Recognizing Textual Entailment:  
A Deep Semantic Approach via Dependency Parsing

The Problem:

Textual entailment asks the question “does the truth of some text, T, imply the truth of some other text, H?” More formally, given a pair of texts (T, H) (also known as the “text” and the “hypothesis”) for which T is true, may we conclude that H is true, false, or neither? For example, if T = “Smith murdered Johnson” and H = “Johnson is dead,” then the entailment relation is true. If, instead, H = “Johnson is alive and well,” then the relation is false. And if H were the statement “Johnson was a helicopter pilot,” then the relation is neither true nor false.

It is important to note that the PASCAL Recognizing Textual Entailment Competition (RTE) defines textual entailment as applying not only when T logically implies H, but also when “a human reading T would infer that H is most likely true.” [1]

Textual Entailment is a problem with many potential useful applications in Natural Language Processing. In question answering (QA), textual entailment may be used to determine if a document (or set of documents) implies the truth of a question (or rather, the statement form of a yes-no question). In machine translation (MT), textual entailment could be used for evaluation, testing whether output translations and their corresponding reference translations entail one another. Similarly, textual entailment may be used for automatic evaluation of paraphrasing systems, or document summary systems.

Why Recognizing Textual Entailment is Difficult:

Detecting textual entailment is a difficult task in that there are a number of ways in which entailment relations may occur. Zaenen, et. al. establish an ontology of four different types of textual entailment. [2]

The first, simple entailment, occurs because of strict logical relationships between word meanings, e.g. “Jane planted a willow” entails “Jane planted a tree” in just this manner, as “willow” is defined as a type of tree.

Another kind of entailment is conventional implicature, in which the hypothesis is upheld through the author’s framing of information in the text. For instance, “Amanda didn’t know that the goods were stolen” implies the hypothesis “The goods were stolen” because the author commits to that fact through the use of the word “know.”

Conversational implicature, on the other hand, implies that the author will share all the information he or she has, e.g. “Mark ate three burgers” entails “Mark did not eat four burgers” under the assumption of conversational implicature.
The fourth, and perhaps most problematic, variety of entailment is that which requires world knowledge. While it is true that the statement “I will attend the next summer Olympics” implies “I will go to Brazil in 2016,” such an inference relies on the knowledge that the next summer Olympics will be held in Brazil in 2016. The degree to which we should require RTE systems to reason with background knowledge is still an open question.

**What Approaches Have Been Tried:**

A variety of approaches have been implemented in attempts at tackling the textual entailment problem. The most basic approaches employ simple similarity testing (e.g. bag-of-words with features like stemming, lemmatization, and part-of-speech tagging). Other approaches represent the texts structurally (e.g. as a graph), and estimate entailment through transforming the structures into one another, and computing a cost function on the transformations. [3] Some approaches frame the problem as proof solving, deriving a “proof” of the hypothesis through iterated rule application. The latter-most approach is a general framework that potentially allows for interfacing with arbitrary modules that contribute more background knowledge or linguistic sophistication. [4]

**Summary of Last Semester’s Work:**

In the fall semester, I designed and implemented a basic RTE system that operates in two stages: translation and inference. First, dependency parses are computed for the text and hypothesis, using the Stanford Dependencies. [7] Hand-written rules are then used to extract basic semantic representations (relation triples) from the dependency parses, completing the translation stage. In the inference stage, the semantic representations of the text and hypothesis are compared to determine whether an entailment relation exists. It is important to note that I designed this RTE system to work in a very narrow linguistic domain of simple descriptive sentences, like “The green apple is poisonous.”

**Proposal for This Semester’s Work:**

This semester, I plan to continue developing this RTE system, further extending the domain of inputs it may correctly classify. Intended extensions include: use of lexical resources (e.g., WordNet) to identify pertinent relations like synonymy and hyponymy, analysis of relative clauses (“The apple that is green...”), consideration of conjunctions (“The apple is green and red.”), as well as analysis of presuppositional contexts (e.g., “He doesn’t know that the apple is green”). I have chosen to focus on these particular domain extensions because the information needed to resolve inputs in these domains appears to be contained in the dependency parses.
Deliverables:

At the end of the semester, I plan to submit:

- All code, with relevant instructions/comments
- A code demo, if applicable
- A written report, explaining the intellectual content of the project, as well as the results of the project, and an analysis of the results

Sources:


