

# SCHEDULING OF PROCESSES IN A SPEECH UNDERSTANDING SYSTEM BASED ON APPROXIMATE REASONING

Renato De Mori

Lorenza Saitta

Istituto di Scienze dell'Informazione  
Università di Torino  
C.so Massimo D'Azeglio 42  
10125 - TORINO (Italy)

A method for scheduling knowledge source instantiations in a Speech Understanding System is presented. It is based on the evaluation of linguistic probabilities and on approximate reasoning.

## 1. INTRODUCTION

Searching for control strategies and methods for scheduling interpretation processes is among the most difficult tasks involved in designing Speech, Language and Image Understanding Systems. Remarkable efforts have been made in the area of Speech Understanding Systems (SUS), whose state of the art has been the object of a recent review by De Mori [1]. Sophisticated control strategies and scheduling policies are required when the world to be interpreted is very complex (for example for large protocols with big dictionaries in a SUS). For such cases, a paradigm of the type "hypothesize-and-test" and knowledge representation by rules of the type "precondition  $\rightarrow$  action" have been proposed [2]. The first attempts to design and evaluate control strategies in such a conceptual framework and in an environment where the input informations are corrupted and redundant appear in papers by Woods et al. [3], Lesser and Hayes-Roth [4] and in the Ph. D. Thesis by Paxton [5]. They mostly use numerical probabilities or heuristic scores to evaluate the hypotheses emitted by the system and to assign priorities to the processes that have to be executed to verify new hypotheses or theories. Valuable as they

are, all these approaches leave room for improvements.

On the other hand, control strategies of systems designed for understanding non corrupted input messages are mostly based on non-numerical rules. The content of this paper tries to establish a connection between the numeric and the non-numeric approach to the design of control strategies.

An attempt has been made for evaluating interpretations of the speech signal by a-posteriori probabilities, using the Bayes theorem [3]. But, in these evaluations, a high degree of imprecision is involved, owing to the approximations made in extracting acoustic features and in evaluating a-priori probabilities. Moreover, an increase of the complexity of an interpretation corresponds very often to having more heuristic and subjective simplification introduced in the computation of probabilities. Thus the a-posteriori probabilities obtained by the Bayes theorem are affected by a degree of imprecision. There is little hope that this imprecision can be removed by refinements involving a practically impossible experimental work. The most realistic position is that of accepting imprecisions and representing a-posteriori probabilities by linguistic variables. A mathematical framework in

which linguistic probabilities can be treated is the theory of fuzzy linguistic variables [6-9]. This theory not only allows one to deal with imprecise concepts, as the linguistic probabilities are, but also to combine probabilistic concepts with concepts of other types, using inference rules for developing an approximate reasoning. Furthermore, a-posteriori probabilities are not the only parameters on which scheduling of further interpretation actions have to be based. In fact, a control system should be designed on the basis of inference rules which take into account also considerations of action importance and suitability, that are not probabilistic in nature.

## 2. ASSIGNMENT OF LINGUISTIC PROBABILITIES

Let  $\Omega$  be a portion of the input data to be interpreted. System knowledge is shared among levels and each level may have a knowledge source associated with it. The lowest phonetic features of the system are "Vocalic", "Non vocalic" and "Sonorant", "Non-sonorant". The latter two features are applicable, in the system, only to consonants. The assignment of linguistic probabilities to interpretations of  $\Omega$  begins with the context-independent phonetic features, given the evidence of the acoustic features. For example, the assignment of the phonetic features "sonorant" and "non-sonorant" to  $\Omega$  is controlled by two fuzzy grammars  $G_s$  and  $G_n$ , described in details in [11]. Parsing the input string under the control of  $G_s$  and  $G_n$  leads to the evaluation of the compatibility values  $\mu_s(\Omega)$  and  $\mu_n(\Omega)$  between  $\Omega$  and the start symbols of  $G_s$  and  $G_n$  respectively. Rather than on the exact values of  $\mu_s(\Omega)$  and  $\mu_n(\Omega)$ , scheduling of subsequent processes depends on considerations that may be expressed by fuzzy propositions like the following: " $\mu_s(\Omega)$  and  $\mu_n(\Omega)$  are approximately equal" or " $\mu_s(\Omega)$  is much larger than  $\mu_n(\Omega)$ ", etc.. These fuzzy propositions lead to the definition of fuzzy sets on the interval  $[-1, +1]$  of the variable:

$$x = \mu_s(\Omega) - \mu_n(\Omega) \quad (1)$$

Let  $A \in \mathcal{A}$  be one of such fuzzy sets. Let  $J$  be the cardinality of  $\mathcal{A}$ ; the evidence of the features "sonorant" and "non-sonorant" in  $\Omega$  will be represented by the fuzzy linguistic variable  $A$  such that:

$$\mu_A(x(\Omega)) = \max_{j \in [1, J]} \mu_{A_j}(x(\Omega)) \quad (2)$$

Let  $H_s$  and  $H_n$  be the hypotheses:  $\Omega$  is a sonorant consonant and  $\Omega$  is a non-sonorant consonant respectively. By the Bayes theorem one gets:

$$\Pr(H_i | A) = \frac{\Pr(A | H_i) \Pr(H_i)}{\Pr(A)} \quad (i=s, n) \quad (3)$$

The evaluation of  $\Pr(H_i | A)$  is rather complex because  $A$  is a fuzzy event instead of a precisely defined event. Given the probability density  $p(x | H_i)$  and the definition of  $A$  on the interval  $[-1, +1]$  of the variable  $x$  and applying the extension principle and the definition of fuzzy integrals [10] it is possible to compute a possibility value  $\mu_{\Pr(H_i | A)}(z)$  for every value  $z \in [0, 1]$  of the probability  $\Pr(H_i | A)$ . This possibility|probability distribution can be compared with linguistic probabilities, defined as fuzzy sets on the interval  $[0, 1]$ . Such definitions can be found in [10]. Let  $\widetilde{P}(H_i | A)$  be the fuzzy set obtained with the (3) and  $\lambda_j$  a linguistic probability value: the value  $\lambda_j$  ( $j \in [1, J]$ ) that best approximates  $\widetilde{P}(H_i | A)$  is retained instead of an exact value of the left side probability in (3). The problem of linguistic approximation is treated in details by Bellman and Zadeh [13].

The computations mentioned so far are performed once forever and the results are represented by tables giving the best linguistic approximation for a probability|possibility distribution, given the evidence of the features in  $\Omega$  represented by fuzzy sets  $A_i \in \mathcal{A}$ .

## 3. PRIORITY ASSIGNMENT

A Speech Understanding System is organized by levels; let  $\langle l_1, \dots, l_j \rangle$  be the set of such levels. The knowledge of the system strategy may be represented in procedural form by rules of the type:

$\langle \text{precondition} \rightarrow \text{action} \rangle$ . For example, in the project referred to in this paper, the lexical knowledge is represented by frames, whose composition is described in another paper [14]. Each frame  $w_i$  contains a set of preconditions  $P_i$  and a set of actions  $A_i$ . A precondition  $p_{ik} \in P_i$ , that allows the action  $A_{ih} \in A_i$  to be requested, is expressed in terms of items corresponding to hypotheses already verified. Combining the probability values of these items, it is possible to evaluate the linguistic probability value of the precondition  $p_{ik}$ ; let  $\lambda_{ik}$  be such value. The assignment corresponds to the following proposition:

$$p_{ik} \text{ is } \lambda_{ik} \quad (4)$$

Priority assignment is performed by the understanding controller on the basis of two types of evaluation. The first refers to the possibility of the success of an action, given the probability of the preconditions. The second refers to the suitability of the action with respect to a knowledge like computational complexity and goal satisfaction. The latter evaluation is based on a proposition of the following type:

IF  $p_{ik}$  is likely THEN  $A_{ih}$  is suitable (5)  
where "suitable" is a fuzzy linguistic variable belonging to a dictionary  $\Sigma$ .

The (5) is an example of the rules of inference of the understanding controller. Such rules may be learned by experiments or they may be initially settled by the designer. During the understanding process, the controller will assign a linguistic value to the process executing the action  $A_{ih}$  according to the result of the following approximate reasoning:

$$p_{ik} \text{ is } \lambda_{ik} \quad (\text{premise}) \quad (6)$$

IF  $p_{ik}$  is likely THEN  $A_{ih}$  is suitable (implication)

Execution of  $A_{ih}$  is  $\sigma$ -suitable (conclusion)  
When more actions compete for execution with different degrees of local suitability, the priority have to be assigned to them on the basis of a few general inference rules of the control strategy. Such rules establish general degrees of importance of possible actions, given premises that are assumed to be true or false.

#### 4. AN EXAMPLE OF APPLICATION

We defined the four fuzzy sets:

$A_1 \triangleq "$   $\mu_S(\Omega)$  and  $\mu_N(\Omega)$  approximatively equal";  $A_2 \triangleq "$   $\mu_S(\Omega)$  slightly higher than  $\mu_N(\Omega)$ ";  $A_3 \triangleq "$   $\mu_N(\Omega)$  much higher than  $\mu_S(\Omega)$ ";  $A_4 \triangleq "$   $\mu_S(\Omega)$  much higher than  $\mu_N(\Omega)$ ". By the method described in [10]

we computed the possibility|probability distribution of the a-posteriori probabilities  $\Pr(H_i|A_j)$  ( $i=S,N$ ),  $j \in [1,4]$ , which we approximated by standard linguistic variables.

Table I and II show the results obtained. By using suitability values as priority values for the competing processes "sonorant" "non-sonorant" the right action is scheduled with higher priority in 84.6% of the cases and in the 14% of the cases the two actions have been scheduled with equal priority. Using a Bayes decision we obtained, in the most favorable case an error rate of 0.8% with 22% undecided cases.

| $A_j$                             | $\lambda(H_S A_j)$ | $\lambda(H_N A_j)$ |
|-----------------------------------|--------------------|--------------------|
| $A_1$<br>$-0.120 \leq x < 0.116$  | Uncertain          | Uncertain          |
| $A_2$<br>$0.116 \leq x < 0.369$   | Likely             | Unlikely           |
| $A_3$<br>$-1.000 \leq x < -0.120$ | Extremely unlikely | Extremely likely   |
| $A_4$<br>$0.369 \leq x < 1.000$   | Very likely        | Very unlikely      |

Table I - Definition of the fuzzy sets  $A_j$  and their a posteriori probability.

|   | $A_1$ | $A_2$ | $A_3$ | $A_4$ |
|---|-------|-------|-------|-------|
| S | 18    | 68    | 0     | 107   |
| N | 33    | 4     | 128   | 1     |

Table II - Assignment of priority according to the linguistic probability value.

## REFERENCES

- [1] R. De Mori: "Recent advances in automatic speech recognition", Invited paper Proc. 4-th Int. J. Conf. Pattern Recognition, Kyoto (1978)
- [2] D. Raj Reddy: "Hearsay speech understanding system: Summary of results of the five-year research effort at Carnegie-Mellon University", Dept. Comp. Sci. Carnegie-Mellon Univ., (1977)
- [3] W. Woods: "Shortfall and density scoring strategies for speech understanding control", Proc. 5-th Int. J. Conf. Artificial Intelligence, Cambridge, MA, 18-26 (1977)
- [4] F. Hayes-Roth, V. Lesser: "Focus of attention in a distributed logic speech understanding system", *ibidem*, 27-35 (1977)
- [5] H. Paxton: "A framework for speech understanding", Ph. D. Thesis, Stanford Univ., Ca (1977)
- [6] L. Zadeh: "PRUF - A meaning representation language for natural languages", *Int. J. Man-Machine Studies* 10, 395-460 (1978)
- [7] B. Gaines: "Foundations of fuzzy reasoning", *Int. J. Man-Machine Studies* 6, 623-668 (1976)
- [8] L. Zadeh: "A theory of approximate reasoning", Memo M77/58, Electronic Research Lab., Univ. of California, Berkeley, Ca. (1977)
- [9] E. Mamdani, S. Assilian: "An experiment in linguistic synthesis with a fuzzy logic controller", *Int. J. Man-Machine Studies* 7, 1-13 (1975)
- [10] D. Dubois, H. Prade: "Fuzzy Algebra, Analysis, Logics", School of Electrical Engineering, Purdue Lafayette, I.R. TR-EE 78-13 (1978)
- [11] R. De Mori, P. Laface: "Automatic detection of distinctive features in continuous speech", Proc. 3-rd Int. J. Conf. Pattern Recognition, Coronado, CA., 609-617 (1976)
- [12] L. Zadeh: "The concept of linguistic variable and its application to approximate reasoning, III" *Information Science* 9, 43-80 (1975)
- [13] R. Bellman, L. Zadeh: "Local and fuzzy logics" Memo ERL M-584, Univ. of California, Berkeley (77)
- [14] R. De Mori, L. Lesmo, M. Porcini: "The structure of a lexicon for a speech understanding system", Proc. IEEE Conf. ASSP, Washington D.C., 252-255 (1979)