

06.11.2019

Call for Bachelor/Master Thesis

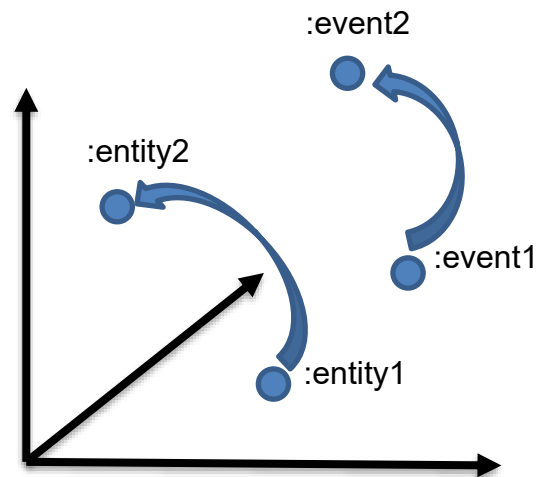
“Learning Machine Learning-based Embeddings for Entities, States, and Events”

(in English/German)

What is the topic?

Machine learning-based embedding techniques have been used with great success in recent years for learning representations of words, entities, relations, etc. in a vector space (see word2vec, doc2vec, fastText, BERT). However, several aspects have not been considered so far:

1. Dynamics: Embeddings have been applied to represent static knowledge (entities, facts, events). How can dynamical aspects (state changes of entities, facts, events) be represented as embeddings and how can embedding changes over time be modeled (e.g., embedding of Donald Trump before, immediately after, and long after the inauguration)?
2. Inter-subjectivity: Learning embeddings entails that the world knowledge to be represented is considered from a single perspective (subject). However, agents (people or databases) might have different world knowledge and might agree on knowledge only to some extent. How can this be represented in embeddings? How can embedding vector spaces be aligned dynamically?



To answer those questions, the student will (after a literature review) train embeddings for entities, facts, events, and states, and will analyze how the embeddings are associated with each other concerning the evolution over time and concerning different agents.

The content of this thesis will be based on the theoretical foundations described in [1]. More information and first literature will be provided in a first meeting.

What are the prerequisites?

- Basic knowledge in machine learning, particularly embeddings (e.g., word2vec, fastText).
- Having fun with coding (preferably in Python).

[1] http://dbis.informatik.uni-freiburg.de/content/team/faerber/papers/Theories_World_CARLA2018.pdf

Contact:
Dr. Michael Färber
michael.farber@kit.edu