THE Ikaros Continuous Speech Understanding System: First Demonstrator

F Monnet, S Jousset, A Demour, P Richard

Laboratoires de Marcoussis
Division informatique
Route de Nozay, 91460 Marcoussis - FRANCE

ABSTRACT

The aim of the Ikaros (Esprit project 954) is to evaluate the interest of Artificial Intelligence techniques in speech understanding. This is achieved by designing and implementing a system whose features include continuous speech, multispeaker, vocabulary of up to one thousand words, natural-type spoken language, multilingualism and railway timetable enquiry as an application.

A set of knowledge sources (KSs) for various levels (acoustic-phonetic, phonological, syntactic semantic and pragmatic) and an architecture that permits the closest cooperation between these KSs have been developed.

Architecture

The Ikaros architecture is based on the blackboard model. Like the BBl system, Ikaros has two blackboards: the domain blackboard contains hypotheses ranging from sub-word units to semantic structures; the control blackboard contains heuristics that drive the solving process. It has been implemented with an object-orientated language and a general purpose inference engine. The latter point has permitted an enlightenment of the blackboard model.

Control

To determine the best action to be performed at each step, we have relied on fuzzy-set and possibility theory. Indeed, the agenda is viewed as the fuzzy-set of the possible actions that satisfy the current heuristic. Furthermore, complex heuristics can be expressed with elementary fuzzy-sets and fuzzy-operations.

Speech Decoding KS

The input of Ikaros is a lattice of sub-word units with confidence scores. For French, it is provided by the decoding system of Esprit SPIN project. This system is based on Hidden Markov Model for training and Dynamic Time Warping procedures for recognition.

Higher-level KS

Ikaros can generate words from reliable anchors in the lattice of sub-word units, from boundaries of already generated words or from lexical structures deduced from the syntactic analysis. This analysis is performed by a chart-parser. The grammar is represented with the lexical-functional grammar (LFG) formalism. As for semantics, a simplified version of the Halvorson formalism has been used.

Reference

F Monnet, S Jousset, A Demour, P Richard "The Ikaros speech understanding system: presentation and first achievements "European Technical Week 88."