An Analysis of Logical and Vector Space Approaches to Semantics

Daoud Clarke

November 1, 2012

1 Introduction

Current approaches to vector-based semantics take the following two-stage approach:

- Find vectors for terms based on their occurrences in large corpora
- Define or learn a composition for these vectors

Here we propose investigating the theory behind a more holistic approach in which the vectors and composition are learnt simultaneously.

This idea is inspired by looking at logical approaches to semantics and contrasting them with current approaches to vector-based semantics.

2 Approaches to Logical Semantics

There are two main approaches to logical semantics, which I shall call the "theorem proving" approach and the "model building" approach. In either case I assume we have the following:

- A collection of sentences forming the **background knowledge**. This may include pseudo-sentences derived from an ontology, such as "all cats are animals".
- The **text** and **hypothesis** sentences

One question we may ask is whether the text implies the hypothesis given the background knowledge. Other questions we may be interested in is whether the text and hypothesis are paraphrases (logically equivalent), or contradictory given the background knowledge.

In both approaches, each sentence is translated to some logical form. The background knowledge is represented by the logical form B, the logical form of the text by T, and the logical form of the hypothesis by H.

Theorem proving approach

In this approach, to determine entailment we simply see if we can prove $B \wedge T \rightarrow H$ using a theorem prover. Normally, we would simultaneously trying to build a model for $\neg(B \wedge T \rightarrow H)$; if we do find a model then we know we can give up trying to prove the theorem.

The other tasks can be attacked similarly, for example the text is inconsistent with the hypothesis given the background knowledge if we can prove that $B \wedge T \rightarrow \neg H$.

Model building approach

There are several ways we can use model builders with natural language semantics, but the basic idea is that given some knowledge, we can build a model of the world that is representative of that knowledge. We then query the model to determine what is true.

Under this approach, the model that is built is only required to be consistent, so any assumptions can be made building the model as long as these are not inconsistent with the knowledge that is presented. This allows for a much more flexible approach than the theorem proving approach.

A simple way you might use a model builder to determine entailment is to build a model for $B \wedge T$ and see if H is true in that model.[1]

3 Analogies with Vector Space Approaches

The approach of Clark et al. (2008) is most closely related to the model building approach of logical semantics. The "model" is the representation of the words as linear operators. Combining these operators for a sentence results in a "truth value" just as when you have a logical model, you can get a truth value for any sentence.

This analogy suggests new ways we can use use Clark et al's approach. For example, we can build a model using the background knowledge (a text corpus) and evaluate the degree to which the hypothesis is true in the model, then a build a model using the background model and the text sentence and evaluate the degree to which the hypothesis is true with respect to the new model. Following Glickman and Dagan (2005), we can view entailment as holding if the degree to which the hypothesis is true is greater when the text is included along with the background knowledge.

It also suggests that the approach of learning vectors and then learning how to compose them is sub-optimal. What we should really be doing is learning vectors and composition together such that when we compose sentences from the background knowledge we get a high value, and other sentences (or non-sentences) give us a low value.

[1] The problem with this approach is that any additional assumptions we make in building the model could potentially make H arbitrarily true or false. There are ways to work around this that are not relevant to the discussion — see Bos and Markert (2006).

References

Johan Bos and Katja Markert. When logical inference helps determining textual entailment (and when it doesn't). In *Proceedings of the Second PASCAL Challenges Workshop on Recognising Textual Entailment*, 2006.

Stephen Clark, Bob Coecke, and Mehrnoosh Sadrzadeh. A compositional distributional model of meaning. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*, pages 133–140, Oxford, UK, 2008.

O. Glickman and I. Dagan. A probabilistic setting and lexical cooccurrence model for textual entailment. In *Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment*, pages 43–48. Association for Computational Linguistics, 2005.