

Knowledge Representation with Conceptual Spaces

Steven Schockaert

School of Computer Science and Informatics, Cardiff University, UK
SchockaertS1@cardiff.ac.uk

Bio

Steven Schockaert is a professor at Cardiff University. His current research interests include common-sense reasoning, interpretable machine learning, vagueness and uncertainty modelling, representation learning, and information retrieval. He holds an ERC Starting Grant, and has previously been supported by funding from the Leverhulme Trust, EPSRC, and FWO, among others. He was the recipient of the 2008 ECCAI Doctoral Dissertation Award and the IBM Belgium Prize for Computer Science. He is on the board of directors of EurAI, on the editorial board of Artificial Intelligence and an area editor for Fuzzy Sets and Systems. He was PC co-chair of SUM 2016 and the general chair of UKCI 2017.

Abstract

Entity embeddings are vector space representations of a given domain of interest. They are typically learned from text corpora (possibly in combination with any available structured knowledge), based on the intuition that similar entities should be represented by similar vectors. The usefulness of such entity embeddings largely stems from the fact that they implicitly encode a rich amount of knowledge about the considered domain, beyond mere similarity. In an embedding of movies, for instance, we may expect all movies from a given genre to be located in some low-dimensional manifold. This is particularly useful in supervised learning settings, where it may e.g. allow neural movie recommenders to base predictions on the genre of a movie, without that genre having to be specified explicitly for each movie, or without even the need to specify that the genre of a movie is a property that may have predictive value for the considered task. In unsupervised settings, however, such implicitly encoded knowledge cannot be leveraged.

Conceptual spaces, as proposed by Grdenfors, are similar to entity embeddings, but provide more structure. In conceptual spaces, among others, dimensions are interpretable and grouped into facets, and properties and concepts are explicitly modelled as (vague) regions. Thanks to this additional structure, conceptual spaces can be used as a knowledge representation framework, which can also be effectively exploited in unsupervised settings. Given a conceptual space of movies, for instance, we are able to answer queries that ask about similarity w.r.t. a particular facet (e.g. movies which are cinematographically similar to Jurassic Park), that refer to a given feature (e.g. movies which are scarier than Jurassic Park but otherwise similar), or that refer to particular properties or concepts (e.g. thriller from the 1990s with a dinosaur theme). Compared to standard entity embeddings, however, conceptual spaces are more challenging to learn in a purely data-driven fashion. In this talk, I will give an overview of some approaches for learning such representations that have recently been developed within the context of the FLEXILOG project.