters development of techniques that will be useful components in an overall agent design.

References

1 Jonathan Schaeffer (2000). The Games Computers (and People) Play, In Proceedings of AAAI-00

2 Jochen Renz, XiaoYu Ge, Matthew Stephenson & Peng Zhang (2019). AI meets Angry Birds, Nature Machine Intelligence volume 1

3.22 On Concepts

Florentin Wörgötter (Universität Göttingen, DE)

Joint work of Florentin Wörgötter, Tomas Kulvicius, Minija Tamosiunaite
Main reference Minija Tamosiunaite, Tomas Kulvicius, Florentin Wörgötter: "Bootstrapping Concept Formation in Small Neural Networks", CoRR, Vol. abs/2110.13665, 2021.
URL https://arxiv.org/abs/2110.13665

The central problem addressed here is how to represent symbolic entities in the brain that can be used (also) for spatio-temporal reasoning. We can operate with symbols (language), but how can we get from stimulus driven experience formation to the formation of abstract thought and then back to motor-signals to execute and action. In the core of this is the question about "Concepts". The substrate "spatio-temporal" would here offer a nice testbed for asking if any representation "works or not".

In this short paper I would like to first provide some definitions and then only discuss some neuronal aspects. Maybe this Dagstuhl seminar wants to take this up and extend these thoughts into a more technical, computer science direction. I wrote this quickly and this is certainly not yet up to high scientific standard. It's just a collection of (wild) thoughts.

On concepts (some definitions):

1. Concepts are different from Categories because Categories are "un-reflected clusterings" within a feature space, whereas Concepts are the narrative, the rule, the essence of the thereby captured entities. This is related to the Piagetian "Schemas", but goes – in our view – beyond mere "feature constellations", which you could obtain by clustering, too. For example, Alpha-Go can cluster feature constellations on the Go-board into groups that require the same move, but Alpha-Go cannot speak out about the "rules of Go". How to formalize what we mean when saying the "a concept is given by a narrative" will be specified below.
2. Concepts are (at any moment in time!) "momentarily closed" with a relatively hard boundary to separate them from each other. One step from moving upwards from a Category to the forming of a concept is to perform this closure. Concepts are, thus, fundamentally discrete. There is a "cup" and no "partial-" or "almost-cup". And "cup" is also distinct from "no-cup".
3. The boundary of a Concept can (at the next moment) be modified when the agent's experience of the world changes. Clearly, some Concepts will, this way, have more volatile boundaries than others.
4. Concepts are compositional. They are (usually) composed from other (narrower) concepts.
5. They form dividable sets: "Apple" divides into "red, green, yellow Apple", etc. This way also hierarchies of Concepts can be formed. All this is very well known from ontologies in AI.
6. From the above we would argue that Concepts come into being by relational operations. Here the most fundamental one is the enclosure versus exclosure operator, but also operators like "part-of", "combined-from", etc. This allows us to arrive at a formal definition of how to form a Concept.
7. Naturally, for us Concepts can be (and are usually) expressed by language ("narrative" of an entity) and/or captured in AI using ontologies. Concepts can be expressed by single words, too. "Cup" is a concept. But this comes then with an explanatory narrative (which you can express in case you want to explain "your concept" – of a "Cup" – to someone else).
8. Thus, the act of expressing a Concept is fundamentally linked to symbolic representations (language, at least in humans). Semiotic processes need here to take place and the closure and separation (the making-discrete) of a Concept relative to other Concepts happens by this.
9. This makes it difficult to verify the existence of Concepts in non-linguistic agents (animals). Possible the only other way to verify this is to find planning-capabilities in another agent assuming that planning sequences and planning operators are directly related to (are the same as?) Concepts (for this see further below).

How many concepts do we have?

This question could be linked to language (again). How many words does YOUR vocabulary have? While there are >170,000 words existing in active English language (https://englishlive.ef.com/blog/language-lab/many-words-english-language/), native speakers use only about up to 35,000 of them. If we assume that every word is a concept then this would be the same number. Now some concepts may not be "word-expressible", but might require "a phrase". While this could go up to larger numbers this way, we might benefit from the compositionality of concepts. If you have phrased a concept, you might as well "name" it. Phrasing would correspond to the pulling out (from memory) one-by-one all those sub-concepts that form the new one. Hence, the new concept would at first consist of a sequential activation of those different sub-concepts.

The German language is here expressive. An alternative word for "Konzept" is "Begriff"; this refers to the word used to describe an entity/concept. This means literally the being able to "touch and feel" an entity as "greifen" means "to grasp", which we also use in English saying that "we grasp a concept". The "Begriff" for that furry animal that says "mio" and. . . etc., etc., is "cat". The narrative – the rule-set – for cat is then replaced by the Begriff "cat". Hence, this points to the capturing of a concept in quite a literal way and the forming of any novel "Begriff" is the conCATenation of a complex narrative into a single word. Certainly, in the Middle Ages no one had a concept of automobile. However, after its invention, this complex machine that looks like a horse-cart but drives on its own, needed a word that stands for the concept. Accordingly, the word "automobile" had been coined. Inventing a new word happens all the time in language, or loosing words for concepts no longer in use. For the Germans: do you know what a "Kumpf" is?? (I need an English example, here of an obsolete concept and its belonging word)

Thus, here the question arises if one needs now a process that makes this time-serial sequential narrative-representation into a momentarily-arising (simultaneously-arising) one. Can Concepts be spread out over narration-time or is this then "something different" and not yet a Concept? Possibly, we have "pop-up" Concepts, which are very strongly consolidated and reside as "one entity" in memory, while other more complex concepts are rather more sub-sequential. I would argue that only pop-up Concepts are true ones. All others are "in the making".

Neuronal Aspects:

The question of neuronal representations of concepts remains difficult. If we consider that we do not have more than 100,000 pop-up concepts (this number seems too high anyhow) then you could easily represent those by individual static (overlapping) cell assemblies. If we allocate 10,000 neurons per concept, we would need 109 neurons in total for this. Given that many concepts might have a very high degree of overlap, it could well be that this number reduces by- say – a factor of 100 to get us to 107 neurons needed for this. That is a very small number given that there are 100,000 neurons in 1 mm3 of cortical tissue. We need 102 mm3, with a thickness of 1mm (really its rather 1.5mm) this comes down to about 10x10 millimeter of cortical surface area to cover for this. If we go up one order of magnitude to 108 needed neurons (less assembly overlap) then this makes 33x33 millimeters. It would be interesting to measure the degree of psychological (cognitive) concept-overlap. For example asking people to explain "cup" and (after some distractors) "glass" and see how much semantic (linguistic) overlap exists in these explanations. Or some experiment of this kind.

On compositionality

If a Concept has not yet reached the stage of "pop up", then one could consider it as the spatio-temporal cross-section of all "neural activations" that happen and contribute to its (sequential) narrative even across (vastly) different contexts. Naively, the Concept of a "Bilauri" would be the cross-section of **dirty Bilauris, your-memory-of-one-specific-Bilauri, clean Bilauris in the cupboard, drinking from a Bilauri, filling it with water, etc., etc.** If one could invoke all these different entities then the neural-activation cross-section might stand for the concept of "Bilauri".

Did a Concept begin to form for you when reading this?[11] Hence, the process of being able to interpret the narrative and maybe forming the known-to-you-words: Cup, Bottle, Glass (?) translates the sequential narrative between the ** text ** into several pop-up Concepts.

On the role of actions: We assume that a Concept represents the "rule-set" (the narrative) that captures an entity. Certain rules will allow performing some actions and forbid others (like for a game, but also for objects, like apples, which you can: eat, roll around, throw, cut, cook, but not sit-on, stack, lever up, etc., etc.). Then we could infer that the set of all "by-this-rule-set-permitted" actions is a very good (not complete, but highly indicative) indicator of that Concept. Here we are now (finally) again in the realm of space and time! Hence, asking an agent what-all it could do "with it" (with that Concept in a given situation), might be a good way to see that the agents indeed "has" that Concept (has it mentally represented). This might help in the scientific verification problem: how can we show that an agent (an ape, an ANN, a robot) has indeed formed a Concept?

3.23 Discovery of Affordances Using Geometric Proximity Queries

Gabriel Zachmann (Universität Bremen, DE)

Main reference René Weller, Nicole Debowski, Gabriel Zachmann: "kDet: Parallel Constant Time Collision Detection for Polygonal Objects", Comput. Graph. Forum, Vol. 36(2), pp. 131–141, 2017.
URL http://dx.doi.org/10.1111/cgf.13113

Discovering affordances has been a long-term research question, with many different approaches both on the symbolic/ontological level [13, 4, 12] and on the geometric/simulation level [3, 6, 8, 5, 2]. It seems that the latter has, however, not received as many attention as the former. One reason might be that only in the past few years have simulation tools and geometric proximity queries become fast enough to be a viable option for affordance discovery.

We posit that with today's 3D geometric computing and simulation methods, the following problem can be tackled: given a task and a "target" object, find a "helper" object (tool) among a set of given potential helper tools with which the task can be achieved. For example, given a bottle and the task is to open the crown cap, a number of other helper objects can be used, besides the specifically designed tool. In order to make the problem amenable for simulation and geometric computing, a first simplification could be to reduce the task to a force vector and a point to which this should be applied on the target object. Thus, the goal will be to find a helper object (or several), a suitable 6D pose, and a suitable force and torque such that the sought-after force on the target object will result.

Finding suitable poses for the helper objects could perhaps be done using optimization methods with which the 6D space of all possible poses can be explored goal-oriented. Constraints would be that no other collisions occur than at the given target position. At the target position, some penetration is permissible and can be mapped into forces. Parameters for

[11] "Bilauri" is Swahili for "glass".

a goal function could be to maximize average distance between helper object and target object, or to maximize the lever. Suitable optimization methods should not require the computation of derivatives, as this would probably slow down the overall convergence. Therefore, methods such as particle swarm optimization or downhill simplex could be potential candidates.

One of the building blocks for the method above is proximity and penetration computation that should allow for extremely fast query times, no matter the complexity of the given objects. Therefore, a lot of acceleration data structures have been devised, such as Boxtrees [15], higher-order AABB-Trees [14], kd-trees, inner sphere trees [11], irregular grids [7], and many more.

Another trend is to intertwine acceleration data structures with the actual proximity queries, so that they become an integral part of the proximity query algorithm itself. One approach is to only build a skeleton or rudimentary data structure, then update or adapt the data structure at runtime on demand [1, 9]. Another approach is to reduce the sophistication of the acceleration data structure, thus the preprocessing time, and make it suitable for massive parallalization [10], so that it can be rebuilt from scratch every time the geometry has changed.

With the advent of machine learning techniques, other approaches to proximity queries for highly dynamic geometries might be possible as well. One idea could be to combine acceleration data structures with machine learning, so that such data structures can be built on-demand whenever the geometry has changed. This could probably imply that the data structure is suitable for parallel construction (e.g., on the GPU), so that all polygons of the geometry can be processed in parallel. One could also consider hybrid construction methods where the machine learning algorithm is used to guide the construction of the hierarchy in the regular top-down fashion (which is, usually, not optimal anyways). Furthermore, when dealing with pairs of objects or just patches of objects, it might be possible to train regression methods such that the answers to proximiy queries is learnt in their latent space. The latter would, most probably, be very time-consuming preprocessing.

References

1. Sigal Ar, Gil Montag, and Ayellet Tal. Deferred, Self-Organizing BSP Trees. In *Eurographics*, pages 269–278, September 2002.
2. Adam Allevato, Elaine Schaertl, Mitch Pryor, and Andrea Thomaz. Learning Labeled Robot Affordance Models Using Simulations and Crowdsourcing. In *Proceedings of Robotics Science and Systems (RSS)*, 2020.
3. Adam Allevato, Andrea Thomaz, and Mitch Pryor. Affordance Discovery Using Simulated Exploration. In *Proc. of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, AAMAS '18, page 2174 – 2176. International Foundation for Autonomous Agents and Multiagent Systems, 2018.
4. Daniel Bessler, Robert Porzel, Mihai Pomarlan, Michael Beetz, Rainer Malaka, and John Bateman. A Formal Model of Affordances for Flexible Robotic Task Execution. In *ECAI 2020*, G. D. Giacomo et al., Ed., pages 24s5–2432. IOS Press, 2020.
5. Alexandre Manoury, Sao Mai Nguyen, and Cédric Buche. Hierarchical Affordance Discovery Using Intrinsic Motivation. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, HAI '19, page 186 – 193. Association for Computing Machinery, 2019. https://doi.org/10.1145/3349537.3351898.
6. Huaqing Min, Chang'an Yi, Ronghua Luo, Jinhui Zhu, and Sheng Bi. Affordance Research in Developmental Robotics: A Survey. *IEEE Transactions on Cognitive and Developmental Systems*, 8(4):237–255, 2016.
7. Arsène Pérard-Gayot, Javor Kalojanov, and Philipp Slusallek. GPU Ray Tracing Using Irregular Grids. *Comput. Graph. Forum*, 36(2):477–486, may 2017.

8 Karl Sims. Evolving 3D Morphology and Behavior by Competition. *Artificial Life*, 1(4): 353–372, 1994.

9 Marek Vinkler, Vlastimil Havran, Jir? Bittner, and Jir? Sochor. Parallel On-Demand Hierarchy Construction on Contemporary GPUs. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1886–1898, 2016.

10 René Weller, Nicole Debowski, and Gabriel Zachmann. kDet: Parallel Constant Time Collision Detection for Polygonal Objects. In *Computer Graphics Forum (Proc. Eurographics)*, vol. 36, 2017.

11 Rene Weller, and Gabriel Zachmann. Inner Sphere Trees for Proximity and Penetration Queries. In *2009 Robotics: Science and Systems Conference (RSS)*. Seattle, WA, USA, June 2009. `http://cg.in.tu-clausthal.de/research/ist`.

12 Chang'an Yi, Huaqing Min, Jinhui Zhu, and Pengshuai Yin. Affordance Discovery Based on Intrinsic Motivation in Robots. In *Proc. of the IEEE Int'l Conference on Robotics and Biomimetics*. 2016.

13 Natsuki Yamanobe, Weiwei Wan, Ixchel G. Ramirez-Alpizar, Damien Petit, Tokuo Tsuji, Shuichi Akizuki, Manabu Hashimoto, Kazuyuki Nagata, and Kensuke Harada. A brief review of affordance in robotic manipulation research. *Advanced Robotics*, 31(19-20):1086–1101, 2017.

14 Gabriel Zachmann. Rapid Collision Detection by Dynamically Aligned DOP-Trees. In *Proc. of IEEE Virtual Reality Annual International Symposium; VRAIS '98*, pages 90–97. Atlanta, Georgia, March 1998.

15 Gabriel Zachmann. Minimal Hierarchical Collision Detection. In *Proc. ACM Symposium on Virtual Reality Software and Technology (VRST)*, pages 121–128. Hong Kong, China, 11–13 November 2002. `http://www.gabrielzachmann.org/`.

Participants

- Pedro Cabalar
University of Coruña, ES
- Zoe Falomir
Universitat Jaume I –
Castello de la Plana, ES
- Werner Kuhn
ETH Zürich, CH
- Sabine Timpf
Universität Augsburg, DE
- Florentin Wörgötter
Universität Göttingen, DE
- Gabriel Zachmann
Universität Bremen, DE

Remote Participants

- Stefano Borgo
National Research Council –
Povo (Trento), IT
- Anthony Cohn
University of Leeds, GB
- Stefania Costantini
University of L'Aquila, IT
- Simon Dobnik
University of Gothenburg, SE
- Gavin Duffy
Technological University
Dublin, IE
- Anna Foerster
Universität Bremen, DE
- Ulrich Furbach
Universität Koblenz-Landau, DE
- Shyamanta Moni Hazarika
Indian Institute of Technology
Guwahati – Assam, IN
- Thiago Pedro Donadon
Homem
IFSP – Sao Paulo, BR
- David Kirsh
University of California –
San Diego, US
- Felix Kroll
Technische Hochschule
Deggendorf, DE
- Pat Langley
Stanford University, US
- Marcos Lopes
Universidade de São Paulo, BR
- Bertram Ludäscher
University of Illinois at
Urbana-Champaign, US
- Chayanika Deka Nath
Tezpur University – Napaam, IN
- Ana-Maria Olteteanu
Freie Universität Berlin, DE
- Marianna Pagkratidou
Technological University
Dublin, IE
- Adam Pease
Articulate Software –
San Jose, US
- James Pustejovsky
Brandeis University –
Waltham, US
- Paulo E. Santos
Flinders University –
Adelaide, AU
- Simon Scheider
Utrecht University, NL
- Michael Sioutis
Universität Bamberg, DE
- John Stell
University of Leeds, GB
- Kristin Stock
Massey University –
Auckland, NZ
- Thora Tenbrink
Bangor University, GB
- Sabine Timpf
Universität Augsburg, DE
- Maria Vasardani
RMIT University –
Melbourne, AU
- Jago Williams
Bangor University, GB
- Diedrich Wolter
Universität Bamberg, DE