Introducing Structure into Neural Network-Based Semantic Models

Stephen Clark
University of Cambridge and DeepMind
stephen.clark@cl.cam.ac.uk

In this talk I will describe two attempts at introducing syntactic structure into semantic models using neural network architectures. The first study focuses on a particular grammatical construction, namely relative clauses, and centers around the design of a new dataset for testing compositional distributional models. The dataset is called REL-PRON, and consists of pairs of terms and properties, such as

> telescope : device that astronomer uses.

The idea is that a good compositional model will produce a vector representation of the property which is close to the vector for the term.

The second study focuses on deriving semantic vectors for phrases and whole sentences, which are then used in two tasks: sentence entailment, and a dictionary definition-term matching task. A feature of the proposed solution is that the syntactic structure for a sentence is induced in an unsupervised fashion, trained end-to-end in order to optimize for the final task.

I will finish with some thoughts on whether and how insights from traditional models of syntax and semantics can contribute to semantic models based on neural networks.