(520|600).666
Information Extraction from Speech and Text

Homework # 3

Due February 19, 2009.

Review Chapter 2 from *Statistical Methods for Speech Recognition* by Frederick Jelinek.

1. *Two Best Paths*: Devise a two-pass algorithm to find the *two most likely paths* through the HMM trellis. Specifically, let the first pass be the Viterbi algorithm, and devise a second pass over the same trellis to find the second most likely path *knowing* the most likely path. Assume if needed that there are no cycles made up entirely of null arcs.

   (a) Describe your second pass in the same way the Viterbi algorithm is described in Chapter 2.

   (b) Redraw the trellis of Problem 1(e) in Homework #2, run your algorithm on it by hand, and color the two most likely paths.

2. *Outputs on States*: Construct an HMM which produces outputs from states, instead of arcs, that is *equivalent* to the original HMM of Problem 1 in Homework #2. In particular, construct an equivalent HMM with the minimum number of states, and specify the transition probability matrix $p(s'|s)$, and output probabilities $q(0|s)$ for each state.

3. Write a two page summary (including any figures if you wish) of the article

   > L. R. Rabiner and B. H. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, pp 4-16, Jan 1986.

   Note: the copy available on-line through IEEE*Xplore* is missing one page of the article, and you may need to photocopy it from the MSE library.

Start working on Project #1, which is due on the *Tuesday* after this homework is due.