Cones, Negation, and All That (Extended Abstract)

Özgür L. Özçep[1 [0000−0001−7140−2574]], Mena Leemhuis[1 [0000−0003−1017−8921]], and Diedrich Wolter[2 [0000−0001−9185−0147]]

[1] University of Lübeck, Lübeck, Germany
oezcep@ifis.uni-luebeck.de
mena.leemhuis@student.uni-luebeck.de
[2] University of Bamberg, Bamberg, Germany
diedrich.wolter@uni-bamberg.de

Abstract. This paper summarizes results on embedding ontologies expressed in the $\mathcal{ALC}$ description logic into a real-valued vector space, comprising restricted existential and universal quantifiers, as well as concept negation and concept disjunction. The main result states that an $\mathcal{ALC}$ ontology is satisfiable in the classical sense iff it is satisfiable by a partial faithful geometric model based on cones. The line of work to which we contribute aims to integrate knowledge representation techniques and machine learning. The new cone-model of $\mathcal{ALC}$ proposed in this work gives rise to conic optimization techniques for machine learning, extending previous approaches by its ability to model full $\mathcal{ALC}$.
This is an extended abstract of the paper "Cone Semantics for Logics with Negation" to be published in the proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI 2020).

1 Introduction

This extended abstract reports on results related to the general framework of cone-based semantics as developed in [15]. The framework relies on the idea of embedding ontologies into low-dimensional continuous vector spaces. This idea goes back to the idea of embedding words into low-dimensional continuous vector spaces which has been implemented successfully in various algorithms with various applications in the realm of information retrieval [6, 16, 12]. However, these approaches are insensitive to the relational structure of documents. The embedding idea was pushed further (see, e.g., [14, 3] and, for an overview, [17]) in order to design embeddings of knowledge graphs or embeddings of ontologies consisting of axioms in some (expressive) logic [13, 10, 8].

The main aim of our framework is to find embeddings of ontologies that give a better compromise between the geometrical models that can be constructed by means of learning and the (expressivity and consistency) demands of ontologies. *Convex cones* are an ideal data structure for such embeddings, as they combine two desirable properties: On the one hand, computational feasibility is ensured by convexity (see work on convex or conic optimization, e.g., [4] as well as work on conceptual spaces [5]). And on the other hand, sufficient expressivity is ensured by conicity; cones have a well-defined polarity operation that behaves as a negation operator, in fact as an orthonegation operator [7, 9]. Arbitrary convex sets do not provide a well-defined negation operation (convex sets are closed under intersection but not under set-complement or set-union.) The main result of [15] states that an ontology defined over the description logic $\mathcal{ALC}$ [1], which provides full concept negation, is satisfiable in a classical sense iff it is satisfiable by a geometric model that interprets all concept descriptions as axis-aligned cones, for short *al-cones*. And one can even ensure that the embedding is faithful: The cone-based geometric models used in [15] are partial and thus allow some uncertainty to be retained, i.e., if x is only known to be a member of the union of two atomic concepts, then our partial model will not commit to saying to which atomic concept x belongs. A *faithful* partial model will represent exactly those axioms derivable from the ontology.

2 Embedding $\mathcal{ALC}$ Ontologies with Al-Cones

The core of our cone-based semantics evolves around the notion of the polarity operator, which is defined for arbitrary convex cones X, i.e. sets fulfilling: If $v, w \in X$, then also $\lambda v + \mu w \in X$ for all $\lambda, \mu \geq 0$. The *polar cone* X° for X is defined for Euclidean spaces with a scalar (dot) product $\langle \cdot, \cdot \rangle$ as follows:

$$X^\circ = \{v \in \mathbb{R}^n \mid \forall w \in X : \langle v, w \rangle \leq 0\}$$

The use of the polarity operation for concept negation $\neg$ is motivated by the idea of providing an operator that always maps a concept to a disjoint concept such that the disjoint concept is maximally so w.r.t. the underlying similarity structure $\langle \cdot, \cdot \rangle$ (see also Farkas' classical lemma on polarity).

Interpreting set intersection as concept-conjunction $\sqcap$ and using de Morgan's rule to define concept-disjunction $\sqcup$ one already has the main ingredients to interpret arbitrary Boolean $\mathcal{ALC}$ concepts. But, as arbitrary cones do not fulfil the distributivity property of $\mathcal{ALC}$ concepts w.r.t. $\sqcap$ and $\sqcup$ (Fig. 1, lhs), our embeddings are constrained to axis-aligned cones, for short *al-cones*:

$$X \text{ is al-cone} :\Leftrightarrow X = X_1 \times \cdots \times X_n, X_i \in \{\mathbb{R}, \mathbb{R}_+, \mathbb{R}_-, \{0\}\}$$

As a simple example for embedding (Boolean) $\mathcal{ALC}$ ontologies we consider the case of all concept descriptions over two atomic symbols A, B (Fig. 1, rhs). In the al-cone embedding of Figure 1 the A is interpreted by the left upper quadrant and B by the right upper quadrant. This induces uniquely the positions of all other hyperoctants corresponding to the other boolean concepts.

One can check that the concepts are associated with appropriate al-cones. For example, the negation $\neg A$ of A is indeed the polar cone of the quadrant of A. Similarly, consider $B \sqcap \neg A$, which is interpreted as the positive x-axis $\mathbb{R}_+ \times \{0\}$.

The example demonstrates also the partiality of al-cone models. Consider, e.g., the difference between a_2 and a_3 in the geometric model on the rhs of Fig.

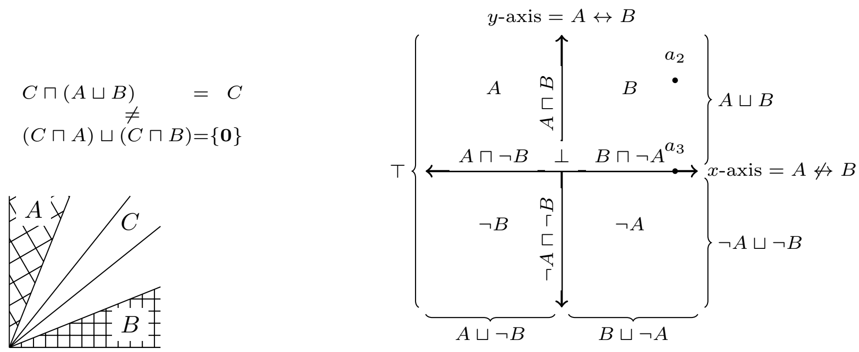

Fig. 1. Counterexample distributivity (lhs) and example al-cone model (rhs)

1. The individual a_3 is completely identified w.r.t. the given concepts A, B: it lies in the extension of B and in the extension of $\neg A$. For a_2 we "only" know that it must be an B, but we do not know whether it is also an A.

Using the construction idea of the example one can prove that a Boolean $\mathcal{ALC}$ ontology is classically satisfiable iff it satisfiable by an al-cone based model. And one can guarentee faithfulness: The geometric model encodes all and only the information of the ontology. Using some more thoughts on how to deal with relations (roles) one can generalize the result to hold for arbitrary $\mathcal{ALC}$ ontologies.

Proposition 1. *$\mathcal{ALC}$ ontologies are classically satisfiable iff they are satisfiable by a faithful geometric model on some $\mathbb{R}^n$ using sets of the form $b_1 \times \cdots \times b_n$ with $b_i \in \{\{0\}, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$.*

This result has important consequences for possible supervised learning algorithms relying on al-cone based geometric models (such as the prototypical multilabelling algorithm described in [11]): If the algorithm is not able to find a model fitting the training data, this is due to a small feature dimension n chosen in the beginning or due to inconsistencies of the ontology. The inconsistency cannot be due to the fact that concepts are represented as al-cones.

3 Conclusion and Outlook

By interpreting negation as a polarity operator it is possible to find embeddings of $\mathcal{ALC}$ ontologies that interpret all concepts as axis-aligned cones. This result adds an interesting alternative to embeddings considered so far.

In [15] we only consider the case where the logic ($\mathcal{ALC}$) has been specified beforehand, not the case of investigating logics induced by the intersection and polarity operators for arbitrary cones. In ongoing work we are investigating non-distributive logics suitable for arbitrary cones. These logics are extensions of so-called orthologics [7]—which describe lattices equipped with an orthonegation. We are able to identify non-trivial rules (weakenings of orthomodularity, a property used for minimal quantum logic [2]) that are fulfilled by cones.

References

1. Baader, F.: Description logic terminology. In: Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.) The Description Logic Handbook, pp. 485–495. Cambridge University Press (2003)
2. Birkhoff, G., von Neumann, J.: The logic of quantum mechanics. Annals of Mathematics **37**(4), 823–843 (1936)
3. Bordes, A., Usunier, N., García-Durán, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data. In: Burges, C.J.C., Bottou, L., Ghahramani, Z., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States. pp. 2787–2795 (2013), http://papers.nips.cc/paper/5071-translating-embeddings-for-modeling-multi-relational-data
4. Boyd, S., Vandenberghe, L.: Convex optimization. Cambridge university press (2004)
5. Gärdenfors, P.: Conceptual Spaces: The Geometry of Thought. The MIT Press, Cambridge, Massachusetts (2000)
6. Goldberg, Y., Levy, O.: word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. ArXiv e-prints (Feb 2014)
7. Goldblatt, R.I.: Semantic analysis of orthologic. Journal of Philosophical Logic **3**(1), 19–35 (1974). https://doi.org/10.1007/BF00652069, https://doi.org/10.1007/BF00652069
8. Gutiérrez-Basulto, V., Schockaert, S.: From knowledge graph embedding to ontology embedding? An analysis of the compatibility between vector space representations and rules. In: Thielscher, M., Toni, F., Wolter, F. (eds.) Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October – 2 November 2018. pp. 379–388. AAAI Press (2018), https://aaai.org/ocs/index.php/KR/KR18/paper/view/18013
9. Hartonas, C.: Reasoning with incomplete information in generalized galois logics without distribution: The case of negation and modal operators. In: Bimbó, K. (ed.) J. Michael Dunn on Information Based Logics, pp. 279–312. Springer International Publishing, Cham (2016)
10. Kulmanov, M., Liu-Wei, W., Yan, Y., Hoehndorf, R.: El embeddings: Geometric construction of models for the description logic EL++. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19) (2019)
11. Leemhuis, M., Özçep, O.L., Wolter, D.: Multi-label learning with a cone-based geometric model. In: Proceedings of the 25th International Conference on Conceptual Structures (ICCS 2020) (2020), (To appear)
12. Levy, O., Goldberg, Y.: Neural word embedding as implicit matrix factorization. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada. pp. 2177–2185 (2014), http://papers.nips.cc/paper/5477-neural-word-embedding-as-implicit-matrix-factorization
13. Mehran Kazemi, S., Poole, D.: SimplE Embedding for Link Prediction in Knowledge Graphs. arXiv e-prints arXiv:1802.04868 (Feb 2018)

14. Nickel, M., Tresp, V., Kriegel, H.P.: A three-way model for collective learning on multi-relational data. In: Proceedings of the 28th International Conference on International Conference on Machine Learning. pp. 809–816. ICML'11, Omnipress, USA (2011), http://dl.acm.org/citation.cfm?id=3104482.3104584
15. Özçep, Ö.L., Leemhuis, M., Wolter, D.: Cone semantics for logics with negation. In: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20 (2020), (To appear)
16. Pennington, J., Socher, R., Manning, C.D.: Glove: Global vectors for word representation. In: EMNLP. vol. 14, pp. 1532–1543 (2014), https://nlp.stanford.edu/pubs/glove.pdf
17. Wang, Q., Mao, Z., Wang, B., Guo, L.: Knowledge graph embedding: A survey of approaches and applications. IEEE Transactions on Knowledge and Data Engineering **29**(12), 2724–2743 (Dec 2017). https://doi.org/10.1109/TKDE.2017.2754499