

Plausible Reasoning Based on Qualitative Entity Embeddings

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

Shoaib Jameel
 Cardiff University, UK
 JameelS1@Cardiff.ac.uk

Abstract

Formalizing and automating aspects of human plausible reasoning is an important challenge for the field of artificial intelligence. Practical advances, however, are hampered by the fact that most forms of plausible reasoning rely on background knowledge that is often not available in a structured form. In this paper, we first discuss how an important class of background knowledge can be induced from vector space representations that have been learned from (mostly) unstructured data. Subsequently, we advocate the use of qualitative abstractions of these vector spaces, as they are easier to obtain and manipulate, among others, while still supporting various forms of plausible reasoning.

1 Introduction

Many applications require forms of plausible inference, i.e. drawing conclusions that cannot be obtained using standard logical deduction but are nonetheless plausible, typically by applying some kind of implicit background knowledge. Broadly speaking, there are at least two types of background knowledge that can be applied to arrive at such plausible conclusions. First, we can use *statistical knowledge*, e.g. only knowing that Tweety is a bird, we can plausibly derive that Tweety can fly, as we know that this is the case for most birds. Note that our available statistical background knowledge could be encoded numerically (e.g. as a Bayesian network) or qualitatively (e.g. as a set of rules in some default logic). Second, we may derive plausible conclusions based on *semantic knowledge*, i.e. background knowledge about how the entities and categories of the given domain are conceptually related. An example of this kind of reasoning is studied in the psychology literature as category-based induction [Osherson *et al.*, 1990], i.e. inducing properties about categories of objects based on known properties of particular instances or subcategories. For instance, knowing that (i) BBC is regulated by Ofcom and (ii) ITV is regulated by Ofcom, we can plausibly derive that all British broadcasters are regulated by Ofcom. Plausible reasoning based on semantic knowledge is also often needed to deal with conflicts between different knowledge bases in a natural way (see Section 2).

The question we focus on in this paper is how we can acquire the semantic background knowledge that is needed to support such forms of plausible reasoning, and how much of this knowledge can be encoded qualitatively. One common approach is to rely on existing open-domain taxonomies such as WordNet. Either we can use such taxonomies qualitatively (e.g. if ITV and BBC are encoded by the taxonomy to be British broadcasters, then statements which are true for both ITV and BBC might be defeasibly generalized to all British broadcasters) or we can use them to estimate a numerical similarity score between different concepts [Resnik, 1999] (e.g. if *bistro* is found to be taxonomically close to *restaurant* then we might plausibly derive that properties that hold for restaurants also hold for bistros).

For many types of inferences, however, taxonomies are too limiting. A key problem is that often there are many ways in which a given group of entities can be partitioned, which means that either arbitrary choices need to be made or that taxonomies end up being too shallow to support meaningful inductive inferences. For example, depending on the intended application, it might be preferable for *beach* to be taxonomically closer to *harbor* than to *desert* (beach and harbour being coastal features), whereas other applications might require the opposite (beach and desert consisting of sand). In the context of plausible reasoning, a better alternative is to explicitly capture the various ways in which entities and concepts can be grouped. For example, [Kok and Domingos, 2007] proposes a method to learn such multi-clusterings that are maximally predictive. A second important limitation of taxonomies as a basis for plausible reasoning is that inductive inferences commonly use non-taxonomic relations, especially when made by domain experts [Coley *et al.*, 2005]. For example, when we only know about some university that its staff are not permitted to travel in business class, we can plausibly derive that its staff will not be allowed to travel in first class either. Note in particular that we would not derive such a conclusion for economy class, despite the fact that economy class and first class are taxonomically equally close to business class. In such examples, conclusions depend on the underlying causal relations (e.g. staff cannot travel in business class because it is too expensive). To automate such forms of plausible reasoning, we typically need to rely on non-taxonomic relations (e.g. first class is *more expensive than* business class).

In the next section, we look more closely at some examples of applications that require plausible reasoning. Section 3 then discusses the use of vector space embeddings as a basis for plausible reasoning. Finally, in Section 4, we argue that qualitative abstractions of vector space embeddings offer several advantages over numerical representations.

2 Motivating examples

Merging logical theories Different ontologies often contain conflicting information. For example, SUMO¹ considers that *rugby ball* is separate from *ball*, reserving the latter concept for objects that have a spherical shape. Wikidata², on the other hand, considers *rugby ball* to be a subclass of *football*, which is in turn considered to be a subclass of *ball*. Finally, OpenCYC³ considers *rugby ball* and *football* to be separate subtypes of *ball*.

Note that, despite the conflict, the information coming from these three ontologies is not erroneous. A natural way to obtain a unified and consistent logical theory, in such a case, is to explicitly model the fact that concepts such as *football* and *ball* can be used with slightly different meanings. An approach for merging conflicting (propositional) theories based on this view was proposed in [Schockaert and Prade, 2011b]. It considers that the precise meaning of a concept is always source-specific, but that we can make assumptions about how these source-specific interpretations are related. Specifically, it starts from the default assumption that all sources assign the same meaning to all concepts, and only weakens this assumption to the extent needed to restore consistency. To specify how the initial assumption can be weakened, it relies on extra-logical background information which can be encoded using a directed graph. For example, Figure 1 encodes how different interpretations of the concept *football* might be related. Because *football* and *soccer ball* are connected, we allow for the possibility that what one source calls *football* might include what another source calls *soccer ball* or *American football*. If two rounds of weakening are needed to restore consistency, we also allow for the possibility that the other source’s interpretation of *football* includes *rugby ball*. Following [Freksa, 1991], concepts such as *football* and *soccer ball*, which are connected in the diagram, are called conceptual neighbors.

Conflicts between logical theories can also be resolved using taxonomies. For example, if one source claims that Mary studied computer science in Cardiff, while another source claims that she only studied in Oxford, the inconsistency can be resolved by weakening both claims as “Mary studied computer science in the UK”. Note that in this case, the required semantic knowledge can be encoded as part of the logical theory itself. Conceptual neighborhood relations, on the other hand, are essentially extra-logical. The advantage of the conceptual neighborhood approach, however, is that it allows us to weaken the initial assertions in a more targeted way, whereas taxonomy-based weakening is often too cautious.

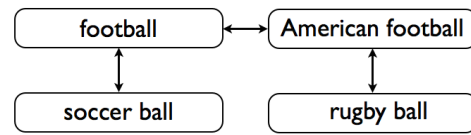


Figure 1: Conceptual neighborhood graph for the concept *football*.

Knowledge base completion Most knowledge bases are incomplete. For example, SUMO encodes that *Piano*, *Violin* and *Guitar* are subclasses of *StringInstrument*, but does not include any facts about *Harpichord*. Even if we did not know anything about the concept *StringInstrument*, we could plausibly derive that *Harpichord* is a subclass of *StringInstrument*. The intuitive reason is because harpsichord is conceptually between piano, violin and guitar, in the sense that properties which tend to be true for pianos, violins and guitars also tend to be true for harpsichords. A method for automatically completing rule bases using this form of inference, which is called interpolation, was introduced in [Schockaert and Prade, 2013] (where a related form of inference, called extrapolation, was also discussed).

Note that taxonomy based approaches cannot be used here, as this example is about dealing with a missing taxonomic relationship. Intuitively, however, we can think of interpolation as a generalization of taxonomy-based induction to multi-clusterings, i.e. the fact that harpsichord is conceptually between piano, violin and guitar could be interpreted as meaning that harpsichord belongs to all natural clusters that also contain piano, violin and guitar.

The aforementioned approaches are all qualitative. Numerical approaches can be used for this task as well, typically by using a form of similarity based reasoning. For example, [Beltagy *et al.*, 2013] presents an approach that uses Markov logic to encode the soft constraint that if a given property holds for a given concept, then it should also hold for similar concepts. Another example of similarity based reasoning is [Raina *et al.*, 2005], which proposed a method for deciding whether one natural language statement entails another. This is formalized as an abductive reasoning problem, where certain assumptions are allowed to be made to prove the entailment (at a cost). The considered assumptions include replacing one concept for a similar one. A key problem of similarity based reasoning methods, however, is the difficulty of linking similarity degrees to probabilities in a principled way.

3 Entity embeddings

The applications discussed in the previous section require semantic background knowledge that is not contained in existing resources such as WordNet, CYC, ConceptNet or Freebase. This includes information about conceptual neighborhood, betweenness, and relative features (e.g. a relation such as “more expensive than” between travel modes). To learn such semantic relations from data, [Schockaert and Prade, 2011a] proposed to induce a vector space by applying multi-dimensional scaling (MDS) to a bag-of-words representation of each entity. In the resulting representation, entities cor-

¹<https://github.com/ontologyportal/sumo/blob/master/Sports.kif>

²<https://www.wikidata.org/wiki/Q2881344>

³<http://sw.opencyc.org/2012/05/10/concept/en/Ball>

respond to points while concepts and properties correspond to convex regions (estimated based on the convex hull of their instances), in accordance with the theory of conceptual spaces [Gärdenfors, 2000]. It was found that geometric betweenness in this vector space is useful for identifying conceptual betweenness (e.g. the region for *Harpsichord*, in a space of musical instruments, should be approximately in the convex hull of the regions for *Piano*, *Violin* and *Guitar*). Furthermore, parallel directions in the space were found to intuitively correspond to analogies. This initial approach was significantly improved in [Derrac and Schockaert, 2015], where it was shown how directions could be identified in the space which model relative features (e.g. “more violent than” in a space of movies). Specifically, these directions allow us to rank the entities of that space according to how much they have the corresponding feature. Recently, [Jameel and Schockaert, 2016] further improved on this method by learning a 300-dimensional open-domain vector space representation for 1.2M entities, in which all instances of a given semantic type are located in a particular lower-dimensional subspace. This is illustrated in Figure 2, which shows the two-dimensional subspace that was found for the entities of type *Civilization* (which is itself embedded in the 300-dimensional space). In this way, semantic relations between entities of different semantic types can be taken into account to improve the representation (e.g. knowing that films A and B have similar directors provides evidence that these films are similar), while the use of subspaces ensures that we can still identify domain-specific relations. For example, in Figure 2, the direction from left to right approximately models the relation “more recent than”, while the upwards direction is related to how populous/influential different civilizations were.

Recently, a number of popular models have been proposed for automatically completing knowledge graphs, i.e. sets of triples of the form (e, r, f) , encoding that entity e is in relation r with entity f . Most of these approaches are based on the view that relations can be modelled as translations in a vector space [Bordes *et al.*, 2013]. For example, to model that Paris is the capital of France, the vector space representation can be constrained such that $p_{france} + r_{capital\ of} \approx p_{paris}$. While these vector space embeddings share some similarities with the aforementioned entity embeddings, there are a number of key differences. For example, knowledge graph embedding models are supervised methods for learning correlations between a given set of relations (often between entities of different types), whereas the aforementioned entity embedding models are used to induce semantic relations between entities of the same type in an unsupervised way (i.e. the salient relations for each domain are identified automatically). Moreover, knowledge graph embedding models primarily capture statistical knowledge, rather than semantic knowledge. Consider for example the following default rule: if “person a works for company b ” and “company b is located in city c ” then typically “person a lives in city c ”. While not deductively valid, this rule may be satisfied for many triples (a, b, c) in a given knowledge graph. In the resulting embedding we then expect $r_{works\ for} + r_{located\ in} \approx r_{lives\ in}$. In principle, a single vector space cannot capture such statistical correlations and at the same time faithfully model semantic relations between

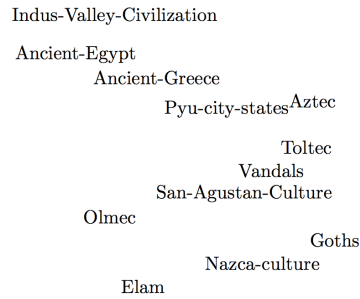


Figure 2: Subspace with entities of type *Civilization*.

the entities. For example, it is possible for persons a and b to be semantically similar (e.g. similar age, profession, hobbies, geographic location) while their brothers are not. In the translation model, on the other hand, if a and b are similar then $a + r_{brother\ of}$ and $b + r_{brother\ of}$ are similar as well⁴. On the other hand, there is some evidence that in practice, learning a single vector space that captures semantic relatedness (induced from bag of words representations) as well as correlations from a knowledge graph can actually be beneficial, in the sense that both the modelling of semantic relatedness and the knowledge graph relations might be improved [Zhong *et al.*, 2015; Jameel and Schockaert, 2016].

4 Qualitative entity embeddings

While vector space embeddings have proven remarkably successful in tasks like knowledge base completion, they also have several drawbacks compared to taxonomies. Taxonomies, even when learned from data, are easy to understand and debug by domain experts, and easier to reuse in new applications. Due to their qualitative nature, taxonomies are also easier to extend, e.g. by extracting taxonomic relations directly from text documents [Hearst, 1992]. Note, however, that the applications discussed in Section 2 only rely on qualitative relations that have been derived from the vector space representation, rather than on the vector space representation itself. In this section, we outline how the advantages of taxonomies and vector space representations could be combined by using a qualitative abstraction of entity embeddings.

Given an embedding of the entities of a given domain (e.g. movies), we can use the method from [Derrac and Schockaert, 2015] to identify the directions d_1, \dots, d_n that correspond to the most salient features of that domain (e.g. more scary than). Each of these directions d_i induces a ranking of the entities, which can be encoded as a mapping σ_i from the set of entities to the set of ranks $\{1, \dots, n\}$. Experiments in [Derrac and Schockaert, 2015] have shown that classifiers using only these rankings can be as accurate as classifiers that directly operate on the vector space representation.

Given the rankings $\sigma_1, \dots, \sigma_n$ of all entities, we can represent concepts by defining for each concept C and each ranking σ_i an interval $[l_C^i, u_C^i]$. An entity e is then assumed to

⁴While extensions of this translation model exist that allow for more flexibility, e.g. by considering translations in a subspace [Wang *et al.*, 2014], the principle remains that knowledge graph constraints could distort the way in which semantic relations are modelled.

belong to the concept C if $l_C^i \leq \sigma_i(e) \leq u_C^i$ for every $i \in \{1, \dots, n\}$. Intuitively, this use of intervals is similar to representing concepts geometrically as hyperrectangles, although the directions corresponding to the rankings need not be orthogonal and the number of rankings can be chosen higher than the number of dimensions. Note that it is straightforward to learn suitable intervals $[l_C^i, u_C^i]$ from positive and negative examples of the concept. Furthermore, we can easily generalize this interval representation to capture the vagueness of category boundaries. To estimate whether concept D is between concepts C_1, \dots, C_k we can simply check whether

$$[l_D^i, u_D^i] \subseteq [\min_{j=1}^k l_{C_j}^i, \max_{j=1}^k u_{C_j}^i] \quad (1)$$

for every i . Note that this condition is not sufficient for D to be in the convex hull of C_1, \dots, C_k in the vector space, although we can always avoid false positives by adding additional directions. Another possibility is to augment the rankings with explicit information about how the regions are spatially related, using an appropriate qualitative spatial calculus [Schockaert and Li, 2013]. Furthermore note that in practice we are mostly interested in approximate betweenness, in which case the inclusion in (1) should be replaced by a soft constraint (i.e. l_D^i should be greater than *or close to* $\min_j l_{C_j}^i$, and similar for the upper bound). Checking whether C and D are conceptual neighbors can be done in a similar way.

The rankings $\sigma_1, \dots, \sigma_n$ thus offer a qualitative view of the entity embedding, which can support plausible reasoning without having to rely on computationally expensive geometric computations (e.g. convex hulls in high-dimensional spaces). Furthermore, given that each ranking corresponds to an interpretable, salient property of the domain of interest, this representation is easy to improve by domain experts, crowdsourcing, or by taking into account information extracted from natural language (e.g. using surface patterns such as “[X] is somewhat between [Y] and [Z]”). Finally, the qualitative representation also makes it easier to take into account the context-dependent nature of many semantic relations, by only considering those rankings that correspond to features that are relevant in the given application context.

5 Conclusions

We have proposed the use of qualitative entity embeddings to represent the semantic background knowledge that is needed for various forms of plausible reasoning. This is motivated by the fact that we can thus combine the data-driven nature of vector space embeddings with the transparency and ease-of-use of symbolic representations, such as taxonomies.

Acknowledgments

This work was supported by ERC Starting Grant 637277.

References

[Beltagy *et al.*, 2013] Islam Beltagy, Cuong Chau, Gemma Boleda, Dan Garrette, Katrin Erk, and Raymond Mooney. Montague meets Markov: Deep semantics with probabilistic logical form. In *Proc. *SEM*, pages 11–21, 2013.

[Bordes *et al.*, 2013] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In *Proc. NIPS*, pages 2787–2795, 2013.

[Coley *et al.*, 2005] John D. Coley, Patrick Shafto, Olga Stepanova, and Elizabeth Baraff. Knowledge and category-based induction. In *Categorization Inside and Outside the Laboratory: Essays in Honor of Douglas L. Medin*. American Psychological Association, 2005.

[Derrac and Schockaert, 2015] J. Derrac and S. Schockaert. Inducing semantic relations from conceptual spaces: a data-driven approach to plausible reasoning. *Artificial Intelligence*, pages 74–105, 2015.

[Freksa, 1991] C. Freksa. Conceptual neighborhood and its role in temporal and spatial reasoning. In M. Singh and L. Travé-Massuyès, editors, *Decision Support Systems and Qualitative Reasoning*, pages 181–187. North-Holland, Amsterdam, 1991.

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

[Hearst, 1992] Marti A Hearst. Automatic acquisition of hyponyms from large text corpora. In *Proc. COLING*, pages 539–545, 1992.

[Jameel and Schockaert, 2016] Shoaib Jameel and Steven Schockaert. Entity embeddings with conceptual subspaces as a basis for plausible reasoning. In *arXiv:1602.05765*, 2016.

[Kok and Domingos, 2007] Stanley Kok and Pedro Domingos. Statistical predicate invention. In *Proc. ICML*, pages 433–440, 2007.

[Osherson *et al.*, 1990] Daniel N Osherson, Edward E Smith, Ormond Wilkie, Alejandro Lopez, and Eldar Shafir. Category-based induction. *Psychological review*, 97(2):185–200, 1990.

[Raina *et al.*, 2005] Rajat Raina, Andrew Y. Ng, and Christopher D. Manning. Robust textual inference via learning and abductive reasoning. In *Proc. AAAI*, pages 1099–1105, 2005.

[Resnik, 1999] Philip Resnik. Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, 11:95–130, 1999.

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

[Schockaert and Prade, 2011a] Steven Schockaert and Henri Prade. Interpolation and extrapolation in conceptual spaces: A case study in the music domain. In *Proc. RR*, pages 217–231, 2011.

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

[Schockaert and Prade, 2013] Steven Schockaert and Henri Prade. Interpolative and extrapolative reasoning in propositional theories using qualitative knowledge about conceptual spaces. *Artificial Intelligence*, 202:86–131, 2013.

[Wang *et al.*, 2014] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In *Proc. AAAI*, pages 1112–1119, 2014.

[Zhong *et al.*, 2015] Huaping Zhong, Jianwen Zhang, Zhen Wang, Hai Wan, and Zheng Chen. Aligning knowledge and text embeddings by entity descriptions. In *Proc. EMNLP*, pages 267–272, 2015.