EEL6586 – Spring 2002
Exam 1 March 20, 2002

NAME: _____

This exam is open-book and calculator. You may use any books or papers that you like. There are four problems on this exam, you have two full class periods. State your assumptions and reasoning for each problem. Justify your steps and clearly indicate your final answers.

1	**/25**
2	**/25**
3	**/25**
4	**/25**
Extra	**/5**
TOTAL	

1. (25 points)

   A formant synthesizer of the type used in Homework #1 is used to generate a voiced phoneme with a pitch of 200 Hz and with a sampling frequency fs=8kHz. Suppose that the first formant is set to have an unrealistic center frequency of 1000Hz, 3dB bandwidth of 200Hz, and a power spectral amplitude of 30. The second formant peaks at 2000Hz, 3dB bandwidth of 400Hz and a power spectral amplitude of 10. Answer the following questions.

   (a) (10 points) Carefully sketch the power spectrum of the resulting speech along with a rough spectral envelope on the same linear plot. Be sure to label all the relevant parameters such as peak values, bandwidths, 3dB heights, center frequencies, etc.

(b) (10 points) Assume that the speech is filtered by a pre-emphasis factor of $(1 - .7z^{-1})$. What are the new **approximate** center frequencies, **approximate** bandwidths and **approximate** power spectral amplitudes of the two formants?

(c) (5 points) Sketch the power spectrum on a linear plot after the pre-emphasis is applied to the speech. Again, be quantitative with all the relevant parameters (peak values, bandwidth, etc). What are the significant changes from the sketch from part(a)?

2. (25 points) Assume that white noise excitation $w(n)$ is filtered by an *all-pole* vocal-tract model $H(z) = 1/(1 + .25z^{-2})$ and then passed through a pre-emphasis filter of $(1 - .7z^{-1})$ to produce a signal $s(n)$. $w(n)$ is defined:

$$E\{w(n)w(m)\} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

In this problem you will use LPC to analyze $s(n)$.

(a) (5 points) Derive the difference equation for $s(n)$. Make sure you get this difference equation correct because you will use it in the next few parts.

(b) (5 points) Compute the autocorrelation values $r(0)$ and $r(2)$ for the speech signal $s(n)$.

(c) (10 points) Compute the autocorrelation value $r(1)$ for the speech signal $s(n)$.

(d) (5 points) Compute the first two LPC coefficients ($p = 2$). Hint: pre-emphasis changes the LPC parameters.

3. (25 points) An HMM model is set up using three **fair** coins.

Coin 1 has "A" on both sides.

Coin 2 has "A" on one side and "B" on the other.

Coin 3 has "B" on both sides.

The transition probabilities are given by

$$a_{11} = a_{22} = \frac{2}{3} \qquad a_{12} = a_{23} = \frac{1}{3} \qquad a_{33} = 1$$

All remaining transition probabilities are zero. The user cannot see which state (i.e. coin) is being flipped but they are told of the result of the flip. The state sequence must always start with coin 1.

(a) (5 points) Explicitly write down the A, B, and π matrices/vectors that completely determine this HMM. (You must write them as matrices or vectors).

(b) (10 points) The model produced the following sequence of observations: A-A-B. What is the most likely state sequence that produced this observation? Also, given this most likely state sequence, what is the probability of generating the observation sequence? Show all of your work.

(c) (10 points) What is the probability that this model produced the sequence of observations: A-A-B? Show all of your work.

4. (25 points) Short Answer.

   (a) (5 points) The cepstral coefficients of a recorded speech signal $x(n)$ are given by $\hat{x}(n)$. How do these cepstral coefficients change when a pre-emphasis factor of $(1 - .7z^{-}1)$ is applied to the speech creating a modified signal $y(n)$? Write an equation for $\hat{y}(n)$.

   (b) (5 points) Describe two major problems with using the original time-domain speech samples as features for a speech recognition system.

(c) (5 points) Pre-emphasis filters are usually of the form $(1 - \alpha z^{-1})$ with typical values of α of 0.95 and 0.97. Why is the value of $\alpha = 0.7$ a poor choice?

(d) (5 points) It is well-known that using windows that are too long can smooth between phonemes and lose information. What is the problem, if any, of using windows that are too short?

(e) (5 points) Use your results from problem 2 to quantify the change in center frequency of the single formant when pre-emphasis is applied. Note: you can get this problem right even you didn't get the correct answer for problem 2.

Do one of the following two extra credit problems. You cannot get credit for both, clearly mark which problem should be graded.

Extra Credit 1: (5 points) Prove that the 3db bandwidth of a formant caused by a single dominant pole can be approximated by

$$bw \approx -\ln(r) f_s/\pi$$

where r is the distance of the pole to the origin and f_s is the sampling frequency in Hz.

Extra Credit 2: (2 points) Name the four speech demos that were shown at this year's E-Fair. 0.5 points for each correct answer, you must be specific to get this correct.