Sentence Interpretation using Stochastic Finite State Transducers

Frederic Bchet, Christian Raymond and Renato De Mori
LIA CNRS BP 1228 , 84911 Avignon Cedex 9 - France
{frederic.bechet, christian.raymond, renato.demori}@lia.univ-avignon.fr

Abstract

An effective way of representing the meaning of a utterance is with frame structures in which a type of sentence is represented by a set of property/value slots. Properties can types of verbs and cases and values are extracted from a sentence and should respect constraints represented by case relations and selectional restrictions involving word senses organized in type hierarchies.

Properties and values can be obtained as the output of Stochastic Finite State Transducers (SFST) based on property specific language models combined with generic n-gam models. In this way, sentence interpretation and recognition are carried out by the same search process.

LM adaptation can be performed by dynamically modifying the probability of each SFST based on system expectations. Phrases accepted by different SFSTs may share words, especially if different SFST recognize constituents of the same frame. For this reason, search for the most likely interpretation has to consider promising (possibly overlapping) hypotheses generated by SFSTs and the best combination of them into an acceptable semantic structure.

Using different types of acoustic confidence measures and indices of consistency, it is possible to evaluate the probability that each semantic component that has been hypothesized is correct. These probabilities can be used by the dialogue strategy to decide about specific clarification and confirmation actions.

SFSTs can be constructed using semi-automatic learning procedures, including the manual analysis of a limited number of cases followed by the automatic generation of examples by analogy or the retrieval of analogous examples from existing corpora of data.

Strategies for clarification and confirmation actions can be learned using classification and regression trees.