# EECS 225d: Midterm, March 19, 1999

Name:

The total test time is estimated as 50 minutes, but you may stay afterwards and work on the test for as long as you would like (up to an additional hour) since the room is free that long.

The test is composed of two main parts. The first part consists of small problems with a point total of 26. The second part consists of 6 larger problems that are worth 8 points each: choose 3 to answer (maximum score for part II is 24).

For multiple choice (or true/false) questions, also write a brief (1-2 sentence) explanation of why the answer is correct; a correct answer without an explanation will only get half-credit. The parenthesized numbers after each question number give the relative point value of each question.

**PART I: Small Problems**

1. (2) An idealized (uniform cross-sectional area tube) vocal tract with a length of 17 cm has its first three resonances at 500 Hz, 1500 Hz, and 2500 Hz. An 8.5 cm vocal tract (for a small child) would have these resonances at

   **a)** 250, 750, and 1250

   **b)** 1000, 3000, and 5000

   **c)** 250, 1250, and 2250

   **d)** 750, 1750, and 2750

2. (2) The ceiling of a cubical room is lifted by a crane so that it is twice as high, and the walls are extended up (to meet the new ceiling position) by using perfectly reflective materials. Ignoring air absorption, frequency dependencies, nonlinearities, or spatial dependencies, what is the effect on the reverberation time (e.g., RT60) of this change?

   **a)** increase the reverberation time for short sounds and decrease it for long sounds

   **b)** double the reverberation time

   **c)** halve the the reverberation time

   **d)** increase the reverberation time by a factor of 4.

3. (2) Two noise monitoring consultants are checking on noise pollution outside of this building using sound pressure level (SPL) meters. Consultant number 1 is using a meter that measures A-weighted SPL, while consultant number 2 is using a meter that measures unweighted SPL. If the noise is primarily below 100 Hz, which would be the higher level?

   **a)** The A-weighted SPL

   **b)** The unweighted SPL

   **c)** The meters would show the same value

   **d)** The meters would explode, leaving the city council no choice but to close down the university.

4. (2) A multivariate Gaussian is fitted to the examples in each of two classes. These two Gaussians are then used to design a minimum probability of error classifier. Under which condition is the decision surface for this classifier guaranteed to be a hyperplane (a linear classifier)?

   **a)** The variance for each feature is the same

   **b)** The covariance matrix is non-diagonal but equal for each class

   **c)** The covariance matrix is diagonal but different for each class

   **d)** The data was collected in Cincinnati

5. (2) The articulation index (AI) for nonsense syllables filtered from 500 Hz to 3 kHz is $x$, and for filtering from 3 kHz to 6 kHz is $y$. According to Fletcher's AI theory, the AI for syllables filtered from 100 Hz to 6 kHz is

   **a)** $x - y$

   **b)** $x + y$

   **c)** at least $x + y$

   **d)** at least $x \times y$

6. (2) Lower spectral resolution at high frequencies is a common property of

   **a)** Linear predictive analysis

   **b)** Critical band filter banks

   **c)** Cepstral analysis

   **d)** Swedish basketball players

7. (2) A Multi-Layer Perceptron (MLP) is composed of units that compute a sigmoidal nonlinearity $f(y) = \frac{1}{1+e^{-y}}$, where $y$ is a weighted sum of inputs to the unit. Which of the following statements is true about this system?

   **a)** The chain rule of differentiation cannot be used directly to optimize the network weights, due to the nonlinearity.

   **b)** The outputs of the system can be anywhere in the range [-1,1].

   **c)** The input layer is often called "hidden" because it is not directly observable from the network outputs.

   **d)** When the number of weights in the MLP is large in comparison with the number of training patterns, a large number of training epochs can lead to overfitting of the training data.

8. (2) Consider a 10th order linear predictor, used for speech that is sampled with a 10 kHz sampling rate. Choose the condition under which the predictor would be expected to have the largest error signal:

   **a)** when the speech is low in energy

   **b)** for female speech

   **c)** at the start of each new period of voiced speech

   **d)** in the middle of each new period of voiced speech

9. (2) Critical bandwidths based on perceptual experiments

   **a)** increase logarithmically over the complete range of hearing

   **b)** increase like a $\frac{1}{3}$ octave bank of filters

   **c)** remain constant over most of the frequency range

   **d)** remain constant for low frequencies and increase logarithmically thereafter

10. (2) True or False: The perception of pitch is only possible when listening to a harmonic series.

11. (3) A typical auditory fiber has a refractory period of about 3 ms. This means that no new firings can occur for 3 ms after a spike. Explain how the auditory system manages to perceive pitch as high as several kilohertz despite this constraint.

12. (3) Define 3 of the following 5 terms (1 sentence per definition):

**a)** neural phase-locking

**b)** critical band

**c)** simultaneous masking

**d)** cochlear apex

**e)** forward masking

**PART II: Larger problems (choose 3 to do)**

13. (8) Let $h(n) = \alpha^n + \beta^n$.

**a)** Find the z-transform $H(z)$ of $h(n)$

**b)** Under what conditions is the digital filter stable? Unstable?

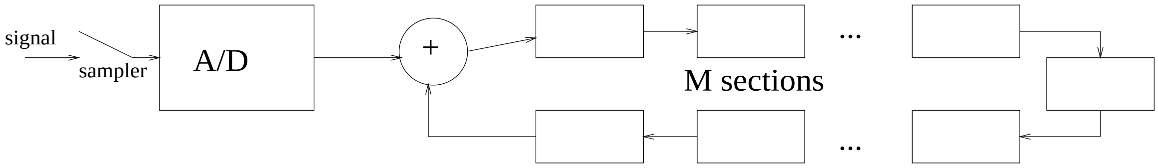
**c)** If $\alpha = 0.8$ and $\beta = -0.8$, find the z-plane pole-zero plot of $H(z)$.

**d)** Sketch the response along the unit circle.

14. (8) Consider the finite sequence $x(n)$, where $x(0) = 1$, $x(1) = 1$, $x(2) = 1$, $x(3) = 2$, $x(4) = 1$, $x(5) = 1$, and $x(6) = 1$.

   **a)** Find the *linear* convolution of $x(n)$ with itself, using direct convolution.

   **b)** Describe an algorithm that should yield the same result, using the DFT.




15. (8) An acoustic tube is represented digitally by a spatial pattern.

    Describe the conditions required to prevent aliasing in time and in space.

16. (8) Describe the EM procedure as it would be used to determine the parameters that maximize the likelihood of the data. Your discussion should not depend on any particular form for the probability estimators or distributions, but should otherwise describe the process as clearly as possible.

17. (8) Describe Voice Onset Time (VOT). Why did experimental results on VOT measurements lead to the concept of "categorical perception?" Describe any physiological evidence that supports the notion that VOT is perceived in connected speech.

18. (8) Speech production, music production by string, and music production by horns (trumpets, trombones, etc.) have in common that an excitation signal drives an acoustic device. In each of these 3 cases, describe the excitation signal or signals and the corresponding acoustic device.