Syntactic Parsing in Humans and Machines

Paola Merlo

Computational Learning and Computational Linguistics,
University of Geneva, Switzerland
`paola.merlo@unige.ch`

[   ]

Abstract

To process the syntactic structures of a language in ways that are compatible with human expectations, we need computational representations of lexical and syntactic properties that form the basis of human knowledge of words and sentences.

Recent neural-network-based and distributed semantics techniques have developed systems of considerable practical success and impressive performance. As has been advocated by many, however, such systems still lack human-like properties. In particular, linguistic, psycholinguistic and neuroscientific investigations have shown that human processing of sentences is sensitive to structure and unbounded relations.

In the spirit of better understanding the structure building and long-distance properties of neural networks, I will present an overview of recent results on agreement and island effects in syntax in several languages. While certain sets of results in the literature indicate that neural language models exhibit long-distance agreement abilities, other finer-grained investigation of how these effects are calculated indicates that that the similarity spaces they define do not correlate with human experimental results on intervention similarity in long-distance dependencies. This opens the way to reflections on how to better match the syntactic properties of natural languages in the representations of neural models.