# Speech Signal Processing
# EC5.408

**Assignment 1**

Aug 8, 2023

## Guidelines

- Do not copy or plagiarise. If you're caught for plagiarism, the penalty will range from **zero** in the assignment to **F** grade in the course.
- Always cite your sources (be it images, papers or existing libraries).
- Mention clearly if any assumptions are being considered.
- Only MATLAB or Python can be used for the coding part.

## Submission Format

Make a directory using the naming format **SSP_A1_RollNo**. The submission might include codes **(.py/.m)** to answer the coding problems, reports **(.pdf)** to answer the theory questions or notebooks **(.ipynb)** to answer both coding and theory questions together. Place the files in their respective folders and zip the main directory using the naming format **SSP_A1_RollNo.zip** and upload this zip file to Moodle.

This is how the final directory structure might look like

```
SSP_A1_RollNo
├── codes
│   ├── code_1.py
│   ├── code_2.ipynb
│   ├── code_3.m
│   └── code_4.py
├── reports
│   ├── report_1.pdf
│   ├── report_3.pdf
│   └── report_4.pdf
└── wavs
    ├── audio_1.wav
    └── audio_3.wav
```

## Resources

- For this assignment you might use **Audacity** or **Wavesurfer** software (will be mentioned)
- Install Audacity: `https://www.audacityteam.org/download/`
- Install Wavesurfer: `https://sourceforge.net/projects/wavesurfer/`

# Questions

**[Maximum marks: 20]**

1. Explain briefly about the following **[6]**
   (a) What is the difference between **sampling** and **quantization**? [1]
   (b) When converting a speech signal into a sequence of frames, why is the frame shift usually smaller than the frame duration? [1]
   (c) Why **Hamming** or **Hann** window is preferred over a rectangular window in speech processing? [1]
   (d) What are the assumptions to be considered while analyzing a speech signal? [2]
   (e) How would you make an impulse train with fundamental frequency F from sine waves? [1]
2. Explain briefly about the following **[4]**
   (a) Co-articulation [1]
   (b) Epochs [1]
   (c) Pitch [1]
   (d) Formants [1]
3. "Female pitch is more when compared to male pitch". **True** or **False**. Justify the statement with proper explanation. **[2]**
4. Record a sentence which should be as "My name is <your name>" **[4]**
   (a) Plot the waveform [1]
   (b) Identify and mark the voiced, unvoiced, silence and plosive regions (plot it over the waveform) [1]
   (c) Acoustic-phonetic description of the regions (MOA and POA) [1]
   (d) Plot the Mel spectrogram and identify the formants [1]

   Use Audacity or Wavesurfer for this question. You are expected to submit a **wav** file, along with the annotated transcription on top of the waveform.
5. Write a code to plot the following for a given audio file **[4]**
   (a) short-time energy [2]
   (b) short-time zero-crossing [2]

   Choose appropriate window sizes, window shifts, and window for the analysis. Explain your choice of these parameters.

# Appendix

In the Acoustic-phonetic description, the following format is expected:

For the example of **kitAb (/k/,/i/,/t/,/A/,/b/)**

It is unvoiced unaspirated velar stop followed by front vowel followed by unvoiced unaspirated dental stop followed by middle vowel followed by voiced unaspirated bilabial stop