Deep Learning and Semantic Interpretation of Natural Language

Abstract:

Classical approaches to formal and computational semantics assign values to the terminal elements of hierarchical syntactic structures and define combinatorial operations on the semantic representations of phrases to compute the values of sentences. While these approaches offer formally elegant models of interpretation, they have not produced wide coverage systems. They do not provide for semantic learning. They have also not succeeded in integrating lexical and compositional semantics in an interesting or computationally efficient way. Recent developments in image caption generation suggest an alternative approach, which can overcome these difficulties. This work formulates the problem of matching images with descriptions as a task in machine translation. Deep neural networks use an encoder to map regions of pixels in an image to vector representations of graphic features, and a decoder to align these features with the distributional vectors of lexical and phrasal items. This approach can be generalized to deep neural networks that identify correspondences between multi-modal data structures and sentences. To the extent that this research program is successful, it will satisfy the core objective of the classical formal semantic program. It will assign truth (fulfilment) conditions to the sentences of a language, where these conditions are specified in terms of multi-modal representations of situations (scenes) in the world. These correspondences are generated not by a recursive definition of a truth predicate in a formal semantic theory, but by an extended deep neural language model.