EmELvar: A NeuroSymbolic Reasoner for the EL++ Description Logic

Biswesh Mohapatra[1], Sumit Bhatia[2], Raghava Mutharaju[3] and G. Srinivasaraghavan[4]

[1]*Inria Labs, Paris*
[2]*Adobe Research, India*
[3]*Indraprastha Institute of Information Technology, Delhi*
[4]*International Institute of Information Technology, Bangalore*

Abstract

This paper summarises our work on Ontology embeddings for the SemRec Workshop at ISWC 2021. Recent work on Knowledge Graph (KG) embeddings provide a low-dimensional representation of entities and relations of a Knowledge Graph and are used successfully for various applications such as question answering and search, reasoning, inference, and missing link prediction. However, most of the existing KG embeddings only consider the network structure of the graph and ignore the semantics and the characteristics of the underlying ontology that provides crucial information about relationships between entities in the KG. Recent efforts in this direction involve learning embeddings for a Description Logic (logical underpinning for ontologies) named $\mathcal{EL}^{++}$. However, such methods consider all the relations defined in the ontology to be one-to-one which severely limits their performance and applications. We provide a simple and effective solution to overcome this shortcoming that allows such methods to consider many-to-many relationships while learning embedding representations. Our proposed solution also paves the way for learning embedding representations for even more expressive description logics such as $\mathcal{SROIQ}$.

Keywords

NeuroSymbolic AI, Ontologies, EL embeddings

1. Introduction

Knowledge Graph (KG) embeddings are an effort to combine the knowledge present in the Knowledge Graphs with the generalisation capability of the neural networks [1]. KG embeddings learn embedding functions that map the entities of Knowledge Graphs to vector space. Different Learning methods for such embeddings have been proposed [2, 3, 4, 5, 6, 7] that try to preserve various properties of these knowledge bases. In case of ontologies, they help create approximate reasoners which could reason over complex ontologies while having low time complexity.

Most of the KG embeddings fail to consider the underlying constraints and characteristics of the ontologies. Hence, reasoning tasks do not perform well on such embeddings of ontologies. In order to tackle this issue, Kulmanov et al. [8] proposed EL Embeddings (EmEl) that incorporate the geometric structure of $\mathcal{EL}^{++}$ description logic ontologies into the embeddings. Mondal et

SemREC'21: Semantic Reasoning Evaluation Challenge, ISWC'21, Oct 24 – 28, Albany, NY
biswesh.mohapatra@inria.fr (B. Mohapatra); sumit.summit@gmail.com (S. Bhatia); raghava.mutharaju@iiitd.ac.in (R. Mutharaju); gsr@iiitb.ac.in (G. Srinivasaraghavan)

al. [9] later added role oriented $\mathscr{EL}^{++}$ constructs into the embeddings through their proposed EmEl^{++} method. While these methods have provided a new technique to perform reasoning tasks on the ontologies, they have a fundamental issue that restricts their performance on $\mathscr{EL}^{++}$ ontologies and restricts them from being used in more complex description logic based ontologies such as $\mathscr{SROIQ}$, which is the basis for OWL 2 DL, a fragment of OWL 2. We provide a simple and effective way to convert the embedding functions such that the roles (relation equivalent in ontologies) can be considered as many-to-many instead of one-to-one functions as is in the case of EmEl. This is significant as most of the roles in ontologies connect a class to multiple classes. For example, the *fatherOf* role can connect an individual to multiple individuals if he is the father of all of them. This issue becomes more important when we try to move to complex description logics like $\mathscr{SROIQ}$ which has properties such as cardinality that depend on such many-to-many roles.

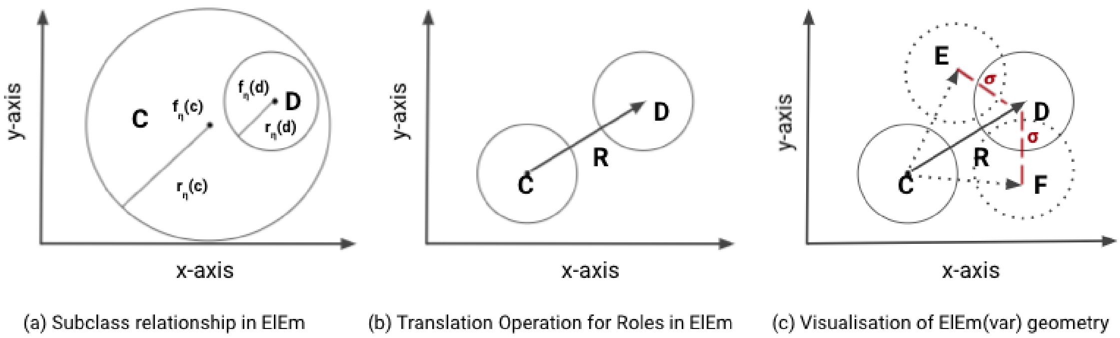

Figure 1: Geometric representation of classes and relations. (a) shows the representation of subclass relation where D ⊑ C and thus the n-ball of D lies inside n-ball of C. (b) shows class C getting translated to D using relation R for a tuple (C,R,D) in the ontology. (c) In case of EmEl(var), the variance σ lets the the entity C relate to multiple entities with the relation R. Any entity which falls within σ distance of C+R are also related to C through R. E and F are the boundary entities for C and R.

Contributions: (1) We provide a simple method to incorporate many-to-many roles in translation based embeddings like TransE [3]. (2) We show how the method could be used to modify the ontological embedding EmEl [8]. (3) We demonstrate the effectiveness of the method on the dataset provided by SemRec. (4) Our work provides a foundation for work on complex DL reasoners.

2. Related Work

Existing works on ontology embedding such as Onto2vec [10] focuses on using word2vec as an underlying model. While the work focuses on encoding the entities and relations, it is unable to handle complex relations in an ontology. [11] provide neuro-symbolic deep deductive reasoners for $\mathscr{EL}^{++}$ DL and first-order logic. [12] pointed out that geometric models are a better way to learn embeddings for ontologies. The simplicity of the translation based models for KG embeddings [3, 2, 4] to measure the correctness of a fact as a distance between entities after being translated by the relation made them popular. EmEl [8] and EmEl^{++} [9] used this translation technique to create embeddings for ontologies which preserve their underlying structures and characteristics. In order to accomplish it, the models use geometric models to learn embeddings. The classes are considered to be n-balls in an n-dimensional space which are translated by the relation vectors to the n-ball of the corresponding class of the fact. This geometric structure provides a way to incorporate various structural properties of an ontology eg. subclass properties. However, like TransE, these models too restrict their triplets to a one-to-one mapping. Not only do these restrictions affect the performance of these models on $\mathscr{E}\mathscr{L}^{++}$ ontologies but also restrict them from being used in more complex description logics such as $\mathscr{S}\mathscr{R}\mathscr{O}\mathscr{I}\mathscr{Q}$.

3. Background on Ontology Embeddings

Kulmanov et al. [8] introduced the concept of incorporating geometric structure of ontologies into the embeddings. They proposed embeddings for the $\mathscr{E}\mathscr{L}^{++}$ description logic (EmEL) that captures the underlying structures and characteristics of the ontology by treating ontology classes as n-balls in n-dimensional space. These n-balls are represented by a center which is a n-dimensional vector and a radius which is a scalar. The relations in the ontology are considered as n-dimensional vectors which are used to translate the class from one point in the space to another. The center and the radius of each class (n-ball), along with the relations can be learnt over multiple iterations. They make up the embeddings for the ontology. Figure 1(a) and Figure 1(b) show the geometric representation of classes and relations in 2-dimensional space.

Hence, they define a geometric ontology embedding η as a pair (f_η, r_η) of functions that map classes and relations in ontology O into R^n. Thus f_η : C ∪ R $\mapsto \mathbb{R}^n$ and r_η : C $\mapsto \mathbb{R}$. Here C is a class and R is a relation and O is defined as $(\mathbb{C}, \mathbb{R}', \mathbb{I}$; ax) where $\mathbb{I}$ are individual symbols, $\mathbb{C}$ is set of class symbols, $\mathbb{R}'$ is set of relation symbols and ax are axioms (facts). Basically, f_η(c) represents center of C, r_η(c) represents radius of C and f_η(r) represents vector of R.

Each axiom, ax, is transformed into its equivalent normal form using a set of conversion rules from [13]. These rules help transform the set of axioms in the ontology into one of four forms without any loss of information. These are (1) Subclass axiom: $C \sqsubseteq D$ (2) Intersection axiom: $C \sqcap D \sqsubseteq E$ (3) Existential restriction (right-hand side): $C \sqsubseteq \exists R.D$ and (4) Existential restriction (left-hand side): $\exists R.C \sqsubseteq D$ where C, D, E ∈ $\mathbb{C}$ and R ∈ $\mathbb{R}$.

EmEl formulates a loss function for each of the four normal forms in order to preserve the semantics of $\mathscr{E}\mathscr{L}^{++}$ in the embeddings. The loss functions are as follows.

$$\begin{aligned} \text{loss}_{C \sqsubseteq D}(c, d) = \max(0, \| f_\eta(c) - f_\eta(d) \| + r_\eta(c) - r_\eta(d) - \gamma) \\ + | \| f_\eta(c) \| - 1| + | \| f_\eta(d) \| - 1| \end{aligned} \tag{1}$$

In Eqn 1, we try to preserve the subclass property of the entities. Here the Euclidean distance between the centers of C and D should be less than the difference between the radius of D and C. Once this is achieved, we ensure that the n-ball representing D is bigger than that of C and that n-ball of C lies completely inside D. Here γ is a hyperparameter called margin. $| \| f_\eta(c) \| - 1| + | \| f_\eta(d) \| - 1|$ ensures that the n-balls lie in the unity sphere.

$$
\begin{aligned}
\text{loss}_{C\sqcap D\sqsubseteq E}(c,d,e) = \max(0, \| f_\eta(c) - f_\eta(d) \| - r_\eta(c) - r_\eta(d) - \gamma) \\
+ \max(0, \| f_\eta(c) - f_\eta(e) \| - r_\eta(c) - \gamma) \\
+ \max(0, \| f_\eta(d) - f_\eta(e) \| - r_\eta(d) - \gamma) \\
+ |\, \| f_\eta(c) \| - 1| + |\, \| f_\eta(d) \| - 1| \, \| f_\eta(e) \| - 1|
\end{aligned} \tag{2}
$$

In Eqn 2, we incorporate the intersection property. The first term ensures that C and D are not disjoint sets. While second and third terms force the center of E to lie in the intersection of D.

$$
\begin{aligned}
\text{loss}_{C\sqsubseteq \exists R.D}(c,d,r) = \max(0, \| f_\eta(c) + f_\eta(r) - f_\eta(d) \| \\
+ r_\eta(c) - r_\eta(d) - \gamma) \\
+ |\, \| f_\eta(c) \| - 1| + |\, \| f_\eta(d) \| - 1|
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
\text{loss}_{\exists R.C\sqsubseteq D}(c,d,r) = \max(0, \| f_\eta(c) - f_\eta(r) - f_\eta(d) \| \\
- r_\eta(c) - r_\eta(d) - \gamma) \\
+ |\, \| f_\eta(c) \| - 1| + |\, \| f_\eta(d) \| - 1|
\end{aligned} \tag{4}
$$

Eqn 3 and Eqn 4 describe the loss function for the third and fourth normal forms respectively. Every point that lies within an n-ball representing a class is a potential instance of that class. The loss functions capture this by applying relations as translations on these points (following the TransE [3] relation model). The relation vector $f_\eta(r)$ when added to the center of class C should be at a maximum distance of the sum of radii of C and D from the center of D. Eqn 4 reverses the direction of translation from Eqn 3.

$$
\begin{aligned}
\text{loss}_{C\sqcap D\sqsubseteq \bot}(c,d) = \max(0, r_\eta(c) + r_\eta(d) - \| f_\eta(c) - f_\eta(d) \| + \gamma) \\
+ |\, \| f_\eta(c) \| - 1| + |\, \| f_\eta(d) \| - 1|
\end{aligned} \tag{5}
$$

Eqn 5 describes the loss function for disjoint classes C and D while Eqn 6 refers to the specific loss function for bottom class whose radius must be equal to zero.

$$
\text{loss}_{C\sqsubseteq \bot}(c) = r_\eta(c) \tag{6}
$$

EmEl++ [9] added role constructors to EmEl. The translation approach of EmEl and EmEl++ is not suitable for more expressive description logics such as *SROIQ* because the translation operator on relations makes them one-to-one. This limits the capabilities of the model as most of the relations are many-to-many. We propose a modification to their approach, named EmEl(var), to overcome the issue.

4. Proposed Approach

In order to address the one-to-one relation restriction, we used a simple yet powerful technique that provides a foundation for further work in embeddings based description logic reasoning. We consider the relations to have a variance (uncertainty) leading to the translation having various possible regions in the vector space. This lets us model one-to-many and many-to-many relations in the ontology. As a result, we can model complex properties such as cardinality.

We consider the variance to be a hard bound. The translation of n-ball C on relation vector R could now be within σ distance of n-ball D in a tuple (C, R, D) where C, D $\in \mathbb{C}$ and R $\in \mathbb{R}'$. Hence all the points within σ distance from the translated space are related to C through R. This removes the one-to-one limitation of previous methods. We call this model EmEl(var).

Every relation has its own σ which is learnt during training. In order to avoid σ from becoming infinite, we keep the absolute value of σ as a loss component for regularisation. Figure 1(c) shows a visual representation of EmEl(var).

Hence, the definition of the geometric ontology embedding η now becomes a tuple $(f_\eta, r_\eta, \sigma_\eta)$ of functions that map classes and relations in ontology O into R^n, where f_η : C ∪ R $\mapsto \mathbb{R}^n$, r_η : C $\mapsto \mathbb{R}$ and σ_η : R $\mapsto \mathbb{R}$. The modified loss functions are provided in Eqn 7 and Eqn 8. Note that the loss function for other normal forms remain the same as in EmEl++.

$$\begin{aligned} \text{loss}_{C\sqsubseteq\exists R.D}(c,d,r) = \max(0, \| f_\eta(c) + f_\eta(r) - f_\eta(d) \| \\ + r_\eta(c) - r_\eta(d) - \sigma_\eta(r) - \gamma) \\ + | \| f_\eta(c) \| -1| + | \| f_\eta(d) \| -1| + \sigma_\eta(r) \end{aligned} \tag{7}$$

$$\begin{aligned} \text{loss}_{\exists R.C\sqsubseteq D}(c,d,r) = \max(0, \| f_\eta(c) - f_\eta(r) - f_\eta(d) \| \\ - r_\eta(c) - r_\eta(d) - \sigma_\eta(r) - \gamma) \\ + | \| f_\eta(c) \| -1| + | \| f_\eta(d) \| -1| + \sigma_\eta(r) \end{aligned} \tag{8}$$

5. Experiments

In order to demonstrate our work and take part in the SemRec challenge, we also train our model with the data provided for the challenge. The code for our work could be found here. Hence in order to test our model on the final data, follow the instructions in the ReadMe. We use the dataset provided for the challenge in order to train our models.

5.1. Model Training

In order to train the embeddings, we use Pytorch [14] and its embedding layers. Pykeen framework [15] was used for implementing TransE, TransH, and DistMult embedding models. For EmEl embeddings, source code provided by the authors was used. In order to learn the embeddings for different models, we first normalize the ontologies, i.e., convert the axioms into one of the four normal forms discussed in Section 3. All individuals in the ontology are considered as nominal classes (containing one instance) and the embeddings learn to make their radii zero resulting in a point in the vector space.

5.2. Evaluation Metrics

Subsumption is one of the reasoning tasks and it checks whether the subclass relation exists between two classes. We chose subsumption to evaluate the effectiveness of the proposed

Table 1
Performance of our model on provided dataset

Model	Top1	Top10	Top100	Median	90th Percentile
OWL2EL2	0	0	0	25769	66577
OWL2EL3	0.08	0.14	0.15	50145	247962
OWL2EL4	0	0	0.2	124.6	248.3
OWL2EL5	0	0	0.1	139.6	241.7

embeddings rather than link prediction because this task makes use of the normalized axioms to infer the subclass relation. The task of subsumption is reduced as a distance-based operation in the embedding vector space. Given a test instance of the form $C \sqsubseteq D$, we use D as source class and rank all other classes in the given ontology in an increasing order of their distance from D in the vector space. Based on the rank at which C is present in the ranked list, we evaluate our model. We hypothesise that an embedding model that successfully captures the ontological information should be able to assign very close vector representations to the two classes in a subclass relation, hence, producing a lower rank for C.

In order to evaluate the performance of different models, we use Hits at ranks 1, 10 and 100 which report the fraction of the test cases where the given class C falls under top 1, 10 and 100 in rank list respectively. Median rank and 90^{th} percentile rank were also considered to compare the overall performance of the models. A median rank of m indicates that for 50% of test cases, the correct answer was found below rank m. Similarly 90^{th} percentile rank indicates the rank below which the correct class was found for 90% of the test cases.

6. Results

The final result on the provided test data for our model is provided in Table 1. Unfortunately, the dataset OWL2EL1 didn't have sub class relations for the subsumption task.

7. Conclusion

The existing Knowledge Graph and ontology embedding approaches assume that relations are one-to-one. This limits the possibility of using these embeddings for more expressive ontologies and for complex reasoning tasks. We have provided a simple yet effective method that overcomes this obstacle and helps embeddings to capture many-to-many relations. The flexibility to model many-to-many relations also opens up the possibility of extending this work for more expressive description logics such as $\mathcal{SROIQ}$.

References

[1] G. Singh, S. Mondal, S. Bhatia, R. Mutharaju, Neuro-symbolic techniques for description logic reasoning (student abstract), Proceedings of the AAAI Conference on Artificial

Intelligence 35 (2021) 15891–15892. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17942.

[2] Z. Wang, J. Zhang, J. Feng, Z. Chen, Knowledge graph embedding by translating on hyperplanes, in: AAAI, 2014.

[3] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in: C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems, volume 26, Curran Associates, Inc., 2013. URL: https://proceedings.neurips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf.

[4] B. Yang, W. tau Yih, X. He, J. Gao, L. Deng, Embedding entities and relations for learning and inference in knowledge bases, 2015. arXiv:1412.6575.

[5] M. Nickel, V. Tresp, H.-P. Kriegel, A three-way model for collective learning on multi-relational data, in: Proceedings of the 28th International Conference on International Conference on Machine Learning, ICML'11, Omnipress, Madison, WI, USA, 2011, p. 809–816.

[6] T. Trouillon, J. Welbl, S. Riedel, E. Gaussier, G. Bouchard, Complex embeddings for simple link prediction, in: M. F. Balcan, K. Q. Weinberger (Eds.), Proceedings of The 33rd International Conference on Machine Learning, volume 48 of *Proceedings of Machine Learning Research*, PMLR, New York, New York, USA, 2016, pp. 2071–2080. URL: http://proceedings.mlr.press/v48/trouillon16.html.

[7] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, in: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI'15, AAAI Press, 2015, p. 2181–2187.

[8] M. Kulmanov, W. Liu-Wei, Y. Yan, R. Hoehndorf, 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, International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 6103–6109. URL: https://doi.org/10.24963/ijcai.2019/845. doi:10.24963/ijcai.2019/845.

[9] S. Mondal, S. Bhatia, R. Mutharaju, Emel++: Embeddings for el++ description logic, in: Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021), volume 2846, ????

[10] F. Z. Smaili, X. Gao, R. Hoehndorf, Onto2Vec: joint vector-based representation of biological entities and their ontology-based annotations, Bioinformatics 34 (2018) i52–i60. URL: https://doi.org/10.1093/bioinformatics/bty259. doi:10.1093/bioinformatics/bty259. arXiv:https://academic.oup.com/bioinformatics/article-pdf/34/13/i52/25098469/bty259_hoehndorf.30.sup.1.pdf.

[11] M. Ebrahimi, A. Eberhart, F. Bianchi, Towards bridging the neuro-symbolic gap: deep deductive reasoners, 2021. doi:10.1007/s10489-020-02165-6.

[12] O. Lutfu Ozçep, M. Leemhuis, D. Wolter, Cone semantics for logics with negation, in: C. Bessiere (Ed.), Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, International Joint Conferences on Artificial Intelligence Organization, 2020, pp. 1820–1826. Main track.

[13] F. Baader, S. Brandt, C. Lutz, Pushing the el envelope, in: Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005, p. 364–369.

[14] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035. URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[15] M. Ali, H. Jabeen, C. T. Hoyt, J. Lehman, The keen universe: An ecosystem for knowledge graph embeddings with a focus on reproducibility and transferability, in: International Semantic Web Conference, Springer, 2019.