Text-based Knowledge Graph Embeddings

Andrea Maurino, Federico Bianchi and Marco Cremaschi
INSID&S Lab
University of Milan-Bicocca

Knowledge Graph Embeddings

Knowledge Graphs (KG) describe **entities** and the **relationships** between them (Figure 1). Generally KG contains also ontological types for entities. KG are now commonly used in many different contexts like recommendation and knowledge exploration [5].

Embedded (i.e., finding a vector representation) entities and relationships are useful for different tasks like similarity evaluation and link prediction [8]

We focus on **Text-based** Knowledge Graph Embeddings: embeddings that are generated starting from text.

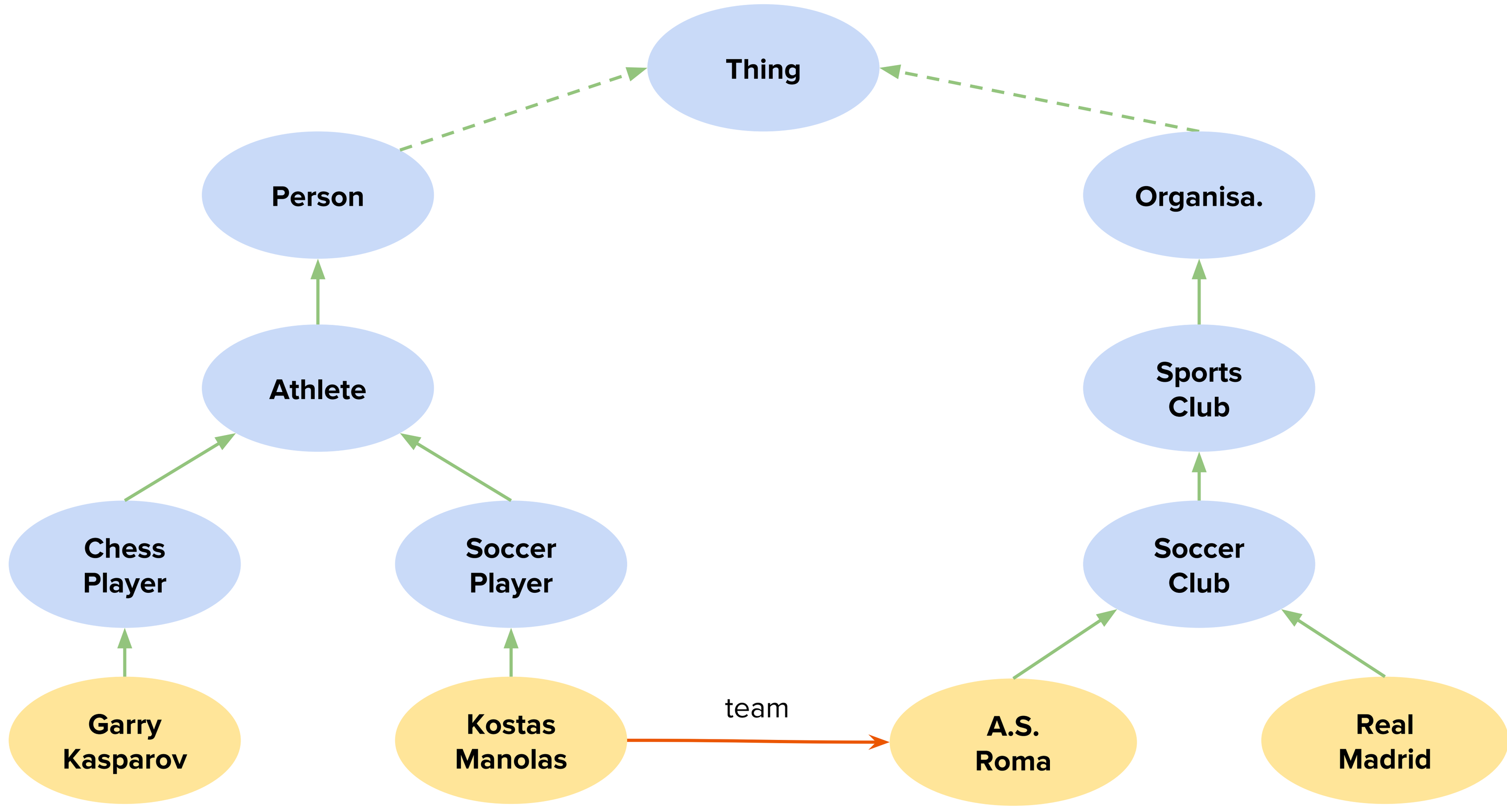

Figure 1. An example of a part of a Knowledge Graph. The entities and the types are taken from the DBpedia KG.

Current Work on Topic

Our current approach takes texts that are semantically annotated, i.e., texts where **mentions of entities** in the KG are recognized [4,6] (see Figure 2, where **dbr** is our reference to the DBpedia KG).

We use word2vec [7] to find representations for entities: two entities that appear in similar textual contexts will be closer in the vector space. Considering that entities have ontological types e.g., Politician, City, etc., we are able to learn embeddings also for types and represent entities and their types as concatenation of these vectors [2]

We have proposed a method to **embed periods of time** into a vector space by considering time descriptions [4].

Since the meaning of words changes over time, we have explored the implications of temporality in word embeddings. We have introduced a new model in the state of the art that can be used to create **word representations of specified time periods** [1].

Training Text-based KG Embeddings

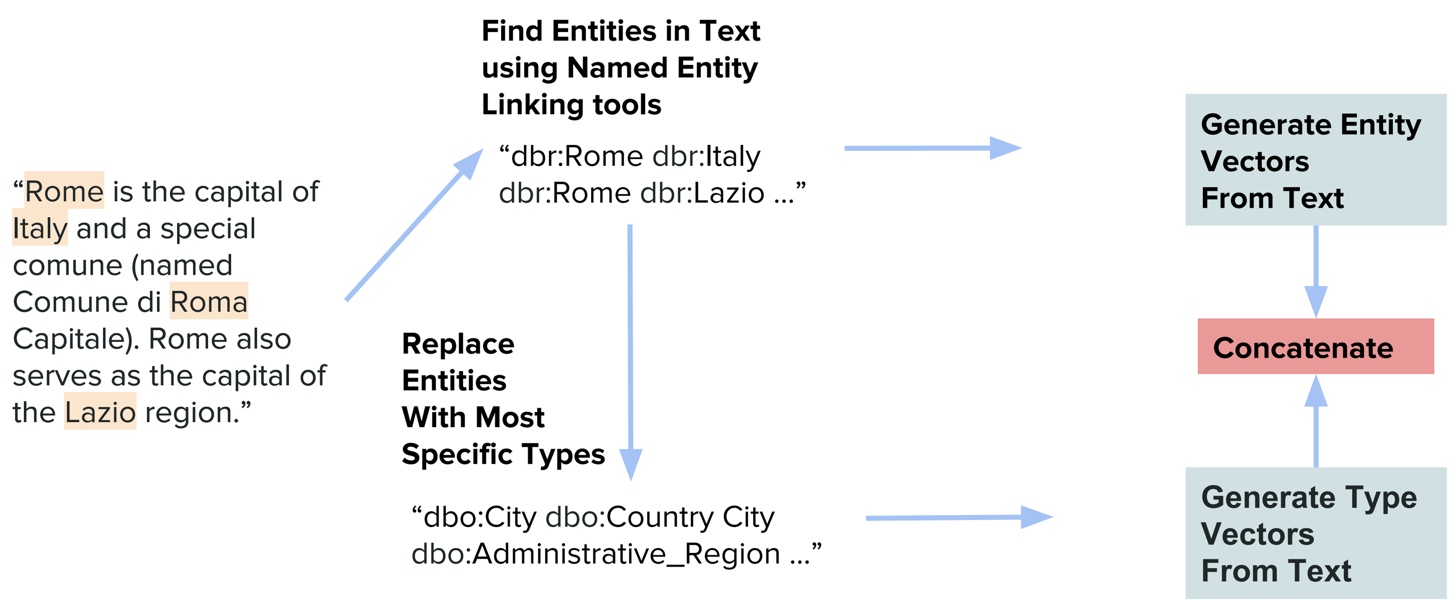

Figure 2. Our current approach used to embed entities in a vector space

Next Steps

- **Text-Enhanced Knowledge Graph Embeddings**
  - Represent relationships in our model since they are currently missing;
  - Explore the use of different semantic annotators to enhance the representations of the embeddings: better annotators = better embeddings;
  - Use our embeddings on news corpora;
  - Apply embeddings to recommendation and link prediction tasks [8].
- **Hyperbolic Embeddings**
  - Explore embeddings that consider hyperbolic geometry instead of standard euclidean [9].
- **Reasoning in Embedded Spaces**
  - Introduce reasoning mechanisms in the embedded spaces (e.g., use a reasoner to make inferences about entities and relationships).

First Results on Reasoning

We explored the use of **Logic Tensor Networks** (LTNs) [10] a model that combines logics and neural networks. LTNs use Neural Networks to compute the truth value of **first order fuzzy logic** formulae.

We recently analyzed the possibilities of this model in context of **deductive reasoning** [3].

Results show that the model can perform deductive reasoning while using a deep neural network.

Our next step is to to **combine** this model with our KG embeddings.

References

[1] Di Carlo, V., Bianchi, F. & Palmonari, M. (2019). *Training Temporal Word Embeddings with a Compass*. AAAI (to appear).

[2] Bianchi, F., Soto, M. Palmonari, M., & Cutrona, V. (2018, June). *Type Vector Representations from Text: An empirical analysis*. in DL4KGS@ESWC.

[3] Bianchi, F. & Hitzler, P. (2019). *On the Capabilities of Logic Tensor Networks for Deductive Reasoning*. AAAI Spring Symposium MAKE (to appear).

[4] Bianchi, F., Palmonari, M., & Nozza, D. (2018, October). *Towards Encoding Time in Text-Based Entity Embeddings*. In ISWC.

[5] Bianchi, F., Palmonari, M., Cremaschi, M., & Fersini, E. (2017, May). *Actively Learning to Rank Semantic Associations for Personalized Contextual Exploration of Knowledge Graphs*. In ESWC.

[6] Bianchi, F. and Palmonari, M. (2017, November). *Joint Learning of Entity and Type Embeddings for Analogical Reasoning with Entities*. In NL4AI@AI*IA.

[7] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality*. In NIPS.

[8] Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). *Translating embeddings for modeling multi-relational data*. In NIPS.

[9] Nickel, M., & Kiela, D. (2017). *Poincaré embeddings for learning hierarchical representations*. In NIPS.

[10] Serafini, L., & Garcez, A. D. A. (2016). *Logic tensor networks: Deep learning and logical reasoning from data and knowledge*. arXiv preprint arXiv:1606.04422.

@fb_vinid @andreamaurino @cremmarco