Generating Descriptions of Entity Relationships

Extended Abstract*

Nikos Voskarides
University of Amsterdam
Amsterdam, The Netherlands
n.voskarides@uva.nl

Edgar Meij
Bloomberg L.P.
London, United Kingdom
edgar.meij@acm.org

Maarten de Rijke
University of Amsterdam
Amsterdam, The Netherlands
derijke@uva.nl

ABSTRACT

Results displayed on a modern search engine result page (SERP) are sourced from multiple, heterogeneous sources. For so-called organic results it has been known for a long time that result snippets, i.e., brief descriptions explaining the result item and its relation to the query, positively influence the user experience [7]. In this paper, we focus on generating descriptions for results sourced from another important ingredient of modern SERPs: knowledge graphs. Knowledge graphs (KGs) contain information about entities and their relationships. A large and diverse set of search applications utilize KGs to improve the user experience. For instance, web search engines try to identify KG entities in queries and augment their result pages with knowledge graph panels that provide contextual entity information [3, 6]. Such panels usually focus on a single entity and may include attributes of the entity and other, related entities.

Entities can be connected with more than one relationship in a KG, however. For example, two actors might have appeared in the same film, be born in the same country and also be partners. Recent work has focused on finding relationships between a pair of entities and ranking the relationships by a predefined relevance criterion [4]. When using relationships in real-world search applications, with SERPs being the prime example, a crucial problem is that they are typically represented in a formal manner that is not suitable to present to an end user. Instead, human-readable descriptions that verbalize and provide context about entity relationships are more natural to use [5]. They can be used, e.g., for entity recommendations [2] or for KG-based timeline generation [1].

Descriptions of KG relationships themselves are usually not included in large-scale knowledge graphs and previous work on automatically generating such descriptions has either relied on hand-crafted templates [1] or on external text corpora [8]. The main limitations of the former are that manually creating these templates is expensive, not generalizable, and thus it does not scale well. The latter approach is limited as the underlying text corpus may not contain descriptions for all certain relationship instances; it will not produce meaningful results for instances that do not appear in the text corpus.

We propose a method that overcomes these limitations by automatically generating descriptions of KG entity relationships. Since there exist textual descriptions of a certain relationship for some relationship instances, we aim to use these descriptions to learn how the relationship is generally expressed in text and use this information to generate descriptions for other instances of the same relationship. Existing relationship descriptions are usually complex and tailored to the entities they discuss. Also, it is likely that the KG does not contain all the information included in a description. For example, the KG might not contain any information about the second part of the following sentence: "*Catherine Zeta-Jones starred in the romantic comedy The Rebound, in which she played a 40-year-old mother of two ...*". Nevertheless, descriptions of the same relationship share patterns that are specific to that relationship. Therefore, we first create sentence templates for a certain relationship and then, for a new relationship instance, we select appropriate templates, which we formulate as a ranking problem, and fill them with the appropriate entities to generate a description.

We propose a method that generates descriptions of entity relationships for a relationship instance given a knowledge graph and a set of relationship instances coupled with their descriptions; we evaluate this method using both automatic and manual evaluation methods, and release the datasets used to the community.[1] We have found that when using information about the relationship instance and the template taken from the KG both automatic and manual evaluation outcomes are improved. We have also found that a supervised method that uses both KG features and other template features (template words, number of entities) consistently outperforms an unsupervised method on all automatic evaluation metrics and also in terms of validity and informativeness.

As to future work, our error analysis showed that we need more sophisticated modeling for capturing the semantic similarity between a relationship instance and a template, especially for capturing temporal dimensions that also involve other relationship instances. We also want to explore more sophisticated methods for selecting the correct surface form for an entity to improve grammaticality. Finally, we aim to evaluate our method on generating descriptions for less popular KG relationships.

REFERENCES

[1] T. Althoff, X. L. Dong, K. Murphy, S. Alai, V. Dang, and W. Zhang. Timemachine: Timeline generation for knowledge-base entities. In *KDD*, 2015.
[2] R. Blanco, B. B. Cambazoglu, P. Mika, and N. Torzec. Entity recommendations in web search. In *ISWC*, 2013.
[3] R. Blanco, G. Ottaviano, and E. Meij. Fast and space-efficient entity linking for queries. In *WSDM*, 2015.
[4] L. Fang, A. D. Sarma, C. Yu, and P. Bohannon. Rex: explaining relationships between entity pairs. In *VLDB*, 2011.
[5] D. Gkatzia, O. Lemon, and V. Rieser. Natural language generation enhances human decision-making with uncertain information. In *ACL*, 2016.
[6] T. Lin, P. Pantel, M. Gamon, A. Kannan, and A. Fuxman. Active objects: Actions for entity-centric search. In *WWW*, 2012.
[7] A. Tombros and M. Sanderson. Advantages of query biased summaries in information retrieval. In *SIGIR*, 1998.
[8] N. Voskarides, E. Meij, M. Tsagkias, M. de Rijke, and W. Weerkamp. Learning to explain entity relationships in knowledge graphs. In *ACL-IJCNLP*, 2015.
[9] N. Voskarides, E. Meij, and M. de Rijke. Generating descriptions of entity relationships. In *ECIR 2017: 39th European Conference on Information Retrieval*, LNCS. Springer, April 2017.

*The full version of this paper was published at ECIR 2017 [9].

[1] https://github.com/nickvosk/ecir2017-gder-dataset/