

Keynote: Learning New Type Representations from Knowledge Graphs

Soumen Chakrabarti
India Institute of Technology Bombay
soumen@cse.iitb.ac.in

Abstract

Beyond words, continuous representations of entities and relations have led to large recent improvements in inference of facts in knowledge bases, as well as applications like question answering. Comparatively less has been done about modeling types and their associated relations (is-instance-of and is-subtype-of). In the first part of the talk, I will present a new representation of types as hyper-rectangles rather than points, which are commonly used to embed words and entities. I will propose an elementary loss function representing rectangle containment. I will also demonstrate that recent work on type representation has used a questionable evaluation protocol, and propose a sound alternative. Experiments using type supervision from the WordNet noun hierarchy show the superiority of our approach. In the second part of the talk, I will move to unsupervised discovery of type representation. The idea is to represent each entity using a type and a residual vector. Each relation is represented by two type-checking vectors and an entity-to-entity compatibility checking vector. We do not use any supervision from KG schema to guide the type (checking) embeddings. Experiments on FB15k and YAGO show two benefits. First, inferring new triples becomes more accurate, exceeding state of the art. Second, the type embeddings are very good predictors of KG types to which the entities belong, although this information was not available during training.

Bio

Soumen Chakrabarti a Professor of Computer Science at IIT Bombay. He got his PhD from University of California, Berkeley and worked on Clever Web search and Focused Crawling at IBM Almaden Research Center. He has also worked at Carnegie-Mellon University and Google. He works on linking unstructured text to knowledge bases and exploiting these links for better search and ranking. Other interests include link formation and influence propagation in social networks, and personalized proximity search in graphs. He has published extensively in WWW, SIGKDD, EMNLP, VLDB, SIGIR, ICDE and other conferences. His work on keyword search in databases got the 10-year influential paper award at ICDE 2012. He is also the author of one of the earliest books on Web search and mining.

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: Joint Proceedings of the First International Workshop on Professional Search (ProfS2018); the Second Workshop on Knowledge Graphs and Semantics for Text Retrieval, Analysis, and Understanding (KG4IR); and the International Workshop on Data Search (DATA:SEARCH'18). Co-located with SIGIR 2018, Ann Arbor, Michigan, USA – 12 July 2018, published at <http://ceur-ws.org>