EUROSPEECH '97
5th European Conference on Speech
Communication and Technology
Rhodes, Greece, September 22-25, 1997

AN ALTERNATIVE AND FLEXIBLE APPROACH IN ROBUST INFORMATION RETRIEVAL SYSTEMS

José Colás, Juan M. Montero, Javier Ferreiros, José M. Pardo

Grupo Tecnología del Habla - Dpto. Ingeniería Electrónica
E.T.S.I. Telecomunicación - Universidad Politécnica de Madrid
Ciudad Universitaria, s/n 28040 Madrid Spain

ABSTRACT

In this paper, we present a flexible architecture to implement a robust information retrieval system based on domain and linguistic modelling by means of a set of conceptual probabilistic and non-probabilistic grammars. It allows certain complexity in the functionality of the application, such as applying non-SQL functions to the results of SQL queries in order to retrieve information not explicitly included in the database, or translating certain natural spoken sentences that would produce difficult embedded queries.

INTRODUCTION

When processing spoken queries, some systems [5] [7] [8] classify them before being translated, using supervised stochastic classifiers based on neural networks, dynamic programming, etc. These classifiers need a great amount of labelled data and a set of complicated rules to fill the slots of the semantic frames recognised by the classification. This approach reduces flexibility when adding further functionality. Other approaches use integrated syntactic and semantic grammars to obtain structural information before applying the translation rules [1] [2].

Our proposed architecture, instead, makes use of several kinds of grammars: a probabilistic finite state network (PFSN) for concept decoding (CD), a context sensitive grammar (CSG_1) for conceptual mapping (CM), a semantic context-free grammar (SCFG) for structure analysis (SA) and another context sensitive grammar (CSG_2) for structure transformation (ST). This combination *allows answering a complex query through the structured execution of multiple simple queries and functions*. Compared to stochastic classifiers, this architecture is more flexible when adding new functionality without the need of collecting and labelling a large amount of additional data. Nevertheless, the PFSN must be retrained and new CM, SA and ST rules could have to be written.

ARCHITECTURE DESCRIPTION

A non-integrated two level approach has been implemented, including an acoustic decoder and an understanding module. The main characteristics of the proposed architecture are:

- Flexibility: different restricted domain tasks are easily modelled, because we have maintained independence between data and procedures.
- Robustness: some acoustic decoding error effects and coverage faults (lexical, syntactic or semantic) can be solved.
- Complexity: more powerful and complete applications can be developed because non-SQL functions can be applied to the results of SQL queries, and the structural linguistic analysis, through the SCFG, decomposes the sentence into multiple simple queries.
- Naturalness: in order to deal with some intrinsic characteristics of Spanish, no restrictions are imposed on phrase ordering; certain features of the verbs, such as tense, number, etc, are taken into account; and the

semantic role of some words (prepositions, conjunctions, adverbs, etc.) has been analysed.

Acoustic decoder

A One-Pass continuous speech recogniser based on semi-continuous HMM (SCHMM) is used [9].

The recogniser also allows N-gram stochastic grammars to guide the decoding process, although we have not used them so far.

Understanding module

It takes an input sentence from the acoustical decoder and produces an answer that, most of the times, is the result of the application of several available non-SQL functions on simple queries. In other cases it is the result of answering a simple query. The following modules compose this understanding system:

1. Probabilistic Conceptual Decoder

It is based on a dynamic programming algorithm and an ergodic regular grammar of concepts, related to the application domain, where each concept is modelled by a PFSN of semantic categories [3] [4] [6]. For the first version we have not trained these probabilities and they are uniform. Robustness and coverage are increased by means of a special “garbage” category.

Stochastic concept automata mainly model the structural information of each concept, but an inter-concept grammar is also trained from labelled data.

Before CD, we use a semantically labelled dictionary to look up each word. Due to ambiguity, the application of the labeller on a sentence produces an oriented category graph. A dynamic programming algorithm processes this graph to get the most probable sequence of concepts and categories, solving the ambiguity problem.

2. Conceptual Mapper

As the CD is unable to model non-contiguous inter-concept relationships, some specific concepts such as numbers or dates, are not correctly labelled. Considering the context, each of these ambiguous concepts can be correctly mapped or assigned to the right concept using a set of mapping and merging CSG$_1$ rules (for instance, a date can be an attribute of more than one entity in the Entity-Relationship scheme). As we can see in the example bellow, due to the presence of a Report concept, Date can be mapped onto Date-Report:

Was	last casualty report	regarding Zeus	in November	?
	Report	Ship	Date=Report Date	

Some contiguous and non-contiguous concepts can be merged. In the example bellow, considering the Length concept in the sentence, Number concept can be incorporated to Length (attribute-value pair):

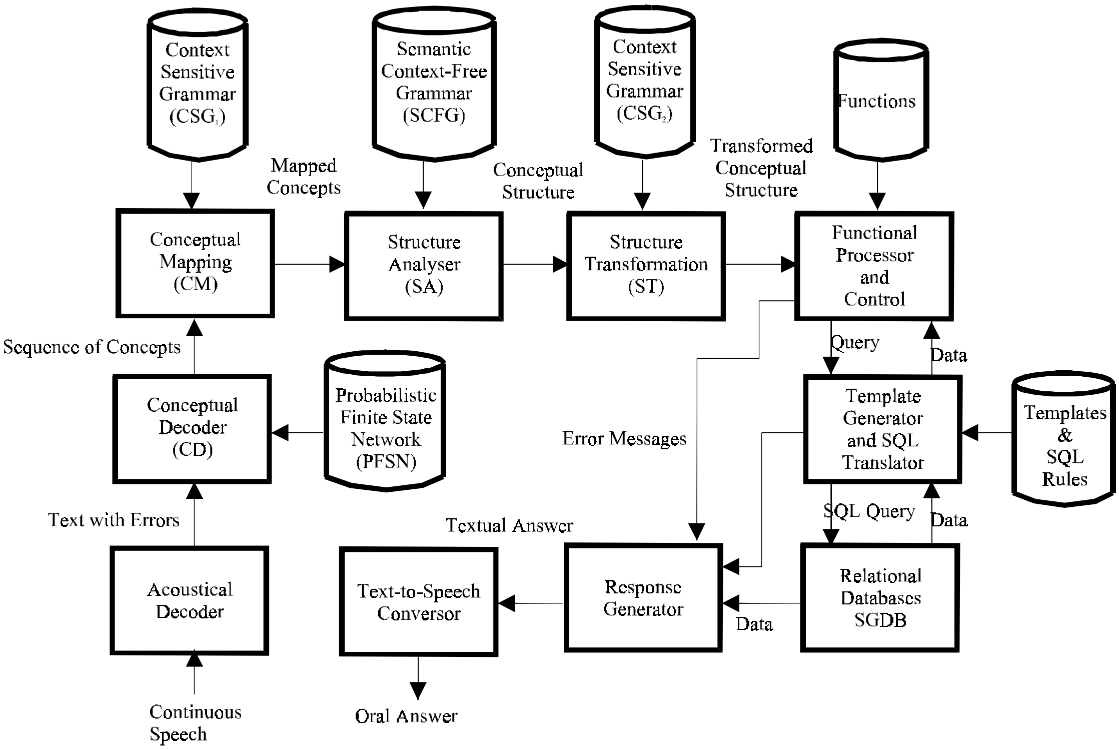
Is	the length	of the fastest	ship	5 meters or more	?
	Length	Speed	Ship	Number=Length	

3. Structure Analyser

Our proposed alternative to stochastic classifiers is based on a SCFG concept parser. The structure obtained this way, allows determining the simple queries and the functions (distance, comparison, etc.) that process them. In order to get a reduced and more general set of grammar rules, a concept taxonomy (verb, function, entity, attribute) is used. To design these rules, the application functionality and certain general linguistic phenomena (co-ordination, relative clauses, negation, etc.) have been taken into account.

4. Structure Transformer

If the SCFG extracts conceptual structure, the ST rules complete subqueries, solving some problems such as concept ellipsis, relative clauses, and so on.



5. Functional Processor and Control

The role of this module is to process the result of the queries using, if required by the answering strategy, the non-SQL function. Additional functions can be easily incorporated if necessary.

6. Template Generator and SQL Translator

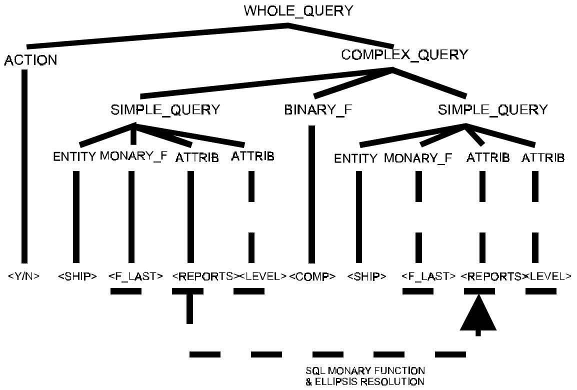
The simple queries detected and completed by previous modules are translated to SQL in two steps: templates are filled with key information from the sentence and, then, these templates are translated into SQL queries by means of a set of rules that include database information.

Examples

One example of complex query processing could be: “Is England’s latest casualty report rated worse than America’s?” It is decomposed into two simple queries (“England’s latest casualty report” and “America’s latest casualty report”) which are the arguments of the comparative function “worse than”. These simple queries are easily translated into SQL after the ST module solves the concept ellipsis problem in the second query. In this example, we include the following illustration that shows the concept structure and the transformation process. The output of CD and CM is:

<Y/N> : is
<SHIP> : England’s
<F_LAST> : latest
<REPORTS> : casualty report rated
<F_COMP> : worse than
<SHIP> : America’s

where concepts appear between brackets. As it is shown in the illustration above, each concept is assigned to one of the classes in the taxonomy.


WHOLE_QUERY
ACTION
COMPLEX_QUERY
SIMPLE_QUERY
BINARY_F
SIMPLE_QUERY
ENTITY MONARY_F ATTRIB ATTRIB
ENTITY MONARY_F ATTRIB ATTRIB
<Y/N> <SHIP> <F_LAST> <REPORTS><LEVEL><COMP> <SHIP> <F_LAST> <REPORTS><LEVEL>
SQL MONARY FUNCTION & ELLIPSIS RESOLUTION

Another example: “Is Neptune closer to Titanic than to Zeus?” It is decomposed into three simple SQL queries, which obtain the geographic position of each ship. With this information two distances are calculated and compared by the control module.

TASK DESCRIPTION AND EVALUATION

An Information Retrieval System has been developed in order to get navy information, allowing not only simple queries but also more complex ones, including non-SQL functions and multi-query questions.

The application has a restricted semantic domain. The vocabulary is about 1100 words. The first version of the system has been implemented using 600 Spanish sentences for training (both text and speech) and 400 for testing. 4 speakers (2 male and 2 female) have uttered the speech. To further evaluate the system we have collected up to 3000 sentences (plain text) from new users and several speakers will speak some of them, after the necessary review.

Preliminary results (on only text corpora), comparing the system output to the reference queries, show that there are *92 % correctly generated queries for the training text corpus, and 89 % for the testing one*. For the first prototype, we are not using grammatical restrictions (up to 85.2% word accuracy).

PLANNED WORK

We are currently evaluating the spoken corpora (4 speakers) and the new collected sentences, to obtain conclusions about the limitations and advantages of using robust concept decoding (with garbage semantic categories) in restricted domain applications

CONCLUSIONS

We have presented a flexible bottom-up approach to deal with complex understanding tasks. To attain this, the architecture integrates robust conceptual decoding and semantic structure analysis to answer a complex query through the structured execution of multiple simple queries and functions. The results that we have obtained so far are encouraging.

ACKNOWLEDGEMENTS

This work has been supported by CICYT MEC (Ministry of Education and Science) TIC 94-0119 *Sicovox II Continuous Speech Recognition Systems*. The authors wish to thank Juan Lestani, Javier Macías, Juan R. González, and, specially, Jorge de Teresa and Samir Kilany for their contribution to the development of this architecture.

REFERENCES

[1] Nagai Akito, Ishikawa Yasushi, Nakajima Kunio, "Integration of Concept-Driven Semantic Interpretation with Speech Recognition", Proc. ICASSP-96,pp.431,434. Atlanta, USA.

[2] S.K. Bennacef, H. Bonneau-Maynard, J.L. Gauvian, L. Lamel, W. Minker, "A Spoken Language System For Information Retrieval", Proc. ICSLP-94, pp. S22-8.1, 8.4, Yokohama, JAPAN

[3] Egidio Giachin, Paolo Baggia, Giorgio Micca, "Language Models for Spontaneous Sppech Recognition: A Bootstrap Method for Learning Phrase Bigrams".Proc. ICSLP-94, pp. 16-17.1, 16-17.4.

[4] Enrique Vidal, Roberto Pieraccini, Esther Levin, "Learning Associations Between Grammars: A New Approach to Natural Language Understanding", Proc. EUROSPEECH-93, pp.1187-1190.

[5] P. Bianchini, P. Ferragina, M. Notturno Granieri, L. Tarricone, "New Techniques for Speech Understanding". Proc. ICASSP-93, pp. II-127, II-130

[6] Ying Cheng, Yves Normandin, Paul Fortier, "Integration of Neural Networks and Robust Parsers in Natural Language Understanding". Proc. Eurospeech-93, pp.1311-1314.

[7] Egidio P. Giachin, "Automatic Training of Stochastic Finite-State Language Models for Speech Understanding". Proc. ICASSP-92, pp. I-173,I-176.

[8] Roberto Pieraccini, Evelyne Tzoukermann, Zakhar Gorelov, Jean-Luc Gauvain, Esther Levin, Chin-Hui Lee, Jay G. Wilpon, "A Speech Understanding System Based on Statistical Representation of Semantics". Proc. ICASSP-92, pp.I-193, I-196

[9] J. Ferreiros, "Contribution to Markov Models training methods for continuous speech recognition". Ph D. Thesis. Universidad Politécnica de Madrid, 1996.