Ricardo de Córdoba Herralde

Person-machine Dialogue Systems

2nd semester
Person-machine Dialogue Systems

Number of credits: 3 theory + 1 practical

Semester: 2

Type: optional

Objectives
This course is devoted to the study of the various modules involved in an interaction system or of human-machine dialog. Starting with an overview on dialogue systems and their problems, to go on to address the key modules that make it up, describing its operation, the research alternatives adopted to achieve optimal system performance and the problems of each.

Each of the modules will be started from a basic level and go up to describing the most advanced algorithms and techniques with which we will get the most robust and reliable systems.

The course is based on lectures to acquire the desired skills, but it also includes a set of application case studies, specially selected, to be solved in common and that allow the application skills to be acquired.

This will enhance the interaction with the students so they can apply the acquired knowledge in a final project of the subject.

Program
The course will be cover the following topics:

1. Dialogue system architecture

2. Fundamentals of production and Speech perception
3. Synthesis and generation of response
4. Speech recognition: parameterization and quantification
5. Speech recognition: hidden Markov models
6. Continuous speech recognition
7. Adaptation
8. Language models
9. Speaker identification and language identification
10. Speech understanding and translation
11. Synthesis and recognition of emotions and multimodal interaction
12. HTS synthesis
13. Design methodologies and user modeling
14. Evaluation of dialogue systems

**Teachers**
Coordinator: Ricardo de Córdoba Herralde

Teachers: Rubén San Segundo Hernández, Roberto Barra Chicote

**Teaching Methodology**
The subject will be taught by:

- Lectures
- We carried out a project related to any of the modules described in the course

**Evaluation**
Students complete the course with a final project of individual character to be presented publicly in English as part of activities to acquire transversal competences of documentation, communication and publication.

The report must be presented in the typical format for IEEE conference papers (http://www.ieee.org/conferences_events/conferences/publishing/templates.html) with aim of encouraging the student, not only through the reading and interpretation of scientific and technical documents, but also its correct wording.
The final project must be eminently practical, and in it should be applied some of the techniques described in the course, preferably, a problem that may be related to research or professional activity of the student.

The written report will be the 70% of the final grade. However, the teacher also will observe the ability of students to communicate effectively and concisely the technical information, knowledge, justifications, etc. and to answer the questions he may pose them. The oral presentation will be the 30% of the grade.

**Comunicación con el profesorado**

- Ricardo de Córdoba Herralde, despacho B-108, [cordoba@die.upm.es](mailto:cordoba@die.upm.es), ext 4209
- Rubén San Segundo Hernández, despacho B-109, [lapiz@die.upm.es](mailto:lapiz@die.upm.es), ext. 4228
- Roberto Barra Chicote, despacho B-112, [barra@die.upm.es](mailto:barra@die.upm.es), ext 4254

**Bibliografía**

All material is made accessible through the Web page of the course well in advance of the delivery of the corresponding lectures. In this way, students have at all times appropriate material for easy tracking of classes.

We recommend the following general bibliography:


For parameterization:


For Vector Quantization:


For Markov Models:


Adaptation of HMMs for:


- The Generation and Use of Regression Class Trees for MLLR Adaptation. Ga-les, MJF, University of Cambridge, August 1996

- Maximum Likelihood Linear Transformations for HMM-based speech recogni-tion. Wales, MJF, Computer Speech and Language, 12, pp. 75-98, 1998


Structural


To Identify speakers:

For language recognition:


For Speech Recognition connected:

The Application of Dynamic Programming to Connected Speech Recognition


To Architectures for recognition:


Spoken Language Processing. Xuedong Huang, Alex Acero and Hsiao-Wuen Hon Prentice Hall PTR. 2001

For Models Language:


Improved Backing off for n-gram Language Modeling. R Kneser and H Ney. ICASSP 1995


For dialogue management:


For evaluation of dialogue systems:


MATERIAL RESOURCES AVAILABLE

The course itself does not currently have a dedicated laboratory equipped with work places in which to implement the techniques introduced. But it does provide trainees with suitable information on possible SW resources that may be available online (open-source software
licensed under GNU-GPL). Some examples of tools related to the techniques described in the subject might be:

- Praat (http://www.praat.org) tool developed by Paul Boersma and David Weenink of the University of Amsterdam, which allows the extraction of acoustic features.

- HTK (http://htk.eng.cam.ac.uk/) is a toolkit for estimating and using hidden Markov models.