Lexical inference as a spatial reasoning problem

Steven Schockaert
Cardiff University, UK
SchockaertS1@cardiff.ac.uk

Abstract

Humans are often able to draw plausible conclusions from incomplete information based on background knowledge about the world. As a substantial part of this background knowledge is of a lexical nature, a crucial challenge in automating plausible reasoning consists in learning how different words are semantically related. This paper argues that most of the lexical relations that we need for plausible reasoning can be identified with qualitative spatial relations in semantic spaces, i.e. high-dimensional Euclidean spaces in which words are represented as geometric objects. This leads us to treat lexical inference as a qualitative spatial reasoning problem, and allows us to combine distributional representations with relation extraction methods and existing lexical resources.

1 Introduction

In applications such as question answering, semantic web search and legal informatics, there is a need for systems that can draw plausible conclusions from knowledge that may be vague, inconsistent, uncertain or ambiguous. There are at least two fundamental reasons why classical logic is ill-suited for this purpose. First, while classical logic is aimed at reasoning about tautologies, human knowledge rather tends to capture statistical regularities (e.g. "most birds can fly"). This discrepancy is well-understood and has been addressed in a variety of frameworks for default reasoning [25; 27; 22] and probabilistic extensions of classical logic [18; 28]. Second, in most applications, available knowledge is based on natural language, in the sense that the atomic concepts, propositions or predicates which are considered tend to correspond to natural language concepts (e.g. because this knowledge has been provided by an expert or has been extracted from text documents). In such cases, humans often rely on background knowledge about the meaning of the underlying concepts to find plausible conclusions. Consider the following example:

Subsumption: Knowing that a restaurant only sells pizza, we conclude that it only sells Italian food.

Here we rely on the knowledge that pizzas are a kind of Italian food. Such subsumption relationships are available in many existing lexical resources such as WordNet[1], ConceptNet[2], CYC[3], and many domain-specific taxonomies. Using subsumption relations for logical reasoning is straightforward, although default logics may be more suitable for this purpose than classical logic, since most subsumption relations are not perfect (e.g. the restaurant could also sell American-style pizzas). Inference methods for reasoning about natural language statements based on subsumption relations have also been proposed. For example, the natural logic from [24] can be used to extend subsumption at the word level to subsumption of phrases, e.g. deriving that "only sells pizza" entails "only sells Italian food" while "does not sell pizza" is entailed by "does not sell Italian food". Beyond subsumption reasoning, similarity based reasoning is one of the most common forms of lexical inference:

Similarity: Knowing that Mary enjoys hiking in the Alps, we conclude that she would also enjoy hiking in the Pyrenees.

where we rely on the knowledge that the Alps and Pyrenees share many characteristics. In general, the intuition behind similarity based reasoning is that from "if A then B" and "A is similar to A^*", we can conclude that "if A^* then B is likely". In some cases the conclusion can be weakened to a statement of the form "if A^* then a property similar to B should hold". This inference pattern has been well-studied in the context of fuzzy set theory [43; 30; 12], although similarity based reasoning has also been implemented using neural networks [38] and in probabilistic settings [4]. Clearly, the higher the degree of similarity between A and A^*, the more likely that properties of A transfer to A^*. However, there seems to be no principled way of linking degrees of similarity to probabilities. As a result, in practical applications, similarity based reasoning is mostly used to deal with near-synonyms [4; 17]. Interpolation can be seen as a variation of similarity based reasoning, in which the use of similarity degrees can be avoided:

Interpolation: Knowing that sandwich shops and restaurants in Wales are both required to display food hygiene

[1] `http://wordnet.princeton.edu`
[2] `http://conceptnet5.media.mit.edu`
[3] `http://www.cyc.com`

ratings, we conclude that cafes in Wales are also required to display such ratings.

The assumption underlying interpolative reasoning is that when B is conceptually between A and C, all properties which A and C have in common should apply to B as well. In the example, we are able to plausibly infer something about cafes, which we can regard as being conceptually between sandwich shops and restaurants (e.g. offering a wider range of food than sandwich shops as well as more seating, while being less formal than restaurants). We can think of interpolation as a more robust (albeit computationally less efficient) alternative to similarity based reasoning, which is not restricted to reasoning about near-synonyms. In [11], interpolation was shown to consistently outperform similarity based reasoning in a classification setting. Other approaches that incorporate some of the intuitions behind similarity based reasoning while avoiding the use of similarity degrees include the idea of statistical predicate invention [21] and the use of Bayesian models with a prior probability based on the (latent) structure of the considered domain [39]. The following example illustrates how plausible reasoning can be used to make sense of inconsistent information:

Consensus: Suppose an online shop sells an item which is referred to as 'original art' in the title and as a 'poster' in the detailed description; we conclude that the item may be a limited-edition art print.

The assumption in this example is that 'original art' and 'poster' are disjoints concepts, which makes the description inconsistent, although limited-edition art prints (e.g. lithographs) can be seen as a borderline case for both concepts and thus offer a plausible way to resolve the inconsistency. Resolving inconsistencies in this way amounts to finding a consensus, i.e. finding an interpretation which is similar to interpretations of each of the conflicting pieces of information; see [36] for a formalisation of this idea. Finally, we consider an example where lexical knowledge is combined with other background knowledge:

A fortiori: Knowing that university staff are not permitted to travel in business class, we conclude that staff are not permitted to travel in first class.

The conclusion is based on the commonsense knowledge that the reason why staff are not permitted to travel in business class is because it is too expensive. Given this extra piece of information, the conclusion follows from the lexical knowledge that first class travel is more expensive than business class travel.

What each of the aforementioned examples have in common is that they rely on some kind of lexical relationship. As most of the required lexical information cannot be obtained from existing linguistic resources such as WordNet or ConceptNet, automating plausible reasoning requires us to learn knowledge about the meaning of words from data. The two most popular methods for learning such knowledge are to induce geometric representations from co-occurrence data [9; 42; 26; 20] and to use information extraction techniques to obtain lexical relations from natural language statements [19; 14; 5]. As we will see in Section 2, each of the lexical relations that are needed in the aforementioned examples can be modelled as spatial relations in a high-dimensional Euclidean space. Subsequently, in Section 3, we explain how this view enables us to unify information derived from geometric models with information derived from relation extraction methods, along with any lexical relations that are available in existing linguistic resources. Integrating and reasoning about lexical information from different sources then becomes a spatial reasoning problem, which we discuss in more detail in Section 4.

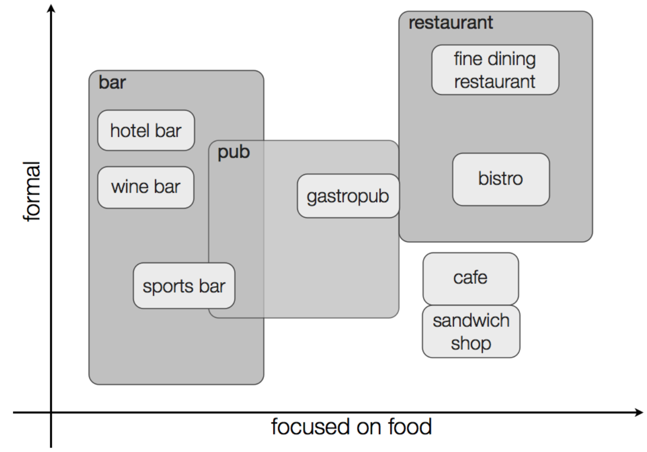

Figure 1: Two-dimensional semantic space of establishments.

2 Lexical relations in semantic spaces

The meaning of words is difficult to capture in logical terms. Wittgenstein, for example, observed that 'game' cannot be defined in terms of necessary and sufficient conditions. Rather, our understanding of this concept is based on similarity, where an activity is considered to be a game if it is sufficiently similar to other activities that we think of as games. This view has given rise to a range of cognitive models that define concepts in terms of distance [29]. Such models are usually implemented by interpreting concepts as regions in a high-dimensional Euclidean space, which is referred to as a semantic space or conceptual space [15].

Figure 1 illustrates the main ideas. Note, however, that while this toy example is based on a two-dimensional space, the semantic spaces used in applications are high-dimensional, often using 100 to 300 dimensions. In the semantic space from Figure 1, categories of establishments (e.g. 'bar', 'wine bar', 'cafe') correspond to regions, while specific establishments (e.g. a particular cafe) would correspond to points (not shown). Crucially, various types of lexical relations correspond to spatial relations in this space. For example, the fact that bar and restaurant are disjoint regions corresponds to the assumption that no establishment can be a bar and a restaurant at the same time. Containment corresponds to subsumption, e.g. the region for 'bistro' is included in the region for 'restaurant' because it is assumed that all bistros are restaurants. Pub and restaurant are adjacent, which corresponds to the view that while they are considered as disjoint concepts, 'restaurant' and 'pub' may share some borderline cases. Adjacent concepts are sometimes also referred to as conceptual neighbours. We could refine how borderline cases are modelled by representing categories as fuzzy sets, in which case regions consist of a core (modelling normal instances of the concept) surrounded by a gradual boundary (modelling borderline cases). The fact that 'pub' is between 'bar' and 'restaurant' corresponds to the assumption that pubs are conceptually between bars and restaurants. This interpretation of geometric betweenness relies on the assumption that natural properties correspond to convex regions, which has been advocated in [15]. Indeed, if a region B is between regions A and C (i.e. included in the convex hull of $A \cup C$), it follows that every convex regions which contains both A and C also contains B, from which we find that all natural properties shared by A and C apply to B as well. Finally, the salient features of the domain correspond to directions in the semantic space. In this example, the more an establishment is located towards the right, to more it is considered to be focused on food, while the more it is located towards the top, the more formal it is considered to be.

Given a suitable semantic space model, we thus have access to all the lexical information that is needed for the examples from the introduction. Unfortunately, learning detailed semantic space models from data is hard, among others because the specification of most regions would require us to select an exponential number of vertices (in the number of dimensions), even if only convex polytopes are considered. Instead of using regions, most methods therefore represent words as points or vectors. Popular techniques for learning vector representations, based on the analysis of word co-occurrence in large text corpora, include the use of matrix factorization [9; 42], neural network embeddings [26] and probabilistic topic modelling [20]. Alternatively, if similarity judgments are available (e.g. provided by domain experts), multi-dimensional scaling can be used to obtain a geometric representation, as is common in cognitive studies.

A few authors have looked at how approximate region boundaries could be induced from point or vector representations. For example, [16] proposes to use the Voronoi tessellation induced by the point representations of a set of concepts, thus implicitly assuming that the concepts are jointly exhaustive and pairwise disjoint. A similar approach has been proposed in [13], which models words as probability distributions, where probability degrees are defined in function of the distance to the point representation of the word. Another possibility is to first learn point representations for several exemplars of a concept, and to derive a region representation from the resulting set of points. For example, [10] learns region representations by considering the convex hull of the representation of exemplars, after removing outliers. A similar approach has been considered in [13], where a k-nearest neighbour classifier is used to define the extent of the regions. Interestingly, similar approaches have been considered in the area of geographic information systems, for learning boundaries for vernacular regions from the web; e.g. [1] uses Voronoi diagrams for this purpose, while [32] uses a k-nearest neighbours based method and [7] uses one-class support vector machines.

An additional challenge when learning semantic spaces from data is that the resulting dimensions correspond to latent features which are often not interpretable. This can be addressed by identifying directions in the space which do correspond to interpretable features. An unsupervised method to find interpretable directions in a learned semantic space has been proposed in [11]. This approach can be used to describe how one entity or concept differs from another, to rank entities according to how much they have a given property, or to identify the most salient properties of the considered domain.

While learning semantic spaces from data has proven very useful for measuring similarity and analogical relationships between words [40; 41; 26], this approach also has a number of important limitations. First, models of similarity which are based on co-occurrence tend to assign high similarity degrees to both synonyms and antonyms [23], whereas in a cognitively plausible semantic space representation, antonyms should not be close to each other. Second, this method is not suitable for words which are very rare, as the co-occurrence vectors for such words are usually too noisy. Finally, the aforementioned methods for estimating region boundaries based on point representations are not suitable for determining whether two similar concepts are disconnected, adjacent, or overlapping, and if they are adjacent, whether the boundary between them is crisp or gradual.

3 Extracting lexical relations from text

While learning accurate semantic space representations is challenging, what mostly matters in applications are the underlying qualitative spatial relationships, e.g. the fact that one region is between two others. To obtain reliable qualitative spatial relations, semantic space models can be complemented with evidence from natural language. For example, the following sentence suggests that 'original art' and 'poster' are disjoint concepts:[4]

> Some, of what we show to you, is original art, some is a poster, and some is a print.

The next sentence suggests that 'limited edition art print' is a kind of 'poster':[5]

> Limited Edition Art Print: This eye catching poster advertising the female rock band, Misdemeanor is the wonderful work of Frank Kozik.

Finally, the following sentence provides evidence that 'limited edition art print' is a kind of 'original art':[6]

> Art Online is where you can purchase affordable and original art, including oil paintings, limited edition art prints and lots more types of stunning artworks.

We can reconcile these statements by considering that 'original art' and 'poster' are disjoint concepts, which both have 'limited edition art print' as a borderline case. In other words, the regions representing 'original art' and 'poster' would have

[4] `http://erikthevermilion.com/art-works-for-sale-online.htm`, accessed 21 May 2015.

[5] `http://www.popartuk.com/art/frank-kozik/misdemeanor-koz29-limited-edition-print.asp`, accessed 20 May 2015.

[6] `http://art-online.co`, accessed 20 May 2015.

disjoint cores, but overlapping gradual boundaries, covering the region representing ‘limited edition art print’. For an automated system to arrive at this conclusion, it would first need to identify the spatial constraints that are implied by the three sentences. This can be accomplished by using existing methods for relation extraction, for example based on lexico-syntactic patterns [19] such as “[NP], including [NP], [NP] and [NP]”. Then, it would need the ability to reason about these spatial constraints, in order to recognise that they are in conflict; we will discuss this spatial reasoning problem in the next section. Finally, the initial spatial constraints need to be weakened to resolve any inconsistencies, e.g. by taking into account that regions may have a gradual boundary and by discarding unreliable pieces of information. While this approach has not yet been evaluated for semantic space representations, a similar strategy was shown in [37] to be useful for learning which spatial relations hold between vague geographic regions.

Semantic space representations and relation extraction methods both have intrinsic limitations, which can be alleviated by combining both approaches. On the one hand, relations extracted from natural language can help us refine a semantic space model, e.g. by identifying whether the regions representing two given words should overlap or be adjacent, or by providing information about rare words for which (reliable) semantic space representations may not be available. On the other hand, relation extraction methods tend to make systematic mistakes. A semantic space model can be used to highlight extracted relations which are implausible, e.g. if a relation extraction method finds that two words are overlapping, while their semantic space representations are far apart; see [37] for an application of this strategy for geographic regions. Note that in addition to semantic space models and relation extraction methods, subsumption/containment and disjointness relations can also be obtained from existing resources such as WordNet, ConceptNet and CYC. The use of qualitative spatial representations means that all these sources of information can be seamlessly integrated.

In addition to subsumption and disjointness relations, it is also possible to extract fine-grained semantic relations from natural language statements. As an example, consider the following sentences:

> Economy class is cheaper than Business Class but at what cost to the employee?[7]
>
> Business class is cheaper than first class, but they are definitely worth the fare. [8]

From sentences like these, we can find out that ‘cheaper than’ is a salient property of travel modes, and accordingly that there should be a direction in a semantic space of travel modes that corresponds to this property. From the given sentences, we can furthermore derive that the representations of the terms ‘economy class’, ‘business class’ and ‘first class’ would occur in that order along this direction. Suitable phrases describing the salient properties of the domain can be found using systems for open-domain relation extraction [14]. As we will see in the next section, however, the resulting spatial reasoning problem is less straightforward than for containment and overlap relations.

Finally, it is also possible to obtain betweenness relations from text. Such relations can be obtained implicitly, based on the orderings that were found for the salient properties of the domain. For example, if we find that ‘economy class’, ‘business class’ and ‘first class’ occur in this or the reversed order for all known salient properties, we may conclude that ‘business class’ is likely to be conceptually between ‘economy class’ and ‘first class’. Betweenness relations can also be obtained explicitly, e.g. using lexico-syntactic patterns, similarly to how containment and overlap relations are found:

> Mediterraneo bar is something between a bar and a club.[9]
>
> A raga is something between a scale and a composition, it is richer then a scale, but not as fixed as a composition.[10]
>
> Brunch is a combination of breakfast and lunch eaten usually during the late morning ...[11]
>
> A phablet is a cross between a smart phone and a tablet, it’s bigger than a phone but smaller than a tablet. [12]

Each of these sentences expresses a betweenness relation which could be extracted using patterns such as “[NP] is something between [NP] and [NP]”.

4 Spatial reasoning about lexical relations

As explained in the previous section, extracting lexical relations from text leads to a spatial reasoning problem. While qualitative spatial reasoning is a well-studied research area, there are key differences between our setting and the problems which are usually considered in this area. First, while normally 2- or 3-dimensional spaces are considered, in our context we need to reason about high-dimensional spaces. Second, in the context of semantic spaces it is often assumed that all regions are convex, whereas most existing methods for reasoning about spatial constraints only require that regions are regularly closed (i.e. that they are equal to the closure of their interior). However, in [35] it has been shown that imposing convexity does not affect the consistency of RCC8 constraints [6], provided that the number of dimensions is sufficiently high (relative to the number of regions). Third, the direction and betweenness relations that we need for lexical reasoning have not yet received much attention.

Figure 2 illustrates how betweenness relations can be formalised in the case of region based representations. We say

[7] `https://companytravel.wordpress.com/category/business-travel/`, accessed 20 May 2015.

[8] `http://ezinearticles.com/?What-You-Need-To-Know-Flying-Business-Class&id=8963765`, accessed 20 May 2015.

[9] `http://www.sitges-tourist-guide.com/en/bars/sitges-bars.html`, accessed 22 May 2015.

[10] `http://www.jazzguitar.be/exotic_guitar_scales.html`, accessed 22 May 2015.

[11] `http://en.wikipedia.org/wiki/Brunch`, accessed 22 May 2015.

[12] `https://www.tigermobiles.com/2014/07/bad-eyesight-smartphones/`, accessed 22 May 2015.

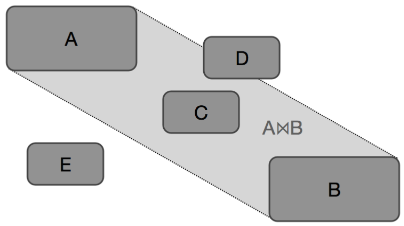

Figure 2: Modelling betweenness for regions.

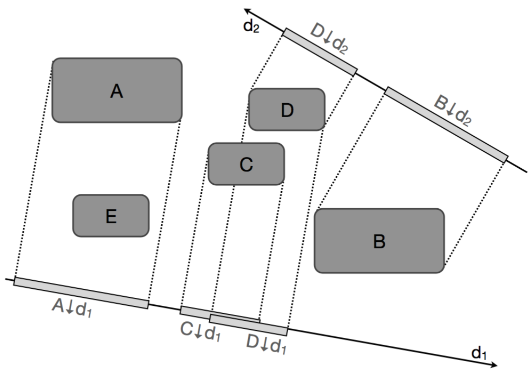

Figure 3: Modelling directions for regions.

that C is fully between A and B, in this example, because C is included in the convex hull of A and B, which we denote by $A \bowtie B$. Similarly we say that D is partially between A and B because D partially overlaps with $A \bowtie B$, and that E is not at all between A and B because E is disjoint from $A \bowtie B$. In other words, we can express betweenness constraints by using the standard topological relations from the region connection calculus in combination with the binary convex hull operator $\bowtie$. It was show in [34] that checking the consistency of such betweenness constraints is tractable when only basic (i.e. non-disjunctive) RCC5 relations are used. Formalising betweenness relations for points is straightforward: a point b is between points a and c iff $\overrightarrow{ab} = \lambda \cdot \overrightarrow{ac}$ for some $\lambda \in]0, 1[$. Checking the consistency of a set of such betweenness constraints was shown in [8] to be $\exists\mathbb{R}$-complete[13].

Figure 3 shows how direction relations between regions can be formalised. Given an oriented line d_1, we can compare regions based on their orthogonal projection on d_1. Since these orthogonal projections correspond to intervals (for self-connected regions), direction relations can be expressed using Allen's interval relations [2]. If all directions are assumed to be orthogonal, then Allen's interval calculus can be used to check consistency (in polynomial time for non-disjunctive relations). However, if directions are not necessarily orthogonal, and in particular if the number of directions is higher than the number of dimensions, this would not be sufficient. Note that in our context both the representation of the directions and the regions are unknown, in contrast to existing calculi such as the rectangle algebra [3], where the directions are given. Currently, no methods are available for checking consistency in this general case. Checking the consistency of direction relations for point representations was studied in [33], where this problem was shown to be $\exists\mathbb{R}$-complete.

A final challenge for lexical reasoning is that it requires us to jointly reason about topological relations (e.g. containment, overlap and adjacency), betweenness and direction. For example, we want to derive that the following statements do not correspond to a consistent set of spatial constraints:

> 'art print' is between 'original art' and 'art poster'
> 'art poster' is a kind of 'poster'
> 'art print' is cheaper than 'poster'
> 'poster' is cheaper than 'original art'

The approach from [34] can be used to reason about betweenness, containment and overlap, but it cannot handle direction relations or adjacency. One important area for future work is thus to extend this approach to include directions and RCC8 relations. In the case of point representations, even reasoning about betweenness or direction relations alone is already $\exists\mathbb{R}$ hard, which means that there is little hope for exact, symbolic inference methods. However, efficient heuristic methods can be designed which generate geometric configurations that as much as possible satisfy a given set of constraint; we refer to [33] for a preliminary evaluation of this approach.

5 Conclusions

In applications that rely on information from the web, plausible reasoning is often necessary to fill in gaps in the knowledge or to make sense of inconsistencies. This typically requires access to fine-grained lexical relations, most of which are not available from existing linguistic resources. This paper has advocated the idea that such lexical relations correspond to qualitative spatial relations in an underlying semantic space. In this way, we are able to combine lexical information derived from learned semantic spaces with relations extracted from natural language, and with whatever lexical information is available from other sources. We highlighted that this leads to a number of challenges for the field of qualitative spatial reasoning, as the consistency problems that arise in this context differ in several ways from those that are encountered in other areas of artificial intelligence.

Acknowledgment

The author has been supported by an ERC Starting Grant (637277).

References

[1] H. Alani, C. B. Jones, and D. Tudhope. Voronoi-based region approximation for geographical information retrieval with gazetteers. *International Journal of Geographical Information Science*, 15:287–306, 2001.

[2] J. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.

[3] P. Balbiani, J.-F. Condotta, and L. F. Del Cerro. A model for reasoning about bidimensional temporal relations. In *Proc. KR*, pages 124–130, 1998.

[13] Recall that $\exists\mathbb{R}$ is the class of decision problems which are as hard as checking the consistency of set of algebraic equations over the reals [31].

[4] I. Beltagy, C. Chau, G. Boleda, D. Garrette, K. Erk, and R. Mooney. Montague meets Markov: Deep semantics with probabilistic logical form. In *Proceedings of the 2nd Joint Conference on Lexical and Computational Semantics*, pages 11–21, 2013.

[5] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. R. Hruschka Jr., and T. M. Mitchell. Toward an architecture for never-ending language learning. In *Proc. AAAI*, pages 1306–1313, 2010.

[6] A. Cohn, B. Bennett, J. Gooday, and N. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica*, 1(3):275–316, 1997.

[7] E. Cunha and B. Martins. Using one-class classifiers and multiple kernel learning for defining imprecise geographic regions. *International Journal of Geographical Information Science*, 28:2220–2241, 2014.

[8] E. Davis, N. Gotts, and A. Cohn. Constraint networks of topological relations and convexity. *Constraints*, 4:241–280, 1999.

[9] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407, 1990.

[10] J. Derrac and S. Schockaert. Enriching taxonomies of place types using Flickr. In *Proc. of the 8th International Symposium on Foundations of Information and Knowledge Systems*, pages 174–192. 2014.

[11] J. Derrac and S. Schockaert. Inducing semantic relations from conceptual spaces: a data-driven approach to plausible reasoning. *Artificial Intelligence*, to appear.

[12] D. Dubois and H. Prade. Fuzzy sets in approximate reasoning, part 1: Inference with possibility distributions. *Fuzzy Sets and Systems*, 100:73 – 132, 1999.

[13] K. Erk. Representing words as regions in vector space. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, pages 57–65, 2009.

[14] O. Etzioni, M. Banko, S. Soderland, and D. S. Weld. Open information extraction from the web. *Communications of the ACM*, 51:68–74, 2008.

[15] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. MIT Press, 2000.

[16] P. Gärdenfors and M. Williams. Reasoning about categories in conceptual spaces. In *Proc. IJCAI*, pages 385–392, 2001.

[17] M. Gardner, P. Talukdar, J. Krishnamurthy, and T. Mitchell. Incorporating vector space similarity in random walk inference over knowledge bases. In *Proc. EMNLP*, pages 397–406, 2014.

[18] J. Y. Halpern. An analysis of first-order logics of probability. *Artificial intelligence*, 46(3):311–350, 1990.

[19] M. A. Hearst. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 14th Conference on Computational Linguistics*, pages 539–545, 1992.

[20] T. Hofmann. Probabilistic latent semantic analysis. In *Proc. UAI*, pages 289–296, 1999.

[21] S. Kok and P. Domingos. Statistical predicate invention. In *Proc. ICML*, pages 433–440, 2007.

[22] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.

[23] D. Lin, S. Zhao, L. Qin, and M. Zhou. Identifying synonyms among distributionally similar words. In *Proc. IJCAI*, pages 1492–1493, 2003.

[24] B. MacCartney and C. D. Manning. An extended model of natural logic. In *Proceedings of the 8th International Conference on Computational Semantics*, pages 140–156, 2009.

[25] J. McCarthy. Circumscription — a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27 – 39, 1980.

[26] T. Mikolov, W.-t. Yih, and G. Zweig. Linguistic regularities in continuous space word representations. In *Proc. NAACL-HLT*, pages 746–751, 2013.

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

[28] M. Richardson and P. Domingos. Markov logic networks. *Machine Learning*, 62(1-2):107–136, 2006.

[29] E. H. Rosch. Natural categories. *Cognitive Psychology*, 4(3):328–350, 1973.

[30] E. Ruspini. On the semantics of fuzzy logic. *International Journal of Approximate Reasoning*, 5:45–88, 1991.

[31] M. Schaefer. Complexity of some geometric and topological problems. In *Graph Drawing*, volume 5849 of *Lecture Notes in Computer Science*, pages 334–344. 2010.

[32] S. Schockaert and M. De Cock. Neighborhood restrictions in geographic IR. In *Proc. SIGIR*, pages 167–174, 2007.

[33] S. Schockaert and J. H. Lee. Qualitative reasoning about directions in semantic spaces. In *Proc. IJCAI*, 2015.

[34] S. Schockaert and S. Li. Combining RCC5 relations with betweenness information. In *Proc. IJCAI*, pages 1083–1089, 2013.

[35] S. Schockaert and S. Li. Realizing RCC8 networks using convex regions. *Artificial Intelligence*, 218:74–105, 2015.

[36] S. Schockaert and H. Prade. Solving conflicts in information merging by a flexible interpretation of atomic propositions. *Artificial Intelligence*, 175(11):1815 – 1855, 2011.

[37] S. Schockaert, P. D. Smart, A. I. Abdelmoty, and C. B. Jones. Mining topological relations from the web. In *Proceedings of the 19th International Workshop on Database and Expert Systems Applications*, pages 652–656, 2008.

[38] R. Sun. Robust reasoning: integrating rule-based and similarity-based reasoning. *Artificial Intelligence*, 75(2):241–295, 1995.

[39] J. B. Tenenbaum, T. L. Griffiths, and C. Kemp. Theory-based bayesian models of inductive learning and reasoning. *Trends in cognitive sciences*, 10(7):309–318, 2006.

[40] P. D. Turney. Measuring semantic similarity by latent relational analysis. In *Proc. IJCAI*, pages 1136–1141, 2005.

[41] P. D. Turney. Domain and function: A dual-space model of semantic relations and compositions. *Journal of Artificial Intelligence Research*, 44:533–585, 2012.

[42] P. D. Turney and P. Pantel. From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37:141–188, 2010.

[43] L. Zadeh. A theory of approximate reasoning. *Machine Intelligence*, 9:149–194, 1979.