

## Semantic Modeling of Textual Entailment

This talk describes an application of a formal semantic model to natural language entailments as demonstrated in the Recognizing Textual Entailment (RTE) corpus (Dagan et al. 2006). The example below illustrates an *entailment pair* from the corpus:

- Premise: The largest search engine on the web, Google, receives over 200 million queries each day through its various services.
- Conclusion: Google operates on the web.

The model incorporates a typed lexicon that encodes the semantic modeling for some of the most common inferential phenomena in the corpus: appositive, intersective and restrictive modification, as well as simple existential and universal quantification. Words are assigned denotations by binding them to lexical items and sentences are analyzed according to the principle of compositionality. Entailment is defined by applying Tarski's truth-conditionality criterion to the sentential terms calculated for the premise and conclusion of a given entailment.

An annotation platform that implements the model was developed in order to apply it to entailment data. The platform combines a stochastic parser, a part-of-speech tagger, an implementation of the typed lexicon, a lambda-calculus engine and a theorem prover in first order logic. The workflow with the platform is semi-automatic, as human intervention is required in order to fix parsing errors, to correct mistakes in the mapping of part-of-speech tags to the lexicon, and to indicate lexical relations. The platform is logically sound with respect to the formal semantic model. Thus, when the theorem prover is able to generate a deductive process from the term of the premise to the term of the conclusion, this indicates that the semantic theory accounts for the entailment.

The platform was used to evaluate the applicability of the semantic model to natural data from the RTE. Entailments categorized as positive were annotated using an annotating-by-proving methodology: a pair is considered well-annotated only if the term calculated for the conclusion can be deduced by the theorem prover from the term of the premise. Entailments categorized as negative are created in a way that isolates a single linguistic phenomenon that blocks the inference. A methodology of simplification was used in order to preserve the natural quality of the data while avoiding semantic phenomena that are not modeled and syntactic constructions that are not supported. This annotation work evolved into the freely available corpus of Semantic Annotation of Textual Entailment (SemAnTE) which consists of 600 pairs in a positive-negative ratio of 2:1.