www.inesc-id.pt

technology
from seed

INESC-ID
Instituto de Engenharia de Sistemas e Computadores Investigação e Desenvolvimento em Lisboa

Rua Alves Redol, 9
1000-029 Lisboa Portugal
Tel. +351.213100300
Fax: +351.213145843
Email: info@inesc-id.pt

L^{2}F – Spoken Language Systems Laboratory

Spoken / Multimodal Dialogue Systems

Human-computer interface is an area where speech, integrated into a multimodal structure, creates a natural and easy way to establish communication.

Goal. Building a platform for research and development of spoken dialogue systems integrated into a multimodal user interface.

Description. This is one of our mainstream lines of activity. We have been developing competences in several of the core technologies for Spoken Dialogue Systems, as Automatic Speech Recognition (AUDIMUS), Text-to-Speech (DIXI+), Synthetic Talking Face (FACE) and Dialog Management (DM). These technologies have been integrated into an Embodied Conversational Agent (ECA) with a unique and sophisticated user interface.

Applications. The spoken dialogue platform has been applied in different scenarios, illustrating the flexibility of our modular system:

(i) home environment: control of a wide range of home devices, such as, lights, air conditioning, hi-fi, and TV, based on X10 and IRDA protocols. We can extend the application to include any infra-red controllable device or any device whose control functions may be programmed by the X10 protocol;

(ii) database access: allowing the user to inquire databases via voice. We have developed prototypes to request weather information, cinema schedules and bus trip information. This type of application can easily be extended to other domains;

(iii) email access: through telephone queries based on a dynamic VoiceXML.

The main areas of activity of L^{2}F are: semantic processing of multimedia contents, spoken/multimodal dialogue systems and speech-to-speech machine translation. Complementing these mainstream areas we are actively involved in many core areas of spoken language research and development.

More information is available by email to info@l2f.inesc-id.pt or directly from the website http://www.l2f.inesc-id.pt/.

inesc id lisboa

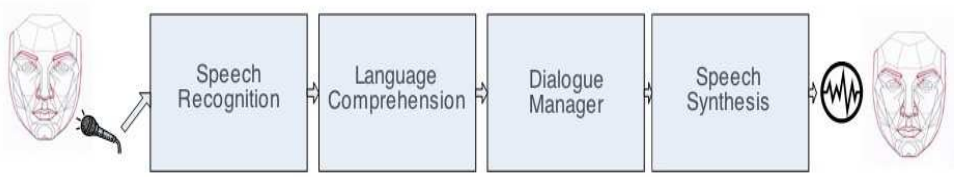
L²F: Spoken / Multimodal Dialogue Systems

Team

Nuno J. Mamede, PhD
João P. Neto, PhD
Luís Oliveira, PhD
Isabel Trancoso, PhD
Luísa Coheur, PhD
David Matos, PhD
Diamantino Caseiro, PhD

Joana Paulo, PhD Student
Porfírio Filipe, PhD Student
Márcio Mourão (got the M.Sc on 2005)
Renato Cassaca, Eng.
Márcio Viveiros, Eng.
Filipe Martins, Student
Pedro Arez, Student

Main Modules



The central component of the spoken dialogue system architecture is the *Dialogue Manager* which integrates several modules:

The Interpretation Manager (IM) receives a set of speech acts from the Language Comprehension module and generates the correspondent interpretations and discourse obligations. Interpretations are frame instantiations that represent possible combinations of speech acts and the meaning associated to each object it contains. To select the most promising interpretation two scores are computed. The recognition score to evaluate the rule requirements already accomplished, and the answer score, a measure of the consistency of the data already provided by the user.

The Discourse Context (DC) manages all knowledge about the discourse, including the discourse stack, turn-taking information, and discourse obligations.

The Behavioral Agent (BA) enables the system to be mixed-initiative: regardless of what the user says, the BA has its own priorities and intentions. When a new speech act includes objects belonging to a domain that is not being considered, the BA assumes the user wants to introduce a new dialog topic: the old topic is put on hold, and priority is given to the new topic. Whenever the system recognizes that the user is changing domains, it first verifies if some previous conversation has already taken place.

The Generation Manager (GM) receives discourse obligations from the BA, and transforms them into text, using template files. The GM uses another template file to produce questions that are not domain specific. For example, domain disambiguation questions, used to decide proceeding a dialogue between two or more distinct domains, or to clarify questions, are defined in this file.

The *Service Manager* is the interface between the spoken dialogue platform and a set of heterogeneous domain devices. Each device description is composed by slots and rules. Slots define domain data relationships, and rules define the system behavior. A rule or service represents an user possible action.

More Information

By email to info@l2f.inesc-id.pt or
directly from the website http://www.l2f.inesc-id.pt/.