Stochastic Modelling of Syntax and Semantics

Johannes Müller, Holger Stahl

Inst. for Human-Machine-Communication, Munich University of Technology
Arcisstrasse 21, D-80290 Munich, Germany
email: {mue,sta}@mmk.e-technik.tu-muenchen.de

Abstract: Within a 'top-down' approach for speech understanding (i.e. for converting a spoken utterance into its semantic representation), the semantic model generates semantic structures, which are semantic representations close to the word level. Corresponding to such a semantic structure, the syntactic model generates according word chains. These models can be integrated for language production or automatic translation within a restricted domain.
Keywords: speech understanding, language production, automatic translation, semantic model, syntactic model, spoken human-machine-dialogue

1. Definition of the semantic structure

In our approach [1], the *semantic structure* S (representing the semantic content) is a tree consisting of a finite number N of semantic units (we simply call them *semuns*) s_n:

$$S = \{s_1, s_2, \ldots, s_N\} \tag{1}$$

Each semun $s_n \in S$ with $1 \le n \le N$ is an (X+2)-tupel of a type $t[s_n]$, a value $v[s_n]$ and X successor-semuns $q_1[s_n], \ldots, q_X[s_n] \in \{s_2, \ldots, s_N, \text{blnk}\} \setminus \{s_n\}$:

$$s_n = (t[s_n], v[s_n], q_1[s_n], \ldots, q_X[s_n]) \quad , X \ge 1 \tag{2}$$

The semun s_1 is defined as the root of the semantic structure S. Every semun $s_2, \ldots, s_N$ is marked exactly once as a successor-semun. The special semun 'blnk' has the type $t[\text{blnk}] = \text{blnk}$, no value and no successor [2].

2. The semantic model

If statistical dependencies are assumed <u>only inside</u> of each semun, the a-priori probability $P(S)$ can be calculated as product of the following probabilities:

$$P(S) = f_{\text{root}} \cdot \prod_{n=1}^{N} (e_n \cdot f_n) \text{ , with ...} \tag{3}$$

- ... f_{root} denoting the a-priori-probability that the root semun s_1 is of the type $t[s_1]$:

$$f_{\text{root}} = P(t[s_1]) \tag{4}$$

- ... e_n denoting the conditional probability that the value $v[s_n]$ occurs with the type $t[s_n]$:

$$e_n = P(v[s_n] \,|\, t[s_n]) \tag{5}$$

- ... f_n denoting the conditional probability that the X successor semuns $q_1[s_n], \ldots, q_X[s_n]$ of the semun s_n with type $t[s_n]$ are of the types $t[q_1[s_n]], \ldots, t[q_X[s_n]]$:

$$f_n = P\left(t\left[q_1[s_n]\right], \ldots, t\left[q_X[s_n]\right] \,\middle|\, t[s_n] \right) \tag{6}$$

3. The syntactic model

We assume the following restrictions for the word chains $W = w_1 w_2 \ldots w_j \ldots w_J$ originated by the syntactic model to express a given semantic content S:

- Every word w_j in the word chain W can be assigned to exactly one semun $s_n \in S$.
- An unbroken part $w_i w_{i+1} \ldots w_j$ of W is originated for each semun s_n <u>and</u> all its successors.
- For each semun $s_n \in S$, one word w_{sig} is produced obligatorily, which depends on the <u>value</u> $v[s_n]$. Another word w_{insig} is produced optionally, which depends only on the <u>type</u> $t[s_n]$. We call these two words the 'significant word' w_{sig} and the 'insignificant word' w_{insig}.

Production Rules: The last assumption above implies to use a stochastic context-free grammar [3] as syntactic model. This grammar, denoted $G_{\text{syn}} = (V, T, \Sigma, P)$ contains the sets V, T and P of variables, terminals and production rules. The derivation always starts rewriting the start symbol $\Sigma \in V$ as variable $A(s_1)$, with s_1 marking the root of the semantic structure S:

$$P\Big(\Sigma \rightarrow A(s_1) \Big) \tag{7}$$

- For the case that $s_n \neq \text{blnk}$ the variable $A(s_n)$ produces a sequence of a variable $B(s_n)$, an optional variable $C(s_n)$ and X variables $A(q_1[s_n]), \ldots, A(q_X[s_n])$ for the successors of s_n:

$$P\Big(A(s_n) \rightarrow \underbrace{B(s_n), \overbrace{C(s_n)}^{\text{optional}}, A(q_1[s_n]) \ldots, A(q_X[s_n])}_{\text{in arbitrary order}} \Big| t[s_n] \Big) \tag{8}$$

- $A(\text{blnk})$ always produces the empty string ε:

$$P\Big(A(\text{blnk}) \rightarrow \varepsilon \Big) = 1 \tag{9}$$

The probability of eq. (8) is estimated by a transition network similar to an ergodic hidden markov model [4]. Such a *syntactic module* (SM) consists of X+4 states: $B(s_n)$, $C(s_n)$ and $A(q_1[s_n]), \ldots, A(q_X[s_n])$ represent the corresponding variables, 'strt' and 'end' stand for the entry and the exit of the SM. The path, i.e. the order of passing the states of the SM, is constrained by eq. (8).

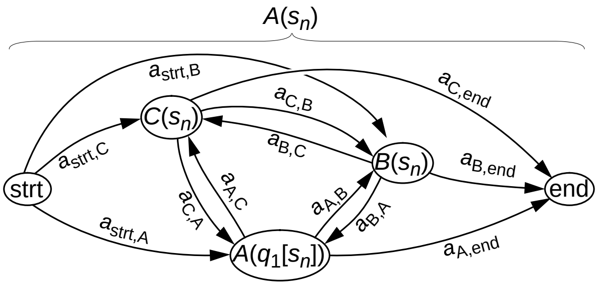

Fig. 1: SM for the semun s_n with X=1 successor

- $B(s_n)$ produces one significant word w_{sig} depending on the value $v[s_n]$:

$$P\Big(B(s_n) \rightarrow w_{\text{sig}} \Big| v[s_n] \Big) \tag{10}$$

- $C(s_n)$ produces one insignificant word w_{insig} depending on the type $t[s_n]$:

$$P\Big(C(s_n) \rightarrow w_{\text{insig}} \Big| t[s_n] \Big) \tag{11}$$

The probability $P(W|S)$ is calculated by maximizing the product of the probabilities concerning all the productions according to eq. (8), (10) and (11) required to derive '$\Sigma \Rightarrow W$' [5].

4. Language production and automatic translation

From a given semantic structure S, the most likely word chain W_g should be created according to the syntactic model:

$$W_g = \underset{W}{\operatorname{argmax}}\, P(W|S) \tag{12}$$

Using eq. (12), the choice and order of the produced words is mostly correct. Since the semantic structure does not contain any grammatical information, case, gender and number of many words in W_g are wrong. Therefore, the syntactic incorrect word chain has to be linguistically purified with an inflectional knowledge base and with the help of the given semantic structure. Combining understanding and production with models of two different languages, a very simple automatic translation system for a restricted domain can be realized.

References

[1] J. Müller, H. Stahl: *Ein Ansatz zum Verstehen natürlicher, gesprochener Sprache durch hierarchisch strukturierte Hidden-Markov-Modelle*, Proc. KONVENS 1994 (Vienna), pp. 260-269

[2] J. Müller, H. Stahl: *Die semantische Gliederung als adäquate semantische Repräsentationsebene für einen sprachverstehenden 'Grafikeditor'*, Proc. GLDV-Jahrestagung 1995 (Regensbg.)

[3] H. Ney: *Stochastic Grammars and Pattern Recognition*, NATO ASI, vol. F75 (1992), pp. 319-344

[4] L.R. Rabiner: *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*, Proc. IEEE, vol. 77, (1989), no. 2, pp. 257-286

[5] H. Stahl, J. Müller: *A Stochastic Grammar for Isolated Representation of Syntactic and Semantic Knowledge*, Proc. Eurospeech 1995 (Madrid), to be published