EECS 225d: Midterm 2, April 24, 1995

Name:

For multiple choice and true/false questions, also write a brief (1-2 sentence) explanation of why the answer is correct. The parenthesized numbers after each question number give the relative point value of each question.

1. (2) The Baum-Welch procedure is

   a) a statistical algorithm for recognizing with an HMM

   b) a statistical algorithm for training an HMM

   c) an approach to filter design

   d) a kind of cosmetic surgery

2. (2) The pitch of a harmonic series with missing fundamental frequency is close to

   a) the fundamental frequency

   b) twice the fundamental frequency

   c) halfway between a) and b)

   d) none of the above

   e) a fastball

3. (10) Briefly describe the different approaches to speech synthesis by 5 of the the following:

   a) Dudley and his Voder

   b) John Holmes and his work with Fant's OVE II Synthesizer

   c) The Haskins Laboratories researchers using the pattern playback.

   d) George Rosen and his Vocal Tract Synthesizer.

   e) Dennis Klatt and his DecTalk.

   f) Synthesis by rule

   g) Sid Fels and GloveTalk II

   Include such topics as motivation, implementation, synthesis strategy.

4. (4) Auditory experiments have determined that an LPC vocoder of order eight is about as good as a channel vocoder that requires 14-16 channels. Briefly present an argument that explains this result.

5. (2) It has often been stated that female speech is more difficult to vocode than male speech. Do you believe this? Explain.

6. (4) User A has an available channel vocoder and user B has an LPC vocoder. Can they communicate without additional hardware? If so, explain; if not, sketch what extra hardware would be needed.

7. (2) If random variable q corresponds to class and x is the feature vector, the minimum probability of error will be achieved by always choosing the value for q that maximizes

   a) $p(q)$

   b) $p(q)p(x|q)$

   c) $p(q)p(q|x)$

   d) $p(x)$

8. (2) Give a statistical condition that would make a linear discriminant function the optimal (minimum probability of error) classifier for a 2-class problem.

9. (2) For a minimum distance classifier, give a reason why Euclidean distance might not be the best one for classification.

10. (4) Give 2 methods of time normalization between reference word templates and input word templates, giving at least one advantage of each.

11. (2) What is a Markov model?

12. (6) Draw a block diagram showing the major components of an isolated word speech recognition system, and explain some reasonable design choices for each major piece. Include an explanation of how the system would be trained.

Answer 3 of the following 5 remaining questions.

13. (6) At different times, the dominant frequency region for pitch perception has been considered to be the region of the fundamental frequency, the region of the 10th harmonic, and the region near harmonics 3, 4, and 5. Discuss the history of these hypotheses in connection with the ideas and experiments of Helmholtz, Schouten, Licklider, and Ritsma.

14. (6) Answer the following about Fletcher's Articulation experiments as presented by Jont Allen:

    a) Why did Fletcher use nonsense CVC's as the database?

    b) What was Fletcher's rationale for defining Articulation bands?

    c) What is the difference between Articulation and Recognition?

15. (6) Why is spectral flattening of a speech signal followed by autocorrelation a good front end for pitch detection? Define autocorrelation and spectral flattening in your answer.

16. (6) Discuss the concepts behind the Pythagorean, Just, and Equal Tempered Scales. Which of these scales is prevalent in western classical music? Why? What is the Pythagorean comma and why is it a problem?

17. (6) A large number of spoken sentences are recorded, stored, shuffled into random order, and split into two groups. The waveforms for one of the groups is processed with a linear time invariant filter. If the unfiltered group is used for training a speech recognizer and the filtered group is used for testing it, how can you choose feature vectors that reduce the effect of this mismatch, without explicitly learning this filter characteristic and compensating for it? Describe any of the techniques we discussed (or any other you can think of that would help).