DT2112 Speech Technology 7.5 credits

Talteknologi

This is a translation of the Swedish, legally binding, course syllabus.

If the course is discontinued, students may request to be examined during the following two academic years

Establishment

Course syllabus for DT2112 valid from Spring 2009

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Electrical Engineering

Specific prerequisites

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

The participants shall after completion of the course be able to:

- Give short descriptions of speech from the acoustic, phonetic, and linguistic perspectives for use in speech technology applications.
- Summarize human hearing and speech perception and describe how hearing impairments influence speech perception.
- Explain how computers recognize speech and speakers and describe common methods to do so, such as HMMs and neural networks.
- Give an overview of different methods used to produce speech with computers and how a computer-animated face may be used to improve speech perception.
- Exemplify speech-driven dialogue systems and choose type of system based on the area of application.
- Give an account of available state-of-the art speech technology and its applications.
- Summarize the current research areas in speech technology and how scientific results may be applied in e.g., mobile systems and IT.

Course contents

The course in Speech Technology discusses processes and systems for information transmission using speech as its medium and also gives basic knowledge of speech, language and hearing. The course contains sections discussing:

Overview of linguistic theory and phonetics.

Basics of physiology and acoustics of speech as a base for speech technology models.

Measuring techniques and signal processing in speech analysis.

Physiology of hearing, psychoacoustics and speech perception with applications in speech understanding systems.

Methods for automatic speaker verification.

Evaluation of speech communication systems.

Studies and experiments with text-to-speech and speech-to-text in systems for human-computer interaction, especially multimodal dialogue systems.

Course literature

Holmes J and Holmes W (2002): Speech Synthesis and Recognition, 2nd ed., Taylor and Francis, London .

Complementary off-prints are distributed during the course.

Examination

- INL1 - Assignments, 1.5 credits, grading scale: P, F
- LAB1 - Laboratory Work, 1.5 credits, grading scale: P, F
- TEN1 - Examination, 4.5 credits, grading scale: A, B, C, D, E, FX, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

Other requirements for final grade

One written examination (4,5 university credits.), a laboratory course (1,5 cr.) and an essay (1,5 university credits.)

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.