EEL6586 – Spring 2004

Exam 1 March 24, 2004

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**
TOTAL	

1. (25 points) A train of impulses is fed through a simple all-pole model to produce a stationary voiced phoneme. The all-pole model is given by:

$$H(z) = \frac{1}{1 + \frac{1}{8}z^{-3}} = \frac{1}{(1 + \frac{1}{2}z^{-1})(1 - \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2})}$$

   (a) (5 points) How many formant frequencies are provided by this $H(z)$? Explain. (Hint: think about what the frequency response of $H(z)$ looks like.)

(b) (10 points) Assuming that the sampling frequency Fs = 2kHz, determine the numerical value of the frequency of each formant.

(c) (10 points) Sketch the resulting speech waveform in the time domain for an input train of pulses with a pitch of 200Hz. Label significant numerical values.

2. (25 points) Assume that white noise excitation $w(n)$ is filtered by a single-pole vocal-tract model

$$H(z) = \frac{1}{1 + \frac{1}{8} z^{-3}}$$

to produce a signal $s(n)$. $w(n)$ is defined:

$$E\{w(n)w(m)\} = \begin{cases} \sigma^2 & m = n \\ 0 & m \neq n \end{cases}$$

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

(a) (5 points) Compute the autocorrelation value $r(0)$ of the speech signal $s(n)$.

(b) (5 points) Compute the autocorrelation values $r(1)$ and $r(2)$ for $s(n)$.

(c) (5 points) Compute the autocorrelation values $r(3)$ for $s(n)$.

(d) (10 points) Compute the first three LPC coefficients by solving the resulting system of equations ($p = 3$). Show ALL of your work.

3. (25 points) An HMM-based phoneme recognizer has been trained to recognize two phonemes e and o. These phonemes are modelled by a 2-state discrete-output HMM, with 3 output symbols in the output distribution (thus, each feature vector is quantized to one of three possible symbols). The three output symbols will be called x, y and z. The states will be referred to as state 1 and state 2. **The models are left-to-right and must start in the initial state (state 1) and end in the final state (state 2).** The two trained HMM models have the same transition matrix:

$$A_e = A_o = A = \{a_{ij}\} = \begin{bmatrix} .9 & .1 \\ 0 & 1 \end{bmatrix}$$

The two output probability matrices for phoneme 1 and 2 are given by:

$$B_e = \{b_i(k)\} = \begin{bmatrix} .8 & .7 \\ .1 & .1 \\ .1 & .2 \end{bmatrix} \qquad B_o = \{b_i(k)\} = \begin{bmatrix} .2 & .1 \\ .4 & .3 \\ .4 & .6 \end{bmatrix}$$

The columns of the B matrix denote the probabilities of the three symbols (x, y and z) for states 1 and 2.

(a) (5 points) How many valid state sequences are possible for a state sequence of 4 symbols (T=4)? List the sequences.

(b) (20 points) A sequence of observations is created as x-y-z-y. How will the system classify this phoneme, as an e or an o.? Be sure to clearly list the relevant probabilities.

4. (25 points) Short Answer.

(a) (5 points) Compute the complex cepstrum of

$$H(z) = \frac{1}{1 + \frac{1}{8}z^{-3}}$$

(b) (5 points) Suppose that $r(1) = r(2) = r(3) = 0$ for some stochastic signal. Does this necessarily imply that $r(4) = 0$? Explain.

(c) (5 points) Problem 3 discussed the use of an HMM for phoneme recognition. Given that that single phonemes are being recognized, do you expect a two-state HMM to do any better than a single state HMM? Explain why or why not.

(d) (5 points) Exactly how is white noise different from the phoneme /s/?

(e) (5 points) Your TA recorded himself speaking a single word shown below in a time domain plot. Identify the word in the list of the following:

a) fail
b) test
c) obey
d) encyclopedia
e) none of the above