ELEN E6820 Speech & Audio Processing & Recognition **Assignment 9**

Dan Ellis <dpwe@ee.columbia.edu>

Assigned: Thursday 2008-04-03 Due: Thursday 2008-04-10

Reading assignment: "Weighted finite-state transducers in speech recognition," Mohri, Pereira, Riley, Computer Speech and Language, 2002. Add your comments to the Courseworks discussion board as usual.

`http://www.ee.columbia.edu/~dpwe/e6820/papers/MohriPR02-FST.pdf`

Practical assignment:

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

(a) Look at the examples of reading labels, calculating features, and testing classification in this week's Matlab diary, which uses the function mfcc.m (from Malcolm Slaney's Auditory Toolbox) to calculate Mel-frequency cepstral coefficients. Following the example in the diary, use the provided functions to train a neural net classifier using the first 6 cepstral coefficients. You will find all the files you need, including the directory containing the TIMIT speaker data, in this zip file. Repeat the training several times and average the results.

(b) Repeat part (a) using the 6 coefficients *excluding* the first one (c_0, average log energy).

(c) Make plots comparing classifier outputs with ground truth for one of the soundfiles. How does classification accuracy vary across the two versions? Are there any differences in the confusion patterns (i.e. most common mistakes)?

Project: I hope you are all making solid progress; the end is in sight!