Visualizing High-dimensional Ball-Embeddings while Keeping Topological Relations*

First Author[1], Second Author[1], Third Author[1], Forth Author[1], and Fifth Author[1,2]

[1] University of Bonn, Germany
{}@uni-bonn.de
[2] Fraunhofer IAIS, Germany

Abstract. Region-based embeddings have been proven to be useful in knowledge graph reasoning. Existing visualization tool cannot preserve topological relations among high dimensional regions. We present a tool for visualizing high-dimensional balls that keeps topological relations after their dimensions are reduced to 2-dimensions. The first version of this tool provides four functions as follows: (1) Simple diagrammatic reasoning for syllogism; (2) diagrammatic reasoning with back-ground knowledge; (3) visualizing the construction process of ball-embeddings; (4) providing a web-service to impose a taxonomy structure onto its vector embeddings with zero-energy loss. A demonstration video of this tool is available at ....

Keywords: Diagrammatic reasoning · ball embeddings · visualization.

1 Introduction

Bing able to learn from data, and robust to noisy inputs, Deep-Learning has been successful in a variety of AI tasks that have frustrated classic symbolic approaches for decades, such as object classification, machine translation, voice recognition, question-answering [9]. However, Deep-Learning Systems lack of explainability, can be fooled [13, 6, 1], and normally need much more learning data than human does [8]. This introduces potential dangers into safety critical applications, such as autonomous driving. Introducing innate structures into Deep-Learning system has been advocated, and listed as one of the AI research topics in the Townhall meeting at AAAI-19, so that Deep-Learning research shall solve logical reasoning tasks (System 2 of mind) [7, 2]. As logical reasoning, such as Syllogism, is better represented by inclusion relations among regions, instead of translation among vectors [14, 5], recent researches have been attempting to promote vectors into regions to improve performances of logical query, triple classification, link prediction of Knowledge Graph [3, 4, 10–12].

Though a powerful tool for learning from data, Deep Learning limits itself in representing everything as vectors, and can only approximate symbolic

* Supported by organization x.

representation and reasoning. This limitation can be approached by promoting vector embeddings to ball embeddings. Vector embeddings of nodes in a taxonomy structure can be promoted into balls in higher-dimensional space, such that (1) each vector embedding is well preserved by the central point of a ball; (2) child-parent relations are precisely encoded by inclusion relations among balls. Significant results are obtained in experiments on unifying word-embeddings with hypernym trees and on unifying entity-embeddings learned from knowledge-graphs and tree structures. Being able to precisely imposing external symbolic structures onto Deep Learning systems not only paves the way towards resolving the antagonism between connectionism and symbolicism in the literature, but also has tremendous value in real applications. For example, being able to impose traffic rules onto autonomous driving cars would ultimately solve the safety issue.

Ball embeddings are rigorously constructed by a sequence of geometric transformations under the condition that the loss function of the embedding must be zero, which is a condition that is not required and has not been targeted by Deep Learning approaches. This also raises the visualization problem of ball embeddings. The popular tool t-NSE for visualizing vector embeddings cannot be directly applied for visualizing ball embeddings, for the dimensional reduction process of t-NSE does not guarantee the topological relations among balls. In this paper, we demonstrate an open source system that is able to visualize ball embeddings while keeping their topological relations. The main contributions of this system are as follows: (1) it has an vivid interactive user interface that can be used for diagrammatic reasoning among taxonomy; (2) it provides an effective and friendly approach for debugging the geometric construction process of ball embeddings; (3) it provides a batch service that accepts a large scale input for construct ball embeddings.

2 The Architecture

Flask with task queue

draw a picture

3 Services

3.1 Simple Diagrammatic Reasoning

Case 1

- user inputs: Socrates is human, human is animal
- user query: what is the relation between Socrates and animal
- system draw: (1) Socarates ball inside human ball, (2) human ball inside animal ball;
- system merges (1) and (2) [the human balls in (1) and (2) may have difference sizes]
- system conclude: Socrates is animal.

Case 2

– user input: Soccer is not human, human is mortal
– user query: what is the relation between Soccer and mortal
– system draw: (1) Soccer ball outside human ball, (2) human ball inside mortal ball;
– system merges (1) and (2)
– system randomly generates Soccer balls, and keep those balls outside human ball
– system conclude: undecided.

3.2 Diagrammatic Reasoning with Background Knowledge

– user input: Soccer is not human, human is animal
– user query: what is the relation between Soccer and animal
– system replaces ‘Soccer is not human’ with ‘Soccer is entity’, system adds ’animal is entity’ by searching background knowledge

```
>>> from nltk.corpus import wordnet as wn
>>> soccer = wn.synsets('soccer')
>>> soccer
[Synset('soccer.n.01')]
>>> soccer = wn.synsets('soccer')[0]
>>> soccer
Synset('soccer.n.01')
>>> human=wn.synsets('human')[0]
>>> human
Synset('homo.n.02')
>>> soccer.lowest_common_hypernyms(human)
[Synset('entity.n.01')]
>> % check 'animal' and 'entity'
>> ..
```

– system draw: (1) Soccer ball outside animal ball, (2) human ball inside animal ball; (3) they are inside entity ball
– system merges (1), (2), and (3)
– system conclude: Soccer is not animal.

3.3 Visual Debugging

– user input: a tree structure, vector embeddings of tree nodes
– System will demonstrate the geometric construction process of ball embeddings in an interactive manner. That is, user clicks a button, System performs one geometric transformation and update the interface

3.4 Batch Service

- user provide her/his name and contact email.
- user input: a tree structure, vector embeddings of tree nodes
- System will construct ball embeddings at backend, and send the user the link fo the final ball embeddings

4 Conclusion and Outlooks

link to the video

References

1. Yonatan Belinkov and Yonatan Bisk. Synthetic and natural noise both break neural machine translation. *CoRR*, abs/1711.02173, 2017.
2. Joshua Bengio. From system 1 deep learning to system 2 deep learning. In *Keynote at NeurIPS*, page 12. Amherst, MA, 2019.
3. Tiansi Dong, Chrisitan Bauckhage, Hailong Jin, Juanzi Li, Olaf H. Cremers, Daniel Speicher, Armin B. Cremers, and Jörg Zimmermann. Imposing Category Trees Onto Word-Embeddings Using A Geometric Construction. In *ICLR-19*, New Orleans, USA. May 6-9, 2019.
4. Tiansi Dong, Zhigang Wang, Juanzi Li, Christian Bauckhage, and Armin B. Cremers. Triple Classification Using Regions and Fine-Grained Entity Typing. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, pages 77–85, Honolulu, Hawaii, USA. January 27 – February 1, 2019.
5. Janice Glasgow, N. Hari Narayanan, and B. Chandrasekaran, editors. *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* MIT Press, Cambridge, MA, USA, 1995.
6. Robin Jia and Percy Liang. Adversarial examples for evaluating reading comprehension systems. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017*, pages 2021–2031, 2017.
7. D. Kahneman. *Thinking, fast and slow.* Allen Lane, Penguin Books, 2011. Nobel laureate in Economics in 2002.
8. Brenden Lake, Ruslan Salakhutdinov, and Joshua B. Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 12 2015.
9. Yann LeCun, Yoshua Bengio, and Geoffrey E. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
10. Xiang Li, Luke Vilnis, Dongxu Zhang, Michael Boratko, and Andrew McCallum. Smoothing the geometry of box embeddings. In *International Conference on Learning Representations (ICLR)*, 2019.
11. Xin Lv, Lei Hou, Juanzi Li, and Zhiyuan Liu. Differentiating concepts and instances for knowledge graph embedding. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1971–1979. Association for Computational Linguistics, 2018.
12. H. Ren, W. Hu, and J. Leskovec. Query2box: Reasoning over Knowledge Graphs in Vector Space Using Box Embeddings. In *ICLR-20*, 2020.

13. Min Joon Seo, Aniruddha Kembhavi, Ali Farhadi, and Hannaneh Hajishirzi. Bidirectional attention flow for machine comprehension. *CoRR*, abs/1611.01603, 2016.
14. John Venn. On the diagrammatic and mechanical representation of propositions and reasonings. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, 10(58):1–18, 1880.