Invited Talk

Thursday 29th August 2019

Quantitative Computational Syntax: dependencies, intervention effects and word embeddings

Paola Merlo
University of Geneva

Abstract

In the computational study of intelligent behaviour, the domain of language is distinguished by the complexity of the representations and the vast amounts of quantitative text-driven data. In this talk, I will let these two aspects of the study of language inform each other and will discuss current work investigating whether the notion of similarity in the intervention theory of locality is related to current notions of similarity in word embedding space.

Despite their practical success and impressive performances, neural-network-based and distributed semantics techniques have often been criticized as they remain fundamentally opaque and difficult to interpret. Several recent pieces of work have investigated the linguistic abilities of these representations, and shown that they can capture long agreement and thus hierarchical notions. In this vein, we study another core, defining and more challenging property of language: the ability to construe long-distance dependencies. We present results that show that word embeddings and the similarity spaces they define do not correlate with experimental results on intervention similarity in long-distance dependencies. These results show that the linguistic encoding in distributed representations does not appear to be human-like, and it also brings evidence to the debate on narrow or broad definitions of similarity in syntax and sentence processing.

Short bio

Paola Merlo is associate professor in the Linguistics department of the University of Geneva. She is the head of the interdisciplinary research group Computational Learning and Computational Linguistics (CLCL). The group is concerned with interdisciplinary research combining linguistic modelling with machine learning techniques. Prof. Merlo has been editor of Computational Linguistics, published by MIT Press and a member of the executive committee of the ACL. Prof. Merlo holds a doctorate in Computational Linguistics from the University of Maryland, and has been associate research fellow at the University of Pennsylvania, and visiting scholar at Rutgers, Edinburgh, Stanford and Uppsala.