Karlsruher Institut für Technologie

KOLLOQUIUM

Informatik-Sonderkolloquium

Briefly confused? It's not you, it's the context

Ira Assent
(Professor in Computer Science, Aarhus University, Denmark)

Abbreviations are convenient when we refer to a concept again and again, in particular if it is a long or complicated term. Naturally, due to the brevity of typical abbreviations (often just two or three letters), abbreviations are typically ambiguous. As humans we are very good at implicitly translating the abbreviation into the correct long form if we are familiar with its context or when we have been introduced to the abbreviation earlier in the text. For computers, however, the situation is not that straightforward. In natural language processing (NLP) applications such as automatic grouping of documents according to topic (e.g. retrieving research articles relevant for a study), it is therefore important to solve the problem of abbreviation disambiguation. In this talk, we present an approach that is based on capturing the context of abbreviations and their long forms by learning word embeddings from shallow neural networks. This approach lends itself to automatic disambiguation, and to identification of terms for which further usage examples are required.

KIT – Campus Süd, Fakultät für Informatik, Am Fasanengarten 5, 76131 Karlsruhe, www.informatik.kit.edu

Mittwoch, 05.02.20, 14:30 Uhr

Informatik-Hauptgebäude (50.34), SR 348 (3. OG), Am Fasanengarten 5, 76131 Karlsruhe

KIT – Die Forschungsuniversität in der Helmholtz-Gemeinschaft