Module Leader
Prof P Woodland

Lecturers
Prof P Woodland and Prof B Byrne

Timing and Structure
Lent term. 14 lectures + 2 examples classes. Assessment: 100% exam.

Prerequisites
3F1 and 3F3 useful

Aims
The aims of the course are to:

- introduce major techniques for recognising and synthesising speech signals, and for statistical machine translation.

Objectives
As specific objectives, by the end of the course students should be able to:

- understand techniques for speech analysis, recognition and synthesis.
- understand techniques for statistical machine translation.
- apply these techniques in order to build simple speech recognition, synthesis and machine translation systems.
- show awareness of the current state-of-the-art in speech recognition and synthesis and in machine translation technology.

Syllabus

Lecture 1: Overview/Introduction

- Speech production mechanisms, types of speech sound, source-filter model, applications of speech and text processing

Lectures 2-3: Acoustic Analysis

- FFT based methods, Mel scale, cepstral analysis, all-pole filter models, calculation of LP coefficients. Formant and voicing analysis. Front-end analysis for speech recognition.
Lectures 4-5: *ASR Introduction and Isolated Word Recognition*


Lecture 6: *Sub-word Acoustic Models*

- Large vocabulary speech recognition, limitations of word models, context dependent phones, parameter tying.

Lecture 7: *Language Models*

- Perplexity, N-gram language models, discounting, interpolation.

Lecture 8: *ASR Search Issues*

- Continuous speech recognition. Pruning. Integrating context dependent HMMs and N-gram language models.

Lectures 9-10: *Weighted Finite State Transducers for Speech and Language Processing*

- Efficient realization of probabilistic models for sequence processing. Transduction, composition, determinization, minimum-cost search. WFSTs in ASR search and other language processing applications.

Lecture 11: *Introduction to Statistical Machine Translation*

- Statistical pattern processing approaches to translation. Automatic evaluation of translation quality.

Lecture 12: *SMT - Alignment*

- Parallel text as training data. Models of word and phrase alignment in translation. Model estimation procedures.

Lecture 13: *SMT - Translation*

- Phrase-based translation systems. Implementation via WFSTs.

Lecture 14: *Text-to-Speech Synthesis*

- Introduction to TTS. Data-driven synthesis and Hidden Markov Model approaches to TTS.

**Booklists**

Please see the [Booklist for Group F Courses](http://teaching.eng.cam.ac.uk) [2] for references for this module.

**Assessment**

Please refer to [Form & conduct of the examinations](http://teaching.eng.cam.ac.uk) [3].
UK-SPEC

The UK Standard for Professional Engineering Competence (UK-SPEC) [4] describes the requirements that have to be met in order to become a Chartered Engineer, and gives examples of ways of doing this.

UK-SPEC is published by the Engineering Council on behalf of the UK engineering profession. The standard has been developed, and is regularly updated, by panels representing professional engineering institutions, employers and engineering educators. Of particular relevance here is the ‘Accreditation of Higher Education Programmes’ (AHEP) document [5] which sets out the standard for degree accreditation.

The Output Standards Matrices [6] indicate where each of the Output Criteria as specified in the AHEP 3rd edition document is addressed within the Engineering and Manufacturing Engineering Triposes.

Last modified: 02/10/2014 16:50

Source URL (modified on 02-10-14): http://teaching.eng.cam.ac.uk/content/engineering-tripos-part-iib-4f11-speech-language-processing-2014-15

Links
[1] mailto:pw117@cam.ac.uk