Embeddings or Visions of the Future of the Lexicon

Hinrich Schütze
LMU München
dehs2016@cis.lmu.org

One way to model the predictability of a linguistic unit alpha in context gamma is to represent alpha and gamma as vectors a and c in high-dimensional space and estimate $P(\text{alpha}|\text{gamma})$ based on a and c. These types of vector representations are commonly referred to as embeddings. I will give a general talk about embeddings – covering issues like learning, granularity and compositionality – and how they relate to the traditional lexicon in computational linguistics. At the end, I will turn to the question of what the implications are for models of predictability.