UNIVERSITY OF LEEDS

This is a repository copy of *Automated Commonsense Spatial Reasoning: Still a Huge Challenge*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/172683/

Version: Accepted Version

Book Section:

Bennett, B and Cohn, AG orcid.org/0000-0002-7652-8907 (2021) Automated Commonsense Spatial Reasoning: Still a Huge Challenge. In: Muggleton, S and Chater, N, (eds.) Human-Like Machine Intelligence. Oxford University Press . ISBN 9780198862536

This item is protected by copyright. This is an author produced version of a book chapter published in Human-Like Machine Intelligence. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Automated Commonsense Spatial Reasoning: Still a Huge Challenge

Brandon Bennett and Anthony G Cohn
School of Computing, University of Leeds, UK

OXFORD
UNIVERSITY PRESS

20

Automated Commonsense Spatial Reasoning: Still a Huge Challenge

20.1 Introduction

Achieving "commonsense reasoning" capabilities in a computational system has been one of the goals of Artificial Intelligence since its inception in the 1960s (McCarthy and Hayes, 1969; McCarthy, 1989; Thomason, 1991). However, as Marcus and Davis have recently argued (Marcus and Davis, 2019): "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 (Forbus *et al.*, 1991; Egenhofer, 1991; Freksa, 1992; Frank, 1992; Ligozat, 1993; Hernández, 1993; Gahegan, 1995; Zimmermann, 1993; Faltings, 1995; Escrig and Toledo, 1996; Gerevini and Renz, 1998; Moratz *et al.*, 2011; Mossakowski and Moratz, 2012) or in a precise logical language, such as classical first-order logic (Randell *et al.*, 1992; Gotts, 1994; Cohn, 1995; Borgo *et al.*, 1996; Cohn *et al.*, 1997; Galton, 1998; Pratt and Schoop, 1998; Pratt, 1999; Cohn and Hazarika, 2001; Galton, 2004).

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 will be explored in detail in the current chaper, is: what is the role of natural language in relation to commonsense spatial reasoning?

The main purpose of this 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 rest of 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.

Of course we do not claim that there has been no progress in addressing these challenges (and indeed we mention a few examples of works that do below), but it seems to us that these still represent considerable challenges to solve in the general case.

20.2 Commonsense Reasoning

In this section we examine the nature of commonsense reasoning and look at the ways in which research in computational Artificial Intelligence has sought to model and simulate human commonsense reasoning.

20.2.1 The Nature of Commonsense Reasoning

Although the specific processes by which human reasoning occurs are little understood, the meaning of the word 'reasoning' is relatively clear. It refers to any kind of process by which new implicit information is derived from given or assumed information. However, there are several different forms in which information is manifested and communicated. Fig. 20.1 illustrates those types of information and relationships that we consider to be particularly relevant to the understanding of different modes of reasoning.

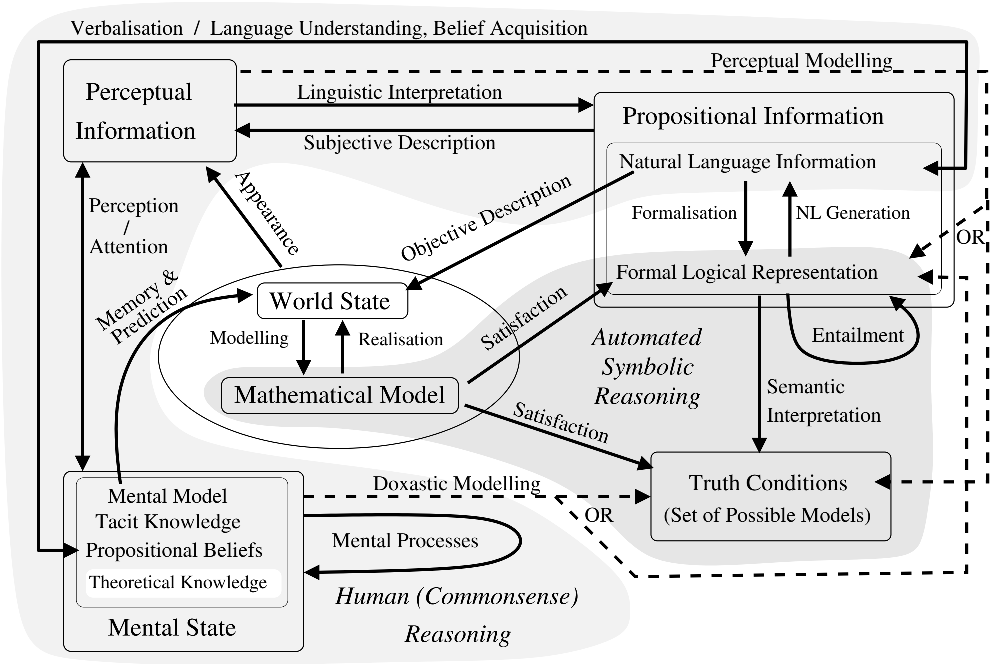


Fig. 20.1 Types of information and the relationships between them. Note that none of the arrows denote causal relationships, perhaps with the exception of the "mental processes" arrow; rather they denote a wide variety of other kinds of relationships such as epistemological and metalogical relationships.

One might define *commonsense* reasoning as that kind of reasoning that humans use in everyday situations, without explicit use of logical, mathematical or scientific theories. As such, the ambit of commonsense reasoning corresponds to the lighter grey region of the diagram, with its primary components being: *mental state*, *perceptual information* and *propositional information* (expressed in natural language).

Although the idea of an agent's 'mental state' is widely used in explanations of the behaviour of humans and animals, its constitution and function are poorly understood and we will not speculate on these; nor do we have the space here to consider the distinction between short and long term representations which are clearly important but not germane to our main argument here. For present purposes, we need only consider what types of information might in some way be stored within an agent's mind. We assume that a mental state includes some kind of *mental model* (Johnson-Laird, 1983), which somehow stores some correlate of received perceptual information in such a way that it can be used to remember or predict useful information about the state of the world. We also assume that the mental state incorporates *tacit knowledge* (Polanyi, 1966; Schacter, 1987; Kimble, 2013), which provides the agent with certain capabilities and skills (either instinctive or learned). Mental models and tacit knowledge are taken to be *non-linguistic* forms of information and hence can be possessed by agents with no linguistic capability. These kinds of information are difficult to articulate in verbal form. Researchers studying tacit knowledge often claim that it is impossible to convert it into an equivalent propositional form. Other researchers (e.g. in symbolic AI) take the view that it is very difficult but not impossible to specify an explicit symbolic correlate of tacit knowledge. We tend towards the latter view.

In the case of beings with linguistic capabilities, their mental state will also somehow store (or be able to generate) propositional information — i.e. internal correlates of natural language sentences. These correspond to the verbalisable beliefs of the agent. A special case of such beliefs would be theoretical knowledge of logic, mathematics or science. Although such theoretical knowledge may be part of the mental state of a sophisticated linguistically capable agent and applied in their processes, its use goes beyond what would be considered *commonsense* reasoning. (Hence, in Fig. 20.1, 'theoretical knowledge' is not within the light grey area of the diagram.)

There are several paths that reasoning can take. The most basic is where the appearance of the world generates perceptual information, which is (somehow) absorbed into an agent's mental state. Mental processes then take place that modify the current mental state to produce a new state that may include the results of some kind of inferential process. (We will not speculate on any details of how mental inference might operate.) Finally, the updated mental state may incorporate some prediction about the world state. This prediction is some piece of information that was not directly present in the perceptual information (nor in information derived by low level processing that takes place as part perception) but has been derived by the reasoning process.

The kind of reasoning just described does not necessarily involve any kind of linguistic information. Hence, it could be carried out by languageless animals. However, the diagram also includes a category of *propositional* information and indicates that information expressed in natural language may also play a part in human commonsense reasoning. Perceptual information may be converted into propositional information by linguistic interpretation. This can then be incorporated into an agent's mental state, in the form of propositional beliefs. Mental processes may then make use of this propositional information (often in combination with other types of information in the mental state) in order to draw inferences by some kind of mental argumentation process.

The reader will have noticed that the diagram also indicates a second mode of reasoning, which is ostensibly very different from commonsense: the darker grey area of the diagram demarcates the types of information that are manipulated by automated symbolic reasoning mechanisms. This kind of reasoning is relatively well-understood by mathematicians and computer scientists. However, it is only indirectly linked to the components of the commonsense reasoning system just described. The most overt connection between commonsense reasoning and automated reasoning occurs within the category of *propositional*[1] information. Here we have both natural language propositions (i.e. assertive sentences) and formulae of some logical language. We may map between these by procedures of formalisation (natural to formal) and natural language generation (formal to natural). However, establishing appropriate mappings has

[1] By propositional information, we mean information expressed in propositions of any form. We do *not* mean that the logical form of expressions is restricted to only atomic propositions and compounds formed using propositional operators. So any natural language assertive sentence or formula of some logical language would be an example of propositional information.

proved to be extremely difficult, especially in the direction natural-to-formal. There is huge controversy over how this should be done and even over whether it can be done at all in a general and reliable way. Moreover, even the details of a logical language suitable for capturing the meanings of natural language sentences are disputed, both in terms of the non-logical predicates that will be needed and in terms of the logical operators and structures that will be required.

Another linkage between commonsense and automated reasoning processes occurs indirectly *via* reality itself — i.e. in relation to the *world state* (the physical material and structure of reality). On one side of the connection, the world state interacts with commonsense reasoning in two ways: the world generates perceptual information; and the contents of a mental state somehow enable predictions to be made about the world state. On the other side of the link, the world state is regarded as having some correspondence (albeit usually very coarse grained) with a mathematical model that provides an interpretation (i.e. a *model* in the sense of model theoretic semantics) for the formal logical representation. Here again there is great controversy over what form an appropriate mathematical model should take and even whether it is possible to provide an adequate model at all.

We have also included in the diagram some further linkages indicated by dashed arrows. These are of a more putative nature. One possibility is that one might carry out some kind of 'perceptual modelling', that would map perceptual information either into a formal logic or into some other representation of truth conditions. we also indicate that some kind of 'doxastic modelling' could provide a mapping from mental state either to some formal logical representation or directly to truth conditions (which would then consist of the set of possible worlds that are compatible with the beliefs held within the mental state (Hintikka, 1962)). Our motivation in adding these links is to allow for the possibility that human commonsense reasoning could be simulated by an automated reasoning system without the need to use natural language as a bridging representation. Nevertheless, it is far from obvious how these links could be substantiated by an actual modelling process (although work on mental models has attempted to explicate a linkage to truth conditions and hence to human reasoning processes (Johnson-Laird, 1983)).

20.2.2 Computational Simulation of Commonsense Spatial Reasoning

A typical approach to developing computational commonsense spatial reasoning within the field of symbolic AI has been to design formal logical representations that are envisaged as being close to natural language forms of spatial description, and therefore similar to the kinds of propositional information used in human commonsense reasoning (Randell *et al.*, 1992; Borgo *et al.*, 1996). Since, reasoning with numerical information is generally considered to be mathematical rather than commonsense reasoning, the formal language is usually restricted to representing qualitative properties and relationships; hence the field is known as Qualitative Spatial Reasoning (QSR) (Ligozat, 2011; Cohn and Renz, 2008) which in turn is part of the larger field of Qualitative Reasoning (Forbus, 2019). For example, the Region Connection Calculus (RCC) has been widely used for a variety of purposes from modelling geographic information to representing activities in video. The two most common variants of RCC are RCC-8 and RCC-5 (8 and 5 referring to the number of jointly exhaustive and pairwise disjoint (JEPD) relationships in the calculus). The RCC-8 relations are depicted in Fig. 20.2; Egenhofer (1991) has postulated a similar set of relations from a different mathematical basis. The RCC calculi, along with most other QSR systems, can not only be structured into "conceptual neighbourhoods" as depicted in the figure, but also one can construct *composition tables* which enable inferences to be made about relationships between spatial relations not already explicit (e.g. from $\mathsf{NTPP}(a,b) \wedge \mathsf{TPP}(b,c)$ infer $\mathsf{NTPP}(a,c)$).

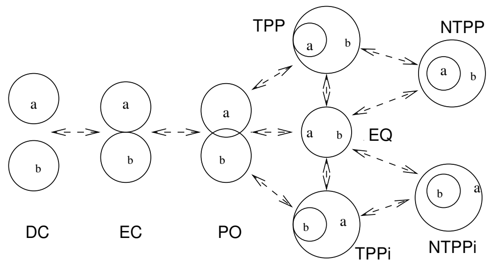


Fig. 20.2 A depiction of the RCC-8 relations. The connecting arcs indicate the "conceptual neighbourhood" – i.e. those neighbouring relations which one relation can transition to immediately, assuming continuous transitions or deformations. These 'conceptual neighbour' relations can be used to perform qualitative spatial simulations in order to predict possible futures, e.g. see Cui *et al.* (1992). RCC-5 is formed from RCC-8 by merging DC and EC, TPP and NTPP, and TPPi and NTPPi.

There are several potential problems with this approach. One is the difficulty of ensuring that the formal language developed is adequate for the kinds of reasoning that can be carried out by human commonsense. Indeed that has certainly not been achieved in a general way. The most that can be claimed is that formal representations have been developed that are capable of simulating some fragment of human commonsense reasoning. Several researchers have conducted psychological experiments to determine the extent to which sets of qualitative relations that have been used as the basis of qualitative spatial representations are comprehensible to human agents and are *cognitively plausible* — i.e. compatible with human spatial reasoning capabilities, e.g. (Klippel *et al.*, 2013; Knauff *et al.*, 1997). The latter concluded from their experiments that the more fine grained RCC-8 relations (rather than RCC-5) "are actually the most promising starting point for further psychological investigations on human conceptual topological knowledge. However, further evidence will be needed before a detailed modeling of human conceptual knowledge is possible". Knauff (1999) also investigated the cognitive adequacy of Allen's Interval Calculus (IA), which has 13 JEPD relations between intervals such as 'before', 'meets', 'overlaps', 'during'; the IA has often been used for reasoning about space. Knauff found that some evidence to support the cognitive adequacy of the IA, particular wrt to the associate composition table. However, whereas it had been postulated in the literature that errors in choice of a relation would normally be conceptual neighbours rather than random relations was not upheld in his experiments. Knauff also found that his results agreed with the 'mental model' theory that has been suggested as a human problem solving paradigm (Knauff *et al.*, 1998).

Another problem is that, as we have elaborated above, propositional information is not the only kind of information involved in human reasoning, and is probably not involved at all in many reasoning tasks, which are effected by manipulation of mental models (Ragni *et al.*, 2005) and/or tacit knowledge. Nevertheless, this need not necessarily be reason not to use formal propositional representations. Such representations have very general expressiveness, so it is plausible that even though the mind of an intelligent agent may be working with non-propositional tacit knowledge and/or neurologically implemented mental models, it might still be possible to encode the relevant information content in a propositional form. This possibility corresponds to the 'doxastic modelling' arrow in Fig. 20.1. Of course one cannot directly access the structure and content of a human mind; so the modelling would need to be done indirectly, by a process of hypothesis and testing.

A different approach is to model the perceptual information received by an agent and use this directly within a reasoning system. This is the approach taken within the *situated* approach to AI promulgated by Rodney Brooks, among others (Brooks, 1986).

20.2.3 But Natural Language is still a Promising Route to Commonsense

Despite the caveats of the last section, we still believe that natural language is likely to provide the most accessible entry point into commonsense spatial reasoning and provide fruitful insight into the semantic distinctions and inference patterns upon which it is based. If we could compute inferences from natural language information (e.g. text) that were judged to be broadly correct by humans, then we would have solved a large part of the problem of automating commonsense reasoning. In so far as non-linguistic information plays a part in commonsense, this would need to be somehow built into the inference generation mechanism. But, given the generality of logical reasoning techniques, there seems to be no obvious reason why this could not be done.

If we do choose to attack the problem of automating commonsense reasoning *via* the analysis of natural language, there are still several different ways in which this can be done. Contrasting views have been put forward by Davis (2013) and Bateman *et al.* (2010, Bateman (2013). Davis' analysis of spatial reasoning required for natural language text understanding begins by analysing the semantics of sentences in terms of the geometrical constraints that they seem to obey (in many cases identifying various constraints corresponding to different interpretations). Bateman's approach is to try to model the semantics of natural language terminology more directly, without attempting to resolve all ambiguities in their geometrical interpretation. The idea is that language-oriented inference rules can be formulated, which generalise over the variety of different ways in which natural language terminology can be employed in spatial descriptions.

Each of these approaches has its own problems:

With the Davis approach, the mapping from natural language to a formal representation is achieved by the expert judgement of a knowledge engineer. This leaves a significant gap in achieving automatic reasoning with natural language. Moreover, questions regarding why a particular interpretation was chosen may be difficult to answer.

The Bateman approach faces the following problem: on the one hand in many cases it will be difficult (sometimes impossible, we would suggest) to specify the semantics of a natural language term in a way that is sufficiently general to capture all its most common uses ; on the other hand it will be difficult so to do in a way that is not so general that it captures inferences that one would expect to only follow from particular applications of a term, which one would expect to have a more specific semantic interpretation. In other words, the Bateman approach may suffer due to a difficulty in distinguishing generality from ambiguity without recourse to some extra-linguistic semantic view point (such as mathematically specified geometrical constraints). An associated problem is that the Bateman approach seems to fall short of supplying truth conditions for propositions. This means that it is unclear what criteria would be used to judge the validity of an inference.

20.3 Fundamental Ontology of Space

Despite having been an object of enquiry for thousands of years, the constitution and structure of space and the material world is still a subject of much controversy. Although, scientific theories provide detailed accounts in terms of particles, fields and forces, the mathematical models developed by physicists are far removed from the terminology and informal inference patterns used in everyday description and reasoning about spatial properties and relationships. Over the past couple of decades the need for theories of space and material object that correspond more closely to natural modes of description have been recognised by many researchers in Artificial Intelligence and information science (Bennett, 2001; Masolo *et al.*, 2003; Grenon and Smith, 2004); nevertheless, some fundamental problems remain.

20.3.1 Defining the Spatial Extent of Material Entities

Fundamental to spatial reasoning is the association between material objects and spatial regions. However, determining the spatial extent of a material entity is complicated by the following considerations:

- The conditions for determining whether a particle is a constituent of a particular entity may be vague. For example, the surface of an animal (i.e. its skin) may have an outer layer incorporating dead or damaged cells, which are only loosely attached, and for which it is unclear which of them should be taken as constituents of the animal. Similarly, a rock may be made up of an agglomeration of rock particles, such that it can be unclear which are actually part of the rock and which are separate but ingrained within a cavity of the rock's surface.
- The exact positions of particles are unknown and intrinsically uncertain.
- Matter is made up of particles that are relatively sparsely distributed in space.
- Many materials (e.g. rock) contain tiny voids, such that it is not clear whether the volume of the voids should be considered as part of the material. To complicate matters, the voids may sometimes be filled with other materials such as water (Hahmann and Brodaric, 2012; Hahmann and Brodaric, 2014).

Let us suppose that intrinsic uncertainty and vagueness can be ignored. That is, we *assume* that:

- Although it may be difficult (or even impossible) to determine in practice, each physical entity is associated with a definite set of atoms (often combined into molecules).
- Although the positions of atoms are uncertain and constantly variable, it is possible in principle (though in most cases not in practice) to establish an assignment of a precise spatial location to each atom (e.g. by numerical coordinates), such that the resulting structure of spatially located atoms is sufficient to capture all aspects of the structure of an entity required to characterise its physical properties — except those that depend on sub-molecular scale details.

Even under the assumption that the particles that ultimately constitute matter have definite spatial locations, we still have the problem that these particles are relatively sparsely distributed in space, so that their combined spatial extent would be more like a scattered cloud of almost point-like regions than a continuously filled volume of space. One method of determining the spatial extent of a material entity would be by constructing an α-volume (Edelsbrunner *et al.*, 1983). This gives a well-defined procedure for determining a reasonable containing volume for an arbitrary set of points. The only problem with this is that it depends on the choice of a parameter, α, that determines what size of gap between points gets filled in to form the volume. When considering most physical entities, an α distance that is microscopic but considerably larger than the length of a molecule would be appropriate, since then the α-volumes of cells and larger entities would be continuous and connected, whereas if a smaller α distance were used, their α-volumes would have many gaps and discontinuities arising from the spaces between molecules. But in considering the structure of a molecule or atom, a much smaller α parameter would be needed, otherwise their α-volumes would be too course-grained to exhibit any distinctive spatial structure.

20.4 Establishing a Formal Representation and its Vocabulary

We may divide the analysis of natural language semantics into two parts: the elicitation of logical form (compositional structure) and the specification of content (meanings of terms).

20.4.1 Semantic Form

The application of automated symbolic reasoning techniques to natural language sentences requires that they be translated into a form that makes explicit their logical structure. For example, *"The pot contains lead"* would be represented by a formula such as $\exists x[\textit{"Pot"}(x) \wedge \textit{"Contains"}(x, \textit{"lead"})]$, which indicates predicative and quantificational structure but retains the vocabulary of the original sentence[2]. Performing this conversion is in general non-trivial. However, even where a sentence conveys spatial information, there is nothing particularly spatial about the logical form of the sentence. Hence, although this conversion is a problem for commonsense reasoning in

[2]Note that for simplicity we render the definite article "The" by the existential quantifier in this example.

general (at least if we want to automatically input information expressed in natural language into our commonsense reasoning system) it is not specifically a sub-problem of automating *spatial* commonsense reasoning.

20.4.2 Specifying a Suitable Vocabulary

By contrast, providing a semantics for the vocabulary of sentences conveying spatial information is a significant part of the commonsense spatial reasoning problem.

Specifying a suitable spatial vocabulary for a commonsense spatial reasoning system is a large and complex task. The concepts and relations used in natural language give a guide to the range of concepts required and the distinctions that one will want to make. However, the expression of these concepts in natural language is often highly ambiguous. Spatial phrases are applied in a wide variety of different situations, so that it is not obvious what is the core meaning or whether there are several different interpretations. To make matters worse, there can be considerable overlap between possible interpretations of different phrases. Such differences or overlapping of senses depend very much on the specific details of a particular spatial situation: in some situations, two phrases may seem to be equally appropriate, whereas in others one phrase will be much more apt than another.

In order to ensure the consistency and semantic rigour of a formal spatial vocabulary, the classification of meanings will need to be conducted in a systematic way. However, there are several different ways in which such a classification might be organised:

- Taxonomic — concepts are identified by successive differentiation of general concepts into more specific refinements.
- Compositional — a limited set of basic concepts/relations is used to construct a more comprehensive vocabulary. This in turn can be achieved in at least two different ways:
  - ∗ Analytic — a set of primitives is identified in order to provide fundamental conceptual units from which more complex concepts can be constructed by definitions that are expressed as structured combinations of the primitives. There is little if any overlap in the meaning of each primitive concept
  - ∗ Synthetic — key general concepts are identified from which more specific concepts can be constructed by combination. The key concepts may overlap so that specialisation may be achieved by their conjunction.

20.4.3 The Potentially Infinite Distinctions among Spatial Relations

A primary reason why a purely taxonomic approach is unlikely to achieve full generality is that ordinary language allows arbitrary elaboration of our descriptions of a spatial situation. An obvious way in which we can express limitless variety in spatial relations is by referring numerically to multiple sub-features of a spatial situation. For example, an entity could have any number of protruding sub-parts and these could be spatially related to some other entity in specific ways. Fig. 20.3 illustrates two relatively simple cases of the potentially infinite number of variants of relations between two disconnected regions that can occur when one region has multiple lobes, each of which protrudes into a distinct open cavity of the other region.

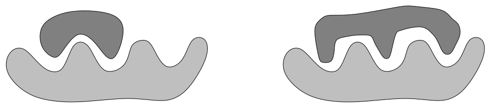
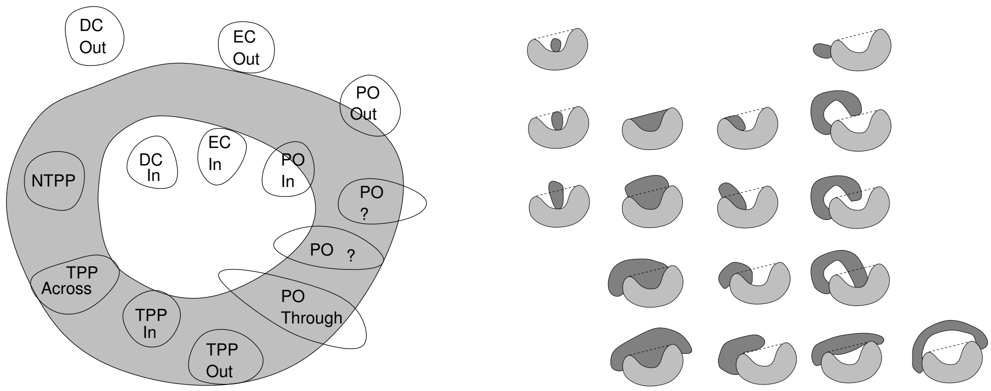


Fig. 20.3 Denumerable variants of disconnectedness with respect to multiple cavities.

Counting sub-features is a rather trivial way in which spatial relationships can be differentiated into more and more sub-types. However, there are other types of relation refinement that give rise to large numbers of distinctive spatial configurations. Fig. 20.4(a) illustrates how, in the context of considering the topological relationship between an arbitrary self-connected region and an entity with an internal cavity (the shaded 'doughnut'), the RCC-8 relations can be refined into various more specialised relations. Fig. 20.4(b) shows a large number of possible refinements of an external connection relation holding with respect to a region with an external concavity (cf (Cohn *et al.*, 1997)).



(a) (b)

Fig. 20.4 (a) Refining RCC-8 (Randell *et al.*, 1992) with respect to a region with an internal cavity (in 2D). (b) Refining the relation of external connection with respect to a region with an external concavity.

These examples suggest to us that the approach used widely in ontology construction, of specifying properties and relations by means of a taxonomy that successively refines concepts from general to specific, may not be the most appropriate for spatial concepts. In (Bennett *et al.*, 2013), it is suggested that a structured classification of spatial concepts consisting of an initial shallow hierarchy of topological relations supplemented by an open ended set of analytic definitions of more specialised relations, formulated by explicitly referring to entities such as surfaces and cavities, may be a better approach.

20.5 Formalising Ambiguous and Vague Spatial Vocabulary

Many spatial concepts have such a wide variety of uses that it seems to be impossible to give any concise definition that would cover all kinds of application of the concept (e.g. the general concept of 'place' (Bennett and Agarwal, 2007)). Nevertheless, such concepts somehow seem to give the impression of conveying a coherent and unproblematic meaning. As well as often involving concepts that are ambiguous — i.e. having several distinct (although possibly overlapping) meanings — spatial concepts may also be *vague* — i.e. subject to gradations of meaning with no clear cut-off point regarding the applicability of the concept (Bennett, 2011). A number of researchers have investigated ways in which typical cut-off points or ranges of cut-off-points can be elucidated from human subjects (Mark and Egenhofer, 1994; Mark *et al.*, 1995; Montello *et al.*, 2003).

In this section we shall consider a variety of examples that illustrate the ubiquity of ambiguity and vagueness in natural language spatial vocabulary and thereby indicate the scale of the difficulty that these phenomena pose to the endeavour of formalising commonsense spatial reasoning based on natural language expression of spatial information.

20.5.1 Crossing

Phrases of the form 'x crosses y' are very common in spatial descriptions, and the notion of one entity crossing another seems to convey a basic notion. But such phrases can have a wide range of different interpretations. The nature of the relationship referred to can often be determined (or at least narrowed down) by knowledge of the type of the entities x and y that are involved, but additional background knowledge may also be needed in order to disambiguate the meaning.

Some different interpretations are as follows:

- A flat elongated entity is part of a surface and runs from one edge of the surface to another. For example a path may cross a park.
- Two elongated entities may intersect (typically, approximately at right angles) at some location that is at a mid-point (or mid-section) of both of them. For example two roads may cross.
- An entity may cross a barrier by passing through a hole in the barrier.
- An entity may cross a barrier and also be part of that barrier. For example a protein that is part of a cell membrane and protrudes both into the cytoplasm and out to the exterior of the cell.
- A line or linear entity or a three-dimensional entity may cross a surface by having a part on one side of the surface and a part on the other side of the surface.
- An entity may cross another entity by going over it from one side to another.
- There are also many dynamic interpretations of 'cross' — as in "The runner crossed the finishing line". The dynamic interpretations vary in similar ways to the static interpretations. (Such interpretations will not be considered further in this report.)

20.5.2 Position Relative to 'Vertical'

The concepts of 'above', 'below', 'over', 'under, 'beneath', depend on having some notion of the directions 'up' and 'down' in relation to the entities being considered. The clearest cases are where 'up' and 'down' are interpreted according to the reference frame of our planet Earth itself. 'Up' normally means away from the centre of the Earth, whereas down is towards the centre of the Earth.

However, there is a certain amount of ambiguity in these relations when we consider the variety of situations in which they might be judged to apply. Some possibilities are illustrated in Fig. 20.5. The cases shown are as follows:

(Fig. 20.5.a) Here every point of the lower region is directly below some point of the upper region (and no point of the upper region is below any point of the lower region).

(Fig. 20.5.b) Here every point of the upper region is directly above some point of the lower region (and no point of the upper region is below any point of the lower region).

(Fig. 20.5.c) Here some points of the upper region are above some points of the lower region (and no point of the upper region is below any point of the lower region).

(Fig. 20.5.d) Here every point of the upper region is higher than every point of the lower region, even though no point of either region is above (or below) any point of the other region.

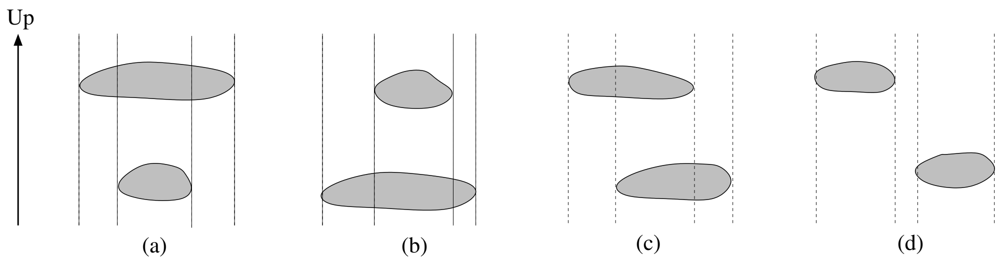


Fig. 20.5 Variants of the Above-Below Relationship.

It is worth noting that only cases (a) and (b) are transitive. Also, out of all the cases, (a) and (b) seem to be the most typical examples that one would describe by saying one region was above (or below) the other. Hence, is tempting to interpret 'x is above/below y as holding whenever either of the situations (a) and (b) occurs. However, if we take $\mathsf{above}(x, y)$ as holding when either of the situations (a) and (b) occurs, then this relation is not transitive. By referring explicitly to the relative vertical positions of points (or parts) of the entities concerned, it is possible to define a variety of different relations that describe the relative vertical positions of extended entities. However, these will have different semantics, and hence, different import with respect to logical entailment. Moreover, the mappings between these and relations referred to in natural language will be ill-defined.

20.5.3 Sense Resolution

In the works (Bennett *et al.*, 2013; Bennett and Cialone, 2014), the authors investigated (by analysing a text corpus obtained from a large Biology text book (Reece *et al.*, 2011)) the use of the spatial relation terms 'contain', 'enclose' and 'surround'. For each occurrence of any of these words (and cognate forms) the geometrical configuration to which the word was being applied was determined (from the text and with the use of auxiliary reference works). The authors found that around 15 different geometric conditions seemed to cover all usages of the terms. They calculated the frequencies with which each word was used to describe a particular geometrical constraint and found that each word could be applied to a variety of different geometrical constraints and that there was considerable overlap in the use of the terms. However, there were also significant differences in the frequencies that a particular term was applied to a given situation, with each having different typical and atypical uses.

Given the many-many correspondence between natural language spatial terminology and geometrical constraints, commonsense spatial reasoning conducted on the basis of linguistic descriptions must employ some method of ascertaining the intended meaning of a given work use. By *sense resolution* we mean the mechanism by which a word or phrase with multiple possible interpretations is associated with a particular axiomatically defined predicate. The following sentences all use the word 'surround', but in each it refers to a different spatial relation: "The embryo is surrounded by amniotic fluid"; "The embryo is surrounded by a shell"; "The cell is surrounded by its membrane"; "The garden is surrounded by a wall"; "The building is surrounded by guards". To reason on the basis of one of these sentences we need to know what spatial relation is intended. Many factors place constraints on possible interpretations of a lexical predicate. An important consideration is the type(s) of thing to which the predicate is applied. Surrounding by a fluid is different from surrounding by a rigid shell, or a wall or a group of people.

As a further illustration, contrast the meaning of 'contains' in "This bottle contains wine" and "The wine contains alcohol". In the first case, the wine is located within a cavity enclosed by, but not overlapping the bottle, whereas in the second alcohol is an ingredient of the wine. These are very different spatial relations. Establishing a robust automated mechanism for sense resolution would be a significant advance towards achieving automated commonsense reasoning.

20.6 Implicit and Background Knowledge

Suppose we take a convincing chain of reasoning expressed in natural language and translate it into a logical language (e.g. first-order logic): we are unlikely to get a formally valid sequence of inferences. It could be that the formal language that we use does not incorporate the types of logical operations required to articulate the inferences. But even if the logical language is sufficiently expressive in terms of its logical expressivity we are still unlikely to get a valid argument. This is because the reasoning will in typical cases also depend heavily on various kinds of implicit knowledge that is covertly utilised within commonsense reasoning processes.

One source of additional information would be the definitions of spatial (and other) terms and the axioms that specify semantic properties of primitives from which these terms are defined — in other words semantic knowledge. In addition to this there is a huge amount of contingent background information that can be potentially drawn upon to facilitate commonsense reasoning. For instance, knowledge about particular spatial properties and configurations of various kinds of object (tools, buildings, people etc). This kind of knowledge is often called *commonsense knowledge* and is critical to effective commonsense reasoning[3]. (Davis, 2017) discusses such knowledge at length and categorises it as, "roughly, what a typical seven year old knows about the world, including fundamental categories like time and space, and specific domains such as physical objects and substances; plants, animals, and other natural entities; humans, their psychology, and their interactions; and society at large". Encoding and storing such background information is a major goal of the long running AI project, CYC (Guha and Lenat, 1990). However, despite several decades of research the original goal of CYC project is still to be achieved.

As well as supplying additional information, background knowledge may also be used to select between different possible interpretations of vague and ambiguous vocabulary terms — i.e. to facilitate *sense resolution*, as described in the previous section. This seems to be particularly important in the resolution of certain ambiguous spatial relationships.

A kind of tacit knowledge that is particular to spatial reasoning is our ability to transfer information between multiple reference frames without any explicit expression of the relationships between these frames or of the reasoning steps that we must somehow be performing.

Geometry and physics represent space and time by coordinate systems defined relative to some reference frame regarded as fixed. Often a single reference frame will suffice, even for a complex physical situation, and, where multiple frames are used, precise mappings between them are defined. This contrasts sharply with natural language, which typically jumps quickly between multiple reference frames. E.g.: "The girl hid behind the curtains, but was visible through the window from the front of the house. The policeman in the garden saw the girl but not the lion in the living room." To model human-like reasoning, we need a representation that can capture such chains of relative location. One object may be used to locate another either directly (using phrases like 'behind', 'through', 'in front of') or indirectly via background knowledge (e.g. the typical relative locations of curtains, windows, houses, gardens and living rooms). Such relative locations have been an important research focus in QSR (Donnelly, 2005) and Scheider's contribution in (Gangemi *et al.*, 2014).

20.7 Default Reasoning

It is widely recognised that commonsense reasoning is often *non-monotonic* in nature. This means that there are cases where is reasonable to infer some conclusion ϕ from

[3]Such implicit knowledge is critical in Winograd Schema Challenge problems. An example such problem from the corpus at `https://cs.nyu.edu/faculty/davise/papers/WinogradSchemas/WSCollection.html` is "Jim signaled the barman and gestured toward his [**empty glass/bathroom key**]. Whose [empty glass/bathroom key]? Answers: **Jim/the barman**". Although this is a spatial reasoning problem, background (implicit) knowledge about objects involved is required to perform the appropriate inference.

some set of information Σ, but if we acquire some additional information α, the conclusion ϕ is no longer warranted. (Symbolically, we may have $\Sigma \mid\!\sim \phi$, but $\Sigma, \alpha \not\mid\!\sim \phi$, where '$\mid\!\sim$' is a commonsense inference relation.) A number of logical calculi have been developed to formalise non-monotonic reasoning, the best known being Reiter's *default logic* (Reiter, 1980) and the *circumscription* theory McCarthy (1986).

Typical examples of default reasoning are inferences such as: If x is a bird, then x can fly (unless you know that x is a penguin or other flightless bird); Gert is German, therefore Gert drinks beer (but not if we know that Gert is 3 years old).

It seems that relatively little work had been done on specifically spatial modes of non monotonic reasoning (though see, e.g., Walega et al (Walega *et al.*, 2017). Yet there are many reasoning examples that suggest that commonsense spatial reasoning is very often supported by simplifying assumptions regarding the spatial properties of objects and configurations. Here are some examples:

- When reasoning with information concerning objects situated in an environment, in many cases we assume that space is empty except for those parts that we know to be occupied by physical objects or matter. (How this affects reasoning about objects moving in space has been considered in detail by (Shanahan, 1995)).
- A spatial extent can be assumed to be convex if nothing is known to the contrary. For example, reasoning about objects fitting into containers typically assumes we are dealing with convex objects in containers whose containing space is convex, unless we have explicit information to the contrary[4].
- If a region is known to be small relative to some other regions then the small region can usually be assumed to behave like a point with regard to inferences involving these regions. For instance, if we know there is a gap between two 'large' objects, we will tend to assume that a 'small' object will fit through it.

There is an interesting relationship between default reasoning and mental models that could be potentially useful in the implementation of commonsense reasoning. A mental model it often regarded as storing a mental correlate of the most typical way in which a set of beliefs could be realised. Thus construction of a mental model may be seen as the limit of default reasoning, where although one's knowledge does not fully pin down the state of the world, one constructs a prototypical example situation that is compatible with that knowledge and uses that as a basis for reasoning (Knauff *et al.*, 1995). Clearly, such reasoning is not deductively valid; but the inferences drawn could be useful in many cases.

20.8 Computational Complexity

From the point of view of traditional computer science, the most obvious difficulty facing the development of automated commonsense spatial reasoning is computational complexity. Indeed many of the intractable or undecidable problems studied in computational complexity theory are spatial in nature. A result of (Grzegorczyk, 1951) proves undecidability of some relatively simple topological theories due to the fact that they

[4]This assumption is required in order to reason appropriately about the *Winograd Schema*: "The trophy would not fit in the case because it was too **big/small**"(Levesque *et al.*, 2012).

can encode arithmetical operators and formulae. It is actually very straightforward to model numbers in terms of multi-piece spatial regions. We simply take the number of components of such a region as representing a number. It is then possible to define equality as a spatial relation and addition and multiplication as spatial constructions.

The extremely high expressive power of spatial concepts is also demonstrated by Tarski's (1956) paper on the definability of concepts, which shows that any concept that can be fully axiomatised within a theory that includes concepts sufficient to axiomatise Euclidean geometry, can actually be defined in terms of the geometrical concepts. This means that the logical properties of any concept can be fully modelled in terms of geometrical properties, with no need for any further axioms, since the standard axioms of geometry together with the geometrical definition of the concept are sufficient (Bennett, 2004).

If we restrict attention to reasoning problems formulated in terms of the limited sets of predicates and operators typically used in QSR, the complexity results are somewhat better, but still rather discouraging. For example, Renz and Nebel (1999) identified maximal tractable subsets of a topological constraint language based on the RCC-8 relation set (Randell *et al.*, 1992). These subsets can be used to carry out useful spatial reasoning tasks, but it is disappointing that it is not possible to reason effectively with more expressive extensions of these languages. Certain tractable extensions have been identified (e.g. by (Gerevini and Renz, 1998), who devised a reasoning algorithm for a combination of topological and size constraints). However, it seems that as we add expressive power in terms of combining different types of spatial property and relation, we very quickly end up with computationally intractable reasoning problems (see e.g. (Davis *et al.*, 1999)) unless we strictly limit other aspects of the representation language. Also, introducing natural global constraints to a reasoning problem, such as requiring regions to be self-connected and/or embeddable in the plane tends to raise complexity of spatial reasoning problems and often results in undecidability (Dornheim, 1998).

However, one should bear in mind that the type of problem for which these unpalatable complexity results arise, is very different from the circumstances to which one would expect commonsense reasoning to be applied and the types of computation performed (e.g. consistency checking of networks of spatial relations) are quite far removed from everyday reasoning tasks. Whereas existing automated spatial reasoning systems typically carry out exhaustive reasoning with respect to large numbers of spatial constraints expressed a very limited set of spatial relationships, commonsense spatial reasoning typically operates with a small number of spatial facts expressed using a rather wide vocabulary of spatial properties and relations. Thus, although computational complexity is clearly a problem for computational spatial reasoning, it is not necessarily a problem for automated *commonsense* spatial reasoning, since existing spatial reasoning algorithms seem to be doing something very different from what one would expect from commonsense reasoning – humans can clearly make comonsense spatial inferences rather quickly.

20.9 Progress towards Commonsense Spatial Reasoning

The preceding sections have been largely negative in tone: pointing out the challenges in endowing machines with commonsense spatial reasoning. Of course there has been progress towards this goal, and indeed we have already mentioned some of this in the earlier part of this chapter. In this section we briefly mention some of the highlights of such work. Foremost in this direction is the work on qualitative spatial representation and reasoning. 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. (Renz and Nebel, 1999). There are toolkits for reasoning with many of these, such as SparQ (Wolter and Wallgrün, 2013) and for extracting QSRs from video data, e.g. QSRlib (Gatsoulis *et al.*, 2016). Moreover there are many implemented systems, particularly in the domain of activity understanding which exploit QSR (e.g. (Duckworth *et al.*, 2019) or which learn about spatial relations (e.g. (Alomari *et al.*, 2017)) 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 Davis and Marcus (2015). Davis has contributed much to the field of commonsense reasoning, and spatial reasoning in particular e.g. his work on liquids (Davis, 2008) and containers (Davis *et al.*, 2017).

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 Cohn and Gotts (1996), or ways of smoothing noisy detections (e.g Sridhar *et al.* (2011), building probabilistic models of QSR, e.g. Kunze *et al.* (2014), or by explicitly reasoning about occlusion, e.g. Bennett *et al.* (2008).

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.

20.10 Conclusions

In this chapter we have decomposed 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 have little discussed is how a commonsense knowledge could be acquired by an automated reasoning system, and in particular spatially related knowledge. 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 (Mitchell *et al.*, 2018) aims to learn such knowledge by learning 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 Alomari *et al.* (2017) 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 Richard-Bollans *et al.* (2020) 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 have hardly discussed is how *embodiment* affects perception and spatial awareness. Tversky, among others, (e.g. 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" (Tversky, 2009). There is work in AI which takes an embodied approach to spatial cognition and spatial commonsense (e.g. (Alomari *et al.*, 2017; Spranger *et al.*, 2014)) but more research on this is certainly needed.

Most of the problems we have discussed actually apply to commonsense reasoning in general, rather than exclusively to spatial reasoning; and yet in the examples we have considered, 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 was not intended to be prescriptive of a particular research direction or methodology[5]. As well as exposing a large number of problems, we have indicated a variety of different approaches that might lead to their solution. Our aim was 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.

20.11 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.

[5]Davis and Marcus (2015) suggests 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 (Hasan *et al.*, 2020)

References

Alomari, Muhannad, Duckworth, Paul, Hogg, David C, and Cohn, Anthony G (2017). Natural language acquisition and grounding for embodied robotic systems. In *Proc. AAAI*.

Bateman, John A. (2013). Space, language and ontology: a response to Davis. *Spatial Cognition & Computation*, **13**(4), 295–314.

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

Bennett, Brandon (2001). Space, time, matter and things. In *Proc. FOIS'01* (ed. C. Welty and B. Smith), Ogunquit, pp. 105–116. ACM.

Bennett, Brandon (2004). Relative definability in formal ontologies. In *Proceedings of the 3rd International Conference on Formal Ontology in Information Systems (FOIS-04)* (ed. A. Varzi and L. Vieu), pp. 107–118. IOS Press, Amsterdam.

Bennett, Brandon (2011). Spatial vagueness. In *Methods for Handling Imperfect Spatial Information* (ed. R. Jeansoulin, O. Papini, H. Prade, and S. Schockaert). Springer-Verlag.

Bennett, Brandon and Agarwal, Pragya (2007). Semantic categories underlying the meaning of 'place'. In *Spatial Information Theory: proceedings of the 8th international conference (COSIT-07)*, Volume 4736, Lecture Notes in Computer Science, pp. 78–95. Springer, Berlin, Heidelberg.

Bennett, Brandon, Chaudhri, Vinay, and Dinesh, Nikhil (2013). A vocabulary of topological and containment relations for a practical biological ontology. In *Proc. COSIT-13* (ed. J. Stell, T. Tenbrink, and Z. Wood). Springer.

Bennett, Brandon and Cialone, Claudia (2014). Corpus guided sense cluster analysis: a methodology for ontology development (with examples from the spatial domain). In *Proc. FOIS-14* (ed. P. Garbacz and O. Kutz). IOS Press.

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.

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.

Brooks, Rodney A. (1986). A robust layered control system for a mobile robot. *IEEE J. Rob. Autom.*, **2**(1), 14–23.

Cohn, Anthony G. (1995). A hierarchical representation of qualitative shape based on connection and convexity. In *Proc. COSIT-95*, pp. 311–326.

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

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.

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

Cohn, Anthony G and Renz, Jochen (2008). Qualitative spatial representation and reasoning. In *Handbook of Knowledge Representation* (ed. F. van Harmelen, Lifschitz, V., and B. Porter), pp. 513–596. Elsevier.

Cui, Z, Cohn, A G, and Randell, D A (1992). Qualitative simulation based on a logical formalism of space and time. In *Proc. of AAAI-92*, pp. 679–684.

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

Davis, Ernest (2013). Qualitative spatial reasoning in interpreting text and narrative. *Spatial Cognition & Computation*, **13**(4), 264–294.

Davis, Ernest (2017). Logical formalizations of commonsense reasoning: A survey. *J. Artif. Intell. Res.*, **59**, 651–723.

Davis, Ernest, Gotts, Nick, and Cohn, Anthony G. (1999). Constraint networks of topological relations and convexity. *Constraints*, **4**(3), 241–280.

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

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

Donnelly, Maureen (2005). Relative places. *Applied Ontology*, **1**(1), 55–75.

Dornheim, C. (1998). Undecidability of plane polygonal mereotopology. In *Proc. KR-98*, pp. 342–353.

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

Edelsbrunner, Herbert, Kirkpatrick, David G., and Seidel, Raimund (1983). On the shape of a set of points in the plane. *IEEE Trans. on Inf. Theory*, **29**(4), 551–558.

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

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

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.

Forbus, K. D. (2019). *Qualitative Represention*s. MIT Press.

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

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

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.

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

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

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

Gangemi, Aldoand Hafner, Verena V., Kuhn, Werner, Scheider, Simon, and Steels, Luc (Eds) (2014). *Dagstuhl seminar 14142 on "Spatial reference in the Semantic Web and Robotics"*.

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*.

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

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.

Grenon, Pierre and Smith, Barry (2004). Snap and span: Towards dynamic spatial ontology. *Spatial cognition and computation*, **4**(1), 69–104.

Grzegorczyk, A. (1951). Undecidability of some topological theories. *Fundamenta Mathematicae*, **38**, 137–152.

Guha, R.V. and Lenat, D.B. (1990). CYC: a mid-term report. *AI Mag.*, **11**(3), 32–59.

Hahmann, Torsten and Brodaric, Boyan (2012). The void in hydro ontology. In *Proc. FOIS 2012*, pp. 45–58.

Hahmann, Torsten and Brodaric, Boyan (2014). Voids and material constitution across physical granularities. In *Proc. FOIS*, pp. 51–64.

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)*.

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.

Hintikka, Jaakko (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, N. Y.

Johnson-Laird, Philip (1983). *MentalModels: Toward a Cognitive Sience of Language, Inference and Consciousness*. Harvard University Press.

Kimble, Chris (2013). Knowledge management, codification and tacit knowledge. *Information Research*, **18**(2), paper 577.

Klippel, Alexander, Wallgrün, Jan Oliver, Yang, Jinlong, Mason, Jennifer S., Kim, Eun-Kyeong, and Mark, David M. (2013). Fundamental cognitive concepts of space (and time): Using cross-linguistic, crowdsourced data to cognitively calibrate modes of overlap. In *Proc.COSIT*, pp. 377–396.

Knauff, Markus (1999). The cognitive adequacy of Allen's interval calculus for qualitative spatial representation and reasoning. *Spatial Cognition and Computation*, **1**(3), 261–290.

Knauff, M., Rauh, R., and Renz, J. (1997). A cognitive assessment of topological spatial relations: Results from an empirical investigation. In *Proc. COSIT'97,*.

Knauff, M, Rauh, R, and Schlieder, C (1995). Preferred mental models in qualitative spatial reasoning: A cognitive assessment of Allen's calculus. In *Proc. 17th Annual Conf. of the Cognitive Science Society*.

Knauff, Markus, Rauh, Reinhold, Schlieder, Christoph, and Strube, Gerhard (1998). Mental models in spatial reasoning. In *Spatial cognition*, pp. 267–291. Springer.

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*.

Levesque, Hector J., Davis, Ernest, and Morgenstern, Leora (2012). The Winograd Schema Challenge. In *Proc. KR*.

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.

Ligozat, G F (2011). *Qualitative Spatial and Temporal Reasoning*. Wiley-ITSE.

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

Mark, D, Comas, D, Egenhofer, M, Freundschuh, S, Gould, J, and Nunes, J (1995). Evaluating and refining computational models of spatial relations through cross-linguistic human-subjects testing. In *Proc. COSIT*, pp. 553–568.

Mark, David and Egenhofer, Max (1994). Calibrating the meanings of spatial predicates form natural language: Line-region relations. In *Proc. SDH*, pp. 538–553.

Masolo, Claudio, Borgo, Stefano, Gangemi, Aldo, Guarino, Nicola, and Oltramari, Alessandro (2003). WonderWeb Deliverable D18: DOLCE Ontology Library. Technical report, Laboratory For Applied Ontology - ISTC-CNR. `http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf`.

McCarthy, J (1986). Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, **28**, 89–116.

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

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.

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.

Montello, Daniel R, Goodchild, Michael F, Gottsegen, Jonathon, and Fohl, Peter (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition & Computation*, **3**(2-3), 185–204.

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.

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

Polanyi, Michael (1966). *The Tacit Dimension*. Routledge and Kegan Paul.

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

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

Ragni, Marco, Knauff, Markus, and Nebel, Bernhard (2005). A computational model for spatial reasoning with mental models. In *Proc. CogSci-05*.

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.

Reece, Jane B., Urry, Lisa A., Cain, Michael L., Wasserman, Steven A., Minorsky, Peter V., and Jackson, Robert B. (2011). *Campbell Biology* (9th Edition edn). Pearson, New York.

Reiter, R (1980). A logic for default reasoning. *Artificial Intelligence*, **13**(1,2), 81–132.

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.

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

Schacter, Daniel L. (1987). Implicit memory: history and current status. *Journal of Experimental Psychology*, **13**(3), 501–518.

Shanahan, M (1995). Default reasoning about spatial occupancy. *Artificial Intelligence*, **74**(1), 147–163.

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.

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.

Tarski, A. (1956). Some methodological investigations on the definability of concepts. In *Logic, Semantics, Metamathematics*. Oxford Clarendon Press. trans. J.H. Woodger.

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.

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

Walega, Przemyslaw Andrzej, Schultz, Carl, and Bhatt, Mehul (2017). Non-monotonic spatial reasoning with answer set programming modulo theories. *Theory and Practice of Logic Programming*, **17(2)**, 205–225.

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.

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