Assigned: Wednesday 2001-03-28 Due: Tuesday 2001-04-03

Dan Ellis <dpwe@ee.columbia.edu>

Background reading:

Read chapters 25 through 27 in Gold & Morgan.

Reading assignment:

"Context independent and context dependent hybrid HMM/ANN systems for vocabulary independent tasks", S. Dupont, C. Ris, O. Deroo, V. Fontaine, J.M. Boite & L. Zanoni, Proc. Eurospeech-97 (Rhodes), vol. 4, pp. 1947-1950. This paper, by some colleagues from Belgium, shows the impact of using different subword units for recognition. Add a summary to your web page. Remember, I want you to try and highlight some detail that stood out for you, or to formulate questions you feel were left unanswered in the paper.

http://www.ee.columbia.edu/~dpwe/e6820/papers/DupRDF97-cdnn.pdf

Practical assignment:

This week we will train a simple classifier on mel-cepstral features.

Look at the examples of reading labels, calculating features, and testing classification in this week's Matlab diary, which uses the function `mfcc` (from Malcolm Slaney's Auditory Toolbox) to calculate Mel-frequency cepstral coefficients. Following the example in the diary, train classifiers using the first 6 cepstral coefficients, the first 12, and 8 coefficients *excluding* the first one (c0, average log energy). You will find all the files you need, including the directory containing the TIMIT speaker in the matlab directory on the web site (be sure to get the new version of `nntrain` that has slightly different initialization behavior). How does classification accuracy vary across these examples? Are there any differences in the confusion patterns (i.e. most common mistakes)?

Projects:

I plan to review each project this week and communicate with you individually via email.