| Short Title: | Interactive Speech Processing APPROVED |
|---|---|
| Full Title: | Interactive Speech Processing |

| Module Code: | MHLT H6016 |
|---|---|

| ECTS credits: | 10 |
|---|---|

| NFQ Level: | 9 |
|---|---|

| Module Delivered in | 1 programme(s) |
|---|---|

| Module Contributor: | Arnold Hensman |
|---|---|

| Module Description: | This module will introduce learners to automated human speech processing techniques that can be applied to a variety of applications such as speaker recognition, speech to text and text to speech systems. Structure of speech signals will be analysed and the challenges of the technology addressed. |
|---|---|

| Learning Outcomes: |
|---|
| *On successful completion of this module the learner will be able to* |

1. Analyse human speech signals in terms of component waveform frequencies
2. Create a synthesised human speech signal
3. Demonstrate the challenges and best practices in speech to text systems
4. Demonstrate how speech processing may be applied to biometric identification
5. Explain how prosody in human speech can be used as an additional element for more efficient automation of signals

## Module Content & Assessment

| Indicative Content |
| --- |
| **Speech Signals**<br>Analysis of voiceprints, waveforms and spectra in order to isolate speech phonemes. |
| **Speech Processing and Synthesis**<br>Use speech formants to create artificial speech and enhance naturalness of resulting outputs. Algorithmic aspects of speech recognition systems including pattern classification, search algorithms, and language modelling techniques. |
| **Speech to text**<br>Creation of speech to text systems and analysing the current obstacles and shortcomings to making applications in the science. Compares and contrasts the various approaches to speech recognition, and describe advanced techniques used for acoustic-phonetic modelling, robust speech recognition, speaker adaptation, processing paralinguistic information, speech understanding, and multimodal processing. |
| **Speech as Biometrics**<br>Use systems that can identify a speaker |
| **Prosody in speech**<br>Augmenting the standard approach of short duration phonemes with the application of speech prosody to isolate emotional markers in the speech signal, |

| Indicative Assessment Breakdown | % |
| --- | --- |
| Course Work Assessment % | 100.00% |

**Course Work Assessment %**

| Assessment Type | Assessment Description | Outcome addressed | % of total | Assessment Date |
| --- | --- | --- | --- | --- |
| Lab work | Create systems that breakdown a speech waveform or spectra into component frequencies for analysis towards a variety of applications, including Text to Speech, Speech to Text, Speaker Identification and Speech Prosody. | 1,2,3 | 30.00 | n/a |
| Case study | Use Speech Analysis Techniques to implement a basic application to demonstrate Speaker Identification and Speech Prosody. | 4,5 | 10.00 | n/a |
| Project | Develop a speech processing system for a chosen application such as speech localisation, speech to text or a biometric system. Produce a report indicating the current state of research in the area as well as effectiveness of system created. | 2,3,5 | 40.00 | n/a |
| In-class test | Ensure learner understanding of technical terminology and appropriate applications within software and hardware. | 3,4,5 | 20.00 | n/a |

| No Final Exam Assessment % |
| --- |

**Indicative Reassessment Requirement**

| **Coursework Only**<br>*This module is reassessed solely on the basis of re-submitted coursework. There is no repeat written examination.* |
| --- |
| **Reassessment Description**<br>Coursework re-submission |

**ITB reserves the right to alter the nature and timings of assessment**

## Indicative Module Workload & Resources

| Indicative Workload: Full Time | |
| --- | --- |
| **Frequency** | **Indicative Average Weekly Learner Workload** |
| Every Week | 2.00 |
| Every Week | 6.00 |
| Every Week | 6.00 |

| Resources |
| --- |
| *Recommended Book Resources* |
| Daniel Jurafsky, James H. Martin 2009, *Speech and language processing*, 7th Ed., Pearson Prentice Hall Upper Saddle River, N.J. [ISBN: 0131873210] |
| Lawrence Rabiner, Ronald Schafer 2014, *Theory and Applications of Digital Speech Processing*, 2 Ed., Prentice Hall [ISBN: 0136034284] |
| *Supplementary Book Resources* |
| Frederick Jelinek 1997, *Statistical methods for speech recognition*, MIT Press Cambridge, Mass. [ISBN: 0262100665] |
| Steven Bird, Ewan Klein, and Edward Loper 2009, *Natural language processing with Python*, O'Reilly Sebastopol, Calif. [ISBN: 0596516495] |
| *This module does not have any article/paper resources* |
| *This module does not have any other resources* |

## Module Delivered in

| Programme Code | Programme | Semester | Delivery |
| --- | --- | --- | --- |
| BN_KMHLT_R | [Master of Science in Computing in Multimodal Human Language Technology](#) | 2 | Elective |