Fast Semantic Extraction Using a Novel Neural Network Architecture

Ronan Collobert and Jason Weston
NEC Laboratories America, Princeton, NJ
ronan@collobert.com and jaseweston@gmail.com

Semantic understanding of text plays an important role in many fashionable tasks such as information (e.g. opinion) extraction, in web-crawling systems, question and answering based systems (e.g. for call centers), as well as in machine translation, summarization and search. End user applications typically have to deal with an enormous quantity of data, e.g. web based systems process large numbers of documents, and interactive human-machine applications require almost instant response.

While complete semantic understanding is still a far-distant goal of natural language processing, several intermediate processing tasks are useful for the tasks described above. These range from syntactic analysis (such as part-of-speech labeling and parsing) to semantic analysis (word-sense disambiguation, semantic-role labeling, named entity extraction, co-reference resolution and entailment). Unfortunately, the state-of-the-art solutions of many of these tasks are simply too slow for the applications previously described. For example, state-of-the-art syntactic parsers typically have cubic complexity in the sentence length [2], and several semantic extraction algorithms use the parse-tree as an initial feature.

In this work, we focus on the *semantic role labeling* problem: being able to give a semantic role to a syntactic consituent of a sentence. (However, our results could extend to other domains as well.) Because of its nature, role labeling seems to require the syntactic analysis of a sentence before attributing semantic labels. Using this intuition, state-of-the-art systems first build a *parse tree*, and syntactic constituents are then labeled by feeding hand-built features extracted from the parse tree to a machine learning system, e.g. an SVM [1]. This is rather slow, partly because of the parse tree component, and partly because of the use of SVMs which are polynomial in training time, and linear in testing time w.r.t the number of examples. In other words, one cannot apply this method to interesting end user applications.

Here, we propose a radically different system: we employ a *new* neural architecture which is able to learn *directly* the semantic role of a *word* in a sentence given a predicate (e.g. verb) in this sentence, *without the knowledge of syntactic constituents in the sentence*. Not only is this system *fast*, but the learning is also "end-to-end": features are learned automatically.

In practice, our system gives similar performance to state-of-the-art systems based on the parse tree, whilst being *several* orders of magnitude faster.

References

[1] S. Pradhan, W. Ward, K. Hacioglu, J. Martin, and D. Jurafsky. Shallow semantic parsing using support vector machines. *Proceedings of HLT/NAACL-2004*, 2004.

[2] D. H. Younger. Recognition and parsing of context-free languages in time n^3. *Information and Control*, 10, 1967.

Topic: text meaning
Preference: oral