Flat start training in HTK

(see also flattrain-word.steps!)

Usually, we have only orthographic (word-by-word) transcripts for the training data. In this case, we can perform "flat training" in HMM.

Resources:

- mfc feature files in the mfc/ folder

- word-level transcripts in the ort/ folder or in a "master label file" (.mlf file)

- pronunciation dictionary: network/dict\_big.dic

Problems that we have to solve:

- phonetic transcription: we use the pronunciation dictionary (pronunciation alternatives are allowed)

- phone boundaries: these will be found be the Baum-Welch algorithm, plus iterative re-alignments

- short pauses between words: we will have a special short pause model for this

1. If the ort files are given one-by-one, concatenate them into one big master label file

- concatlabfiles.vbs trainlist.mfc trainall-word.mlf

- now it is not necessary, we have the transcripts in timit-train.rand-word.mlf

2. Phonetic transcription:

- HLEd.exe -T 2 -l ..\mfc -d ..\network\dict\_big.dic -i timit-train.rand-phone.mlf makephones0.led timit-train.rand-word.mlf

- you have to replace the "../mfc/" by "\*/" in the path in timit-train.rand-phone.mlf

3. Create a global statistics (required for the "flat" model):

- HCompV.exe -T 1 -m -C nobyteswap.config -S timit-train.rand.mfc -f 0.01 -M . -o new ..\hmm\proto3st1g

4. Create the "flat" model for all the phones:

- BuildFlatModel.vbs new vFloors ..\hmm\timitphonemes70 ..\hmm\hmm0\newMacros

5. Baum-Welch training:

- HERest.exe -T 1 -C nobyteswap.config -I timit-train.rand-phone.mlf -S timit- train.rand.mfc -H ..\hmm\hmm0\newMacros -M ..\hmm\hmm1 ..\hmm\timitphonemes70

- Repeat this 4-5 times! Don't forget to create a folder for each model in /hmm!

6. Create the sp model from the sil model (manually, see other side of this paper!)

7. Realignment of the labels using HVite in "forced alignment" mode, now including the sp model

- HVite.exe -T 1 -o W -b "<s>" -C nobyteswap.config -a -H ..\hmm\hmm6\newMacros -i timit-train.rand-realigned.mlf -m -t 250.0 -I timit- train.rand-word.mlf -S timit-train.rand.mfc ..\network\dict\_big+sp.dic ..\hmm\timitphonemes70+sp

8. We can now make further training iterations using HERest:

9. We can increase the number of Gaussians, or create tied-state context-dependent phone models...

- for an example of how to create monophone models with 30 Gaussians, see addgauss.bat

9. For recognition, use runhdecode.bat

10. Evaluation: runhresultsw.bat

- there might be problems with the path names in results.mlf!

Creating the sp model from a copy of the sil model:

- change the name from sil to sp

- change the number of states from 5 to 3

- discard state 2 and state 4

- rename state 3 to state 2

- manually create a new transition matrix (values are shown below)

