University of Sheffield

COM3502-4502-6502
Speech Processing

Main Programming Assignment

John Ayad

Matthew Kinton

Department of Computer Science

December 6, 2018

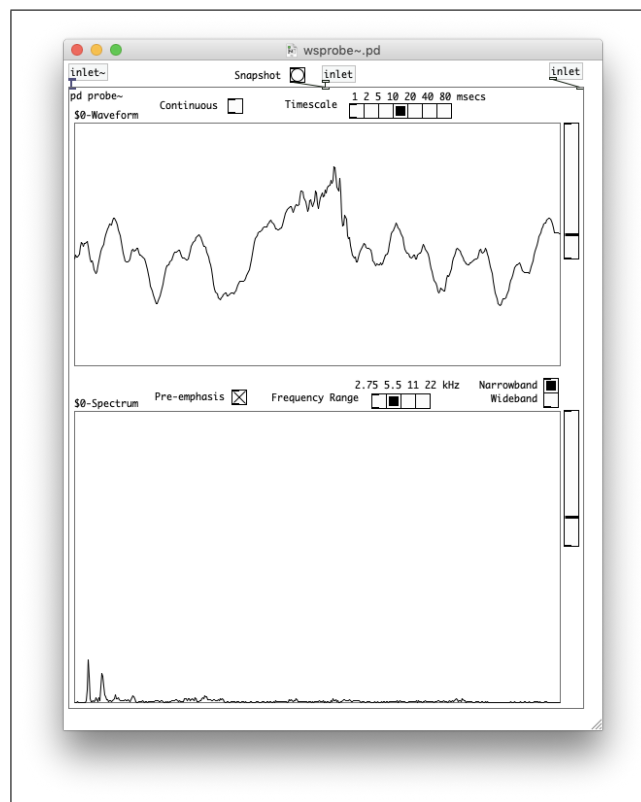**QUESTION 1** *(worth up to 5 marks)*
Provide a screenshot of [wsprobe∼] for a typical voiced sound, and explain the features in the waveform and spectrum that distinguish it from an unvoiced sound. *Hint: use the 'snapshot' feature in [wsprobe∼] to obtain a static display.*

Voiced:
- Energy is present low frequencies (spectrum)
- Pulses are visible due to the glottal closure (Waveform) (Because of the vocal cords vibrating)
Voiceless:
- No energy at low frequencies (Spectrum) (First formant is at a higher frequency)
- No pattern can be seen in the waveform

QUESTION 2 *(worth up to 5 marks)*
Which sounds are most affected when the low-pass cut-off frequency is set to around 500 Hz - vowels or consonants - and why?

Consonants are most affected. This is because vowels are voiced which have lower frequencies (250Hz to 1000Hz). Consonants on the other hand have higher frequencies (1500Hz to 6000Hz). The low-pass filter allows through lower frequencies (vowels) while severely damping higher frequencies (consonants). However, you could still hear dampened consonants as the filter doesn't entirely cut-off the frequencies but rather heavily dampen them at a rate of 6dB per octave.

How is it that the speech is still quite intelligible when the high-pass cut-off frequency is set to 10 kHz?

A high-pass filter is mainly used for noise cancellation. It's given a roll-off frequency (Also known as the cut-off frequency). PD implements a one-pole high-pass filter with a roll-off rate of 6dB per octave. The higher the cut-off frequency, the more aggressive the dampening of the lower frequencies which results in a softened voice that has less noise. PUT SCREEN SHOTS (BEFORE AND AFTER) (Listen to lecture 12)

QUESTION 4 *(worth up to 5 marks)*
COM3502-4502-6502: The [GraphicEqualiser~] object uses an FFT internally; what does FFT stand for and what does an FFT do?
COM4502-6502 ONLY: What is a DFT and how is it different from an FFT?

FFT stands for Fast Fourier Transform. It separates a complex sound signal into it's individual frequency components each with their own amplitude and phase. (RE-WORD) // LIsten to next lecture and lecture 12/13

QUESTION 5 *(worth up to 10 marks)*
With speed = 50 and depth = 0.5, what are the minimum and maximum amplitudes of your LFO output, and how do they vary with changes in these two settings? Also, please provide two screenshots: (a) your [LFO~-help] object and (b) the internal structure of your [LFO~] object.

The minimum amplitude is -0.5 and the maximum is 0.5. Varying the speed doesn't change the amplitude as speed determines the frequency, which ranges from 0 to 50. Varying the depth doesn't change the frequency but rather the amplitude with a maximum of 1 and minimum of -1. MAKE A HELP OBJECT

screenshot

QUESTION 6 *(worth up to 5 marks)*
In your own words[1], why is this effect known as 'ring modulation'?

Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer.

QUESTION 7 *(worth up to 5 marks)*
Why is SSB commonly used in long-distance radio voice communications?

Amplitude modulation on a signal results in two sidebands. One of them is higher in frequency while the other is lower in frequency. To create a SSB, one of these two sidebands (most commonly the one lower in frequency which is known as the lower sideband) along with the carrier wave are removed. This results in an improved and more efficient signal in regards to power and spectrum which is ideal for long-distance radio voice communications.

QUESTION 8 *(worth up to 5 marks)*
COM3502-4502-6502: Why can the voice be shifted up in frequency much further than it can be shifted down in frequency before it becomes severely distorted? /emphHint: look at [`wsprobe~`].
COM4502-6502 ONLY: Your frequency shifter changes all the frequencies present in an input signal. How might it be possible to change the pitch of a voice *without* altering the formant frequencies?

Replace this text with your answer. Replace this text with your answer. Replace

---

[1]I.e. do not plagiarise from Wikipedia.

3

this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer.

QUESTION 9 *(worth up to 5 marks)*
In a practical system, why is it important to keep the feedback gain less than 1?

Feedback being applied is known as Positive feedback. Such feedback results in an overall increased signal in terms of magnitude. On increasing the gain above 1, exponential growth occurs, parameters reach extreme magnitudes causes chaotic extreme behaviour. Exploring through Waveform has shown that the harmonics have increased significantly while the waveform has shown a massive increase in frequency and amplitude. This is why positive feedback tends to cause system instability.

QUESTION 10 *(worth up to 50 marks[2])*
Please provide a short[3] description of the operation of your [VoiceChanger] application, together with a screenshot of your final GUI.

Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer. Replace this text with your answer.

screenshot

---

[2]25 for functionality, 15 for design/layout, 5 for Pd features, 5 for innovations

[3]no more than 200 words