

DSP HW2-1

HMM Training and Testing

教授：李琳山

助教：王君璇

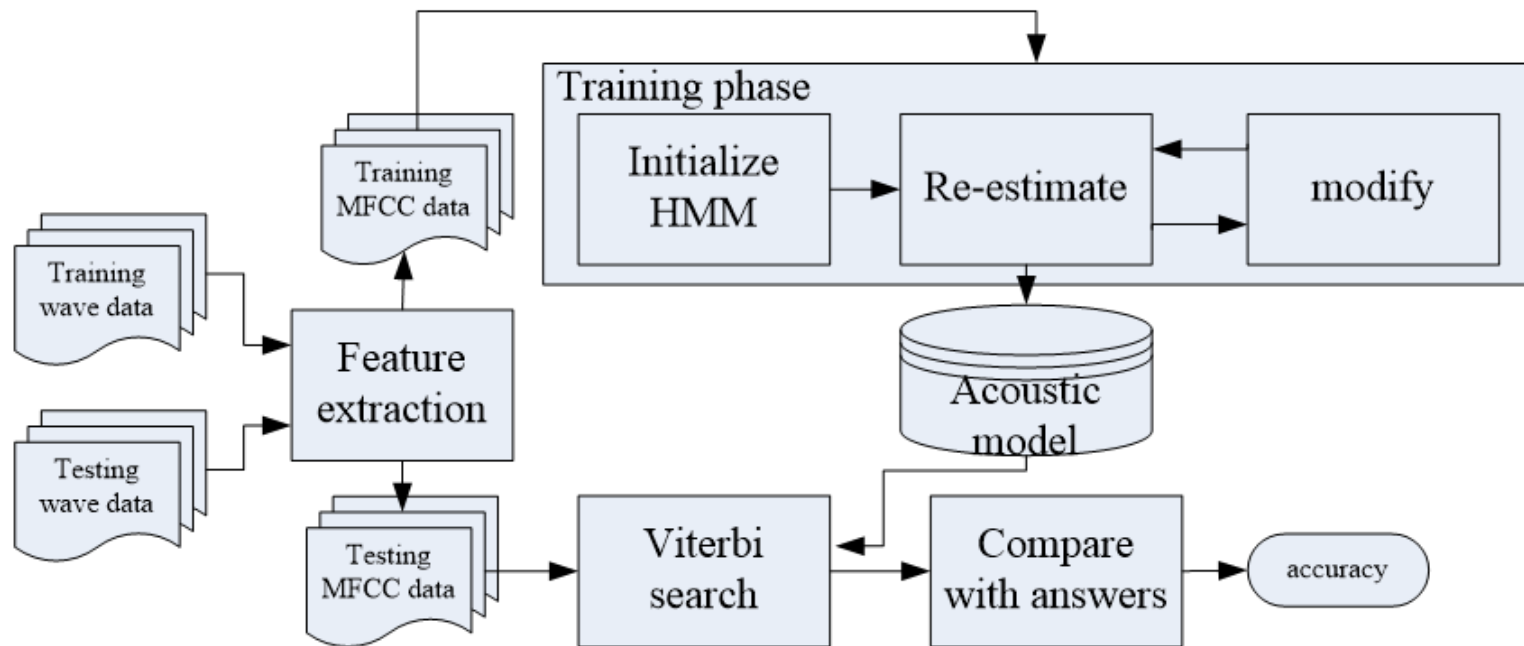
Outline

1. Introduction
2. Hidden Markov Model Toolkit (HTK)
3. Homework Problems
4. Submission Requirements

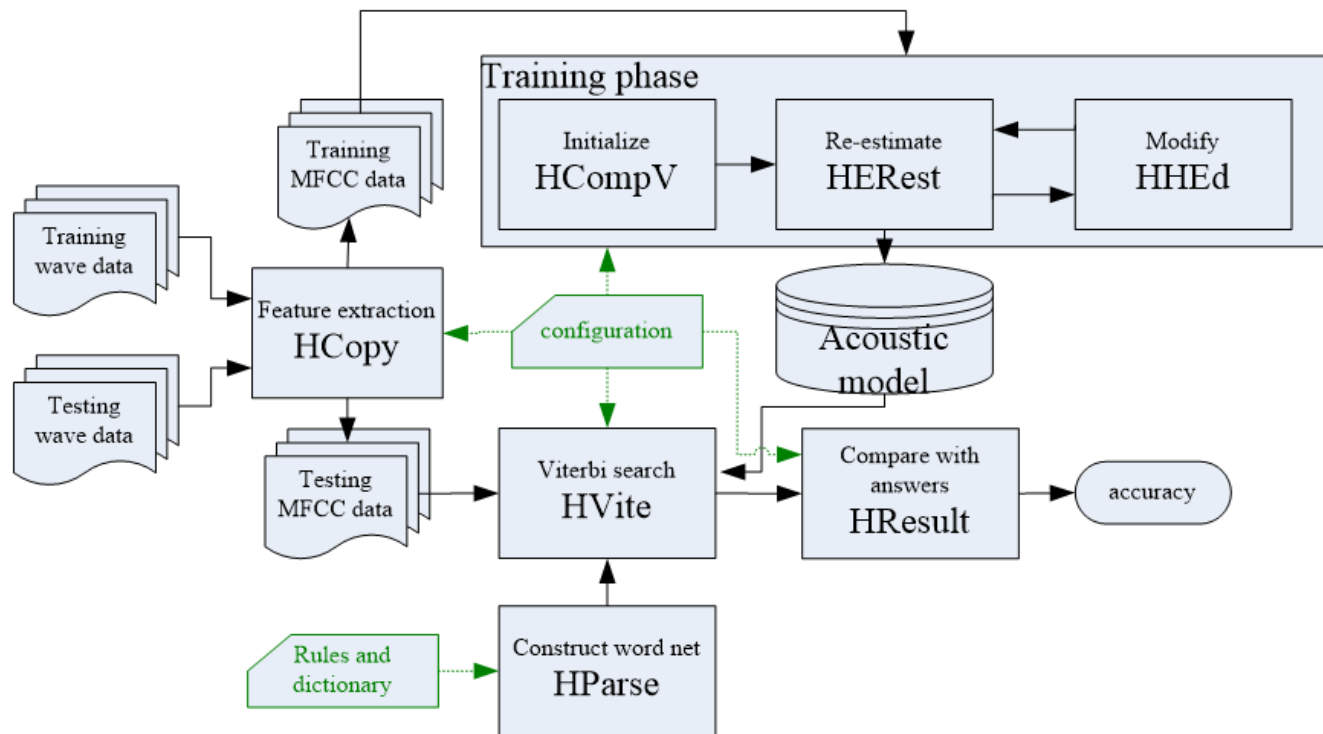
Introduction

- Construct a digit recognizer - monophone
ling | yi | er | san | si | wu | liu | qi | ba | jiu
- Free tools of HMM: Hidden Markov Toolkit (HTK)
<http://htk.eng.cam.ac.uk/>
- Training data, testing data, scripts, and other resources
all are available on
<http://speech.ee.ntu.edu.tw/DSP2018Autumn/>

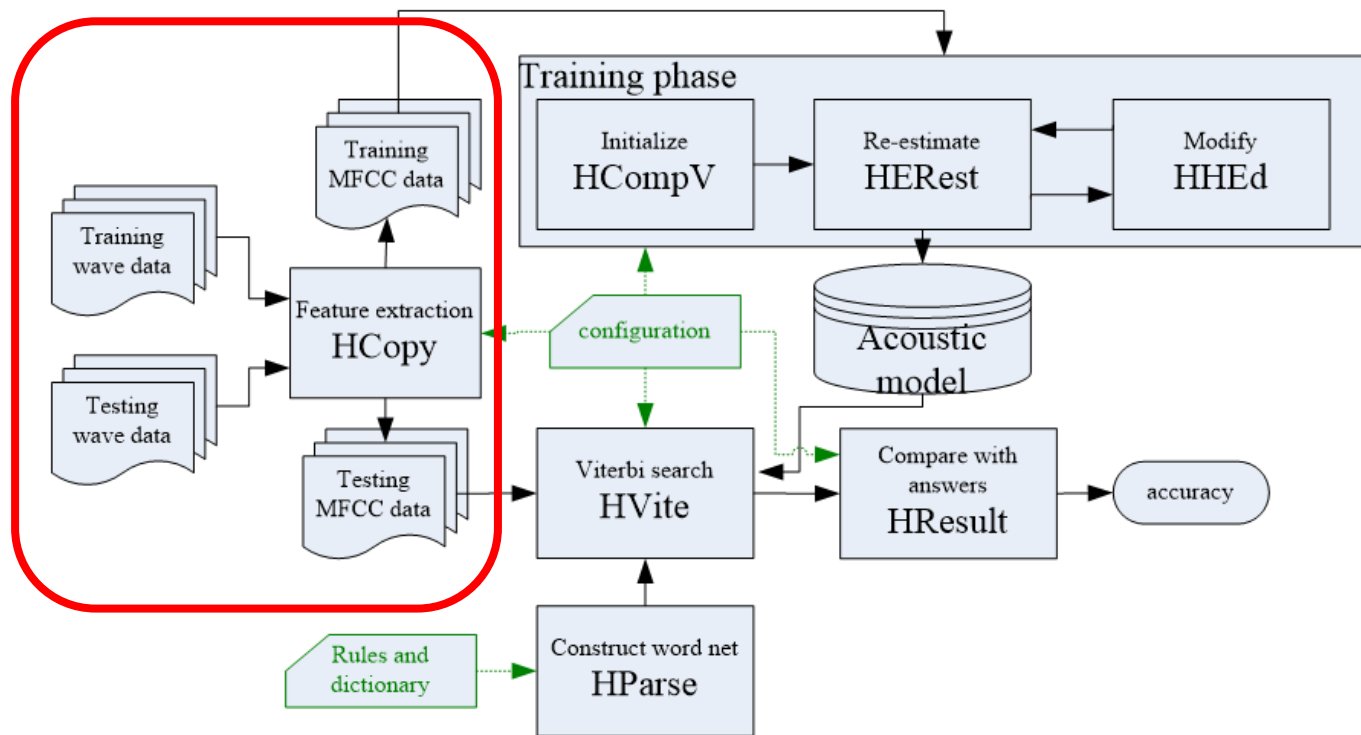
Flowchart



Hidden Markov Model Toolkit (HTK)



Feature Extraction



Feature Extraction - HCopy

```
HCopy -C lib/hcopy.cfg -S scripts/training_hcopy.scp
```

Convert wave to 39 dimension MFCC.

-C lib/hcopy.cfg

- input and output format e.g. wav -> MFCC_Z_E_D_A
- parameters of feature extraction
- [Chapter 7 - Speech Signals and Front-end Processing](#)

-S scripts/training_hcopy.scp

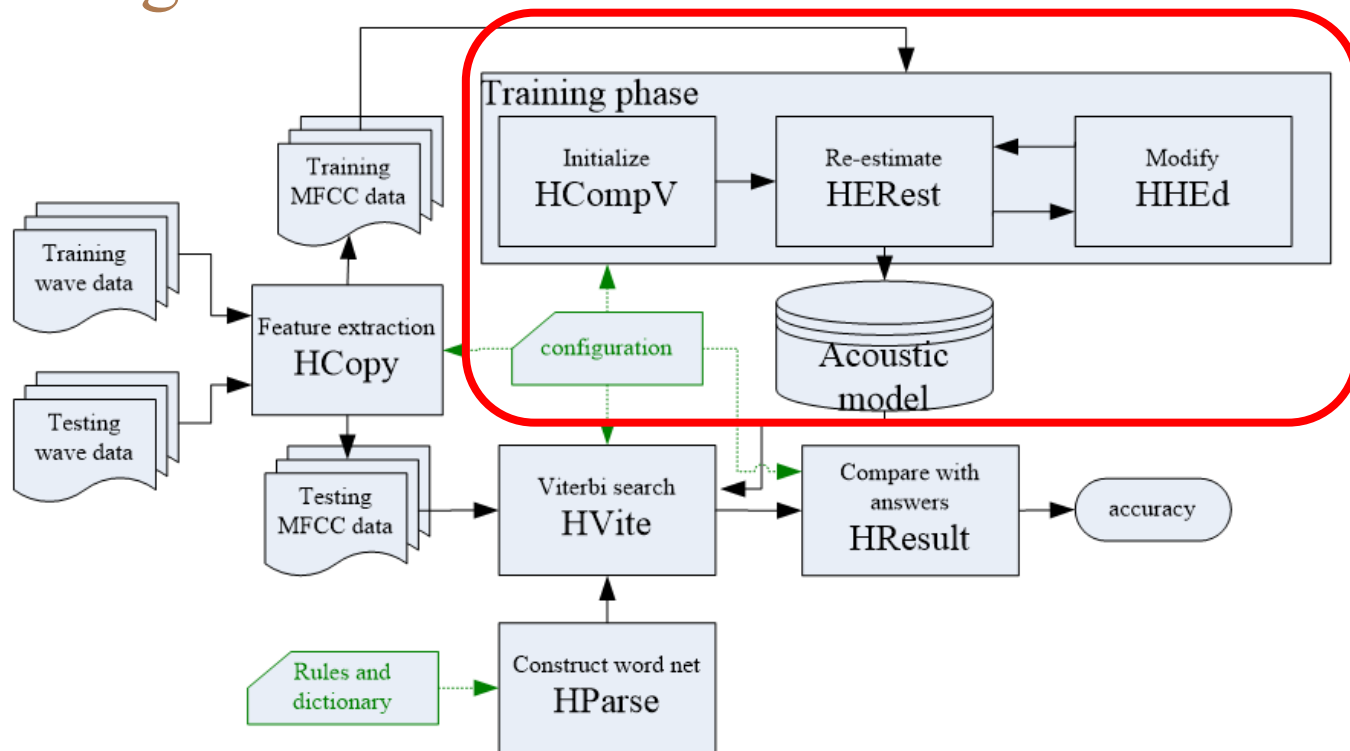
- a mapping from Input file name to output file name

speechdata/training/
N110022.wav

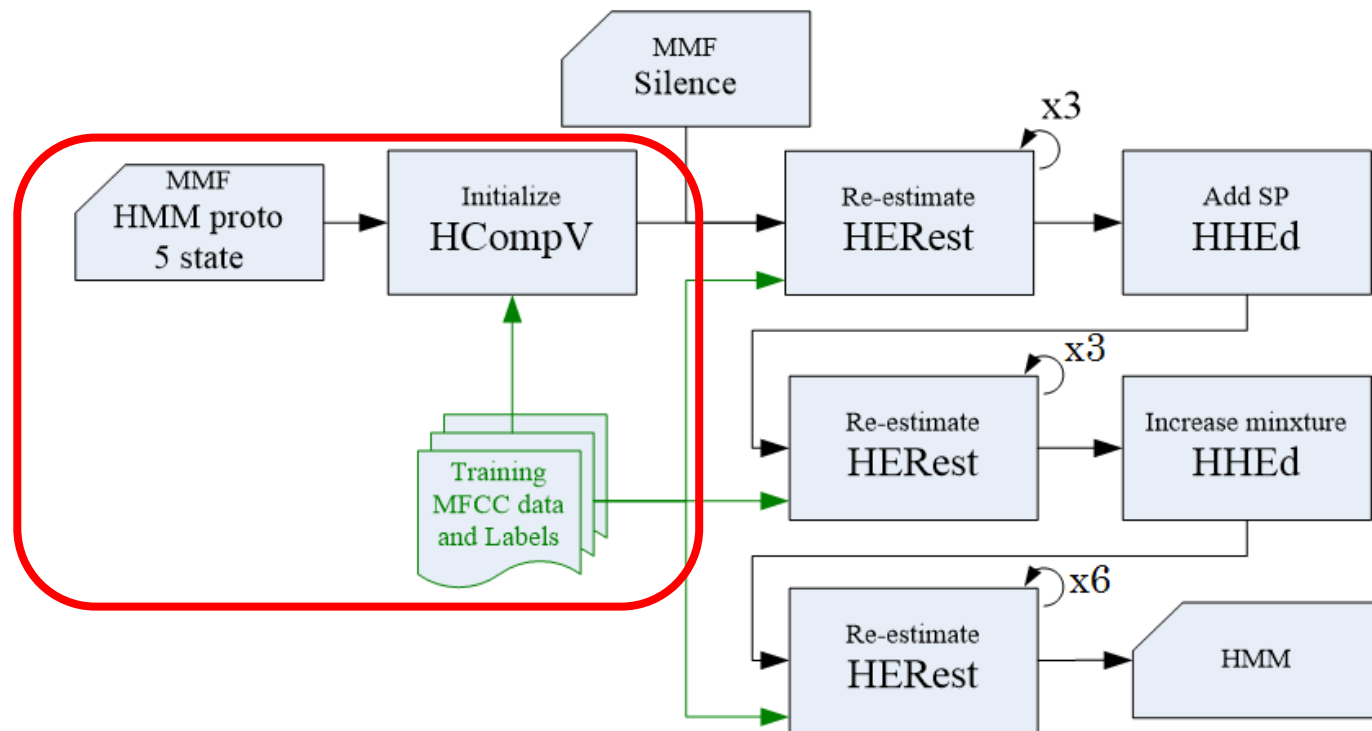


MFCC/training/
N110022.mfc

Training Flowchart



Training Flowchart



Initialize model - HCompV

```
HCompV -C lib/config.cfg -o hmmdef  
-M hmm -S scripts/training.scp lib/proto
```

Compute global mean and variance of features

-C lib/config.cfg

- set format of input feature (MFCC_Z_E_D_A)

-o hmmdef -M hmm

- set output name: hmm/hmmdef

-S scripts/training.scp

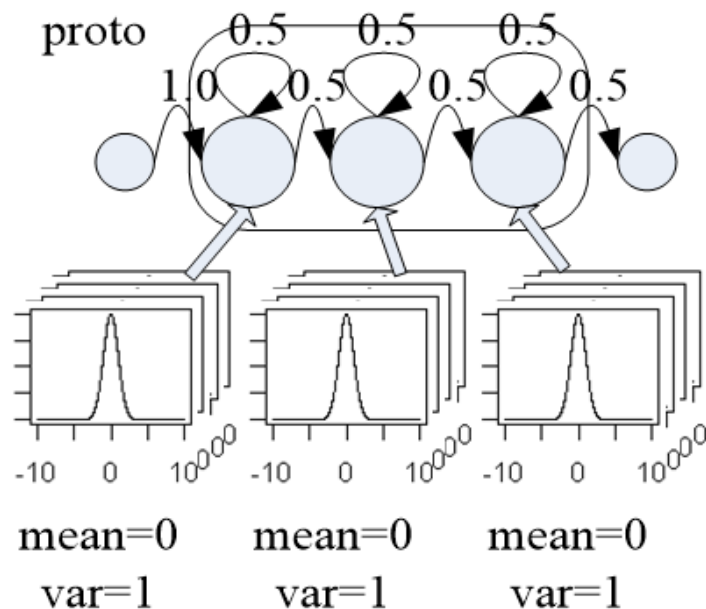
- a list of training data

lib/proto ⇨ you can modify the Model Format here (# states) !

- a description of a HMM model, HTK MMF format

Initial MMF Prototype

MMF: HTKBook chapter 7

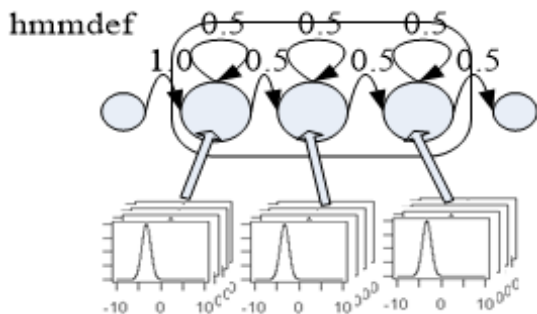


```
~o <VECSIZE> 39 <MFCC_Z_E_D_A>  
~h "proto"  
<BeginHMM>  
<NumStates> 5  
<State> 2  
<Mean> 39  
0.000000.000000.000000.000000.000000...  
<Variance> 39  
1.01.01.01.01.01.01.01.01.01.01.01...  
<State> 3  
<Mean> 39  
0.000000.000000.000000.000000.000000...  
<Variance> 39  
...  
<TransP> 5  
0.01.00.00.00.0  
0.00.50.50.00.0  
0.00.00.50.50.0  
0.00.00.00.50.5  
0.00.00.00.00.0  
<EndHMM>
```

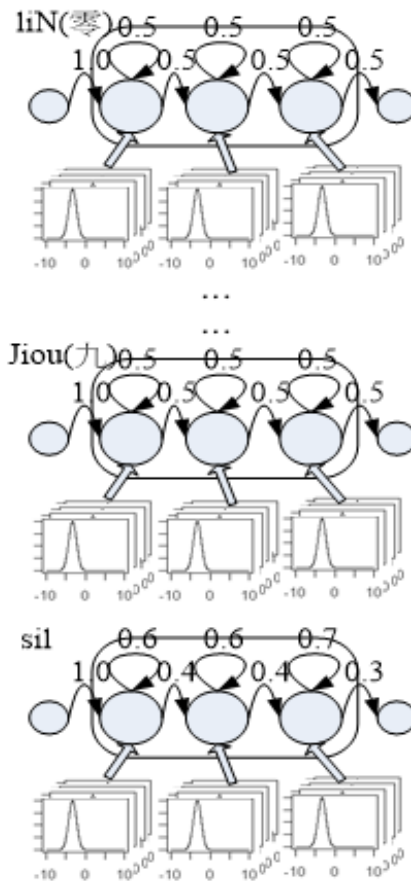
Initial HMM

- bin/macro
Produce MMF contains vFloor
- bin/models_1mixsil
add silence HMM

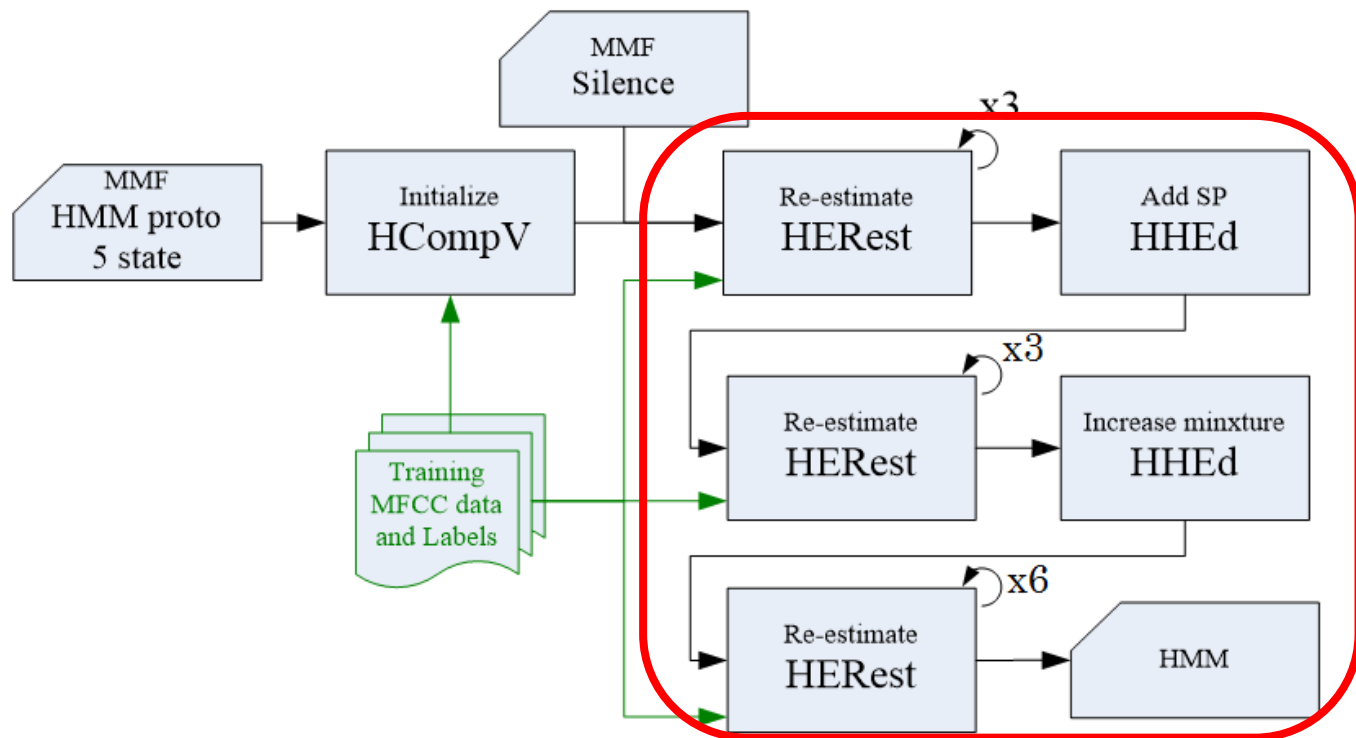
hmm/hmmdef



hmm/models



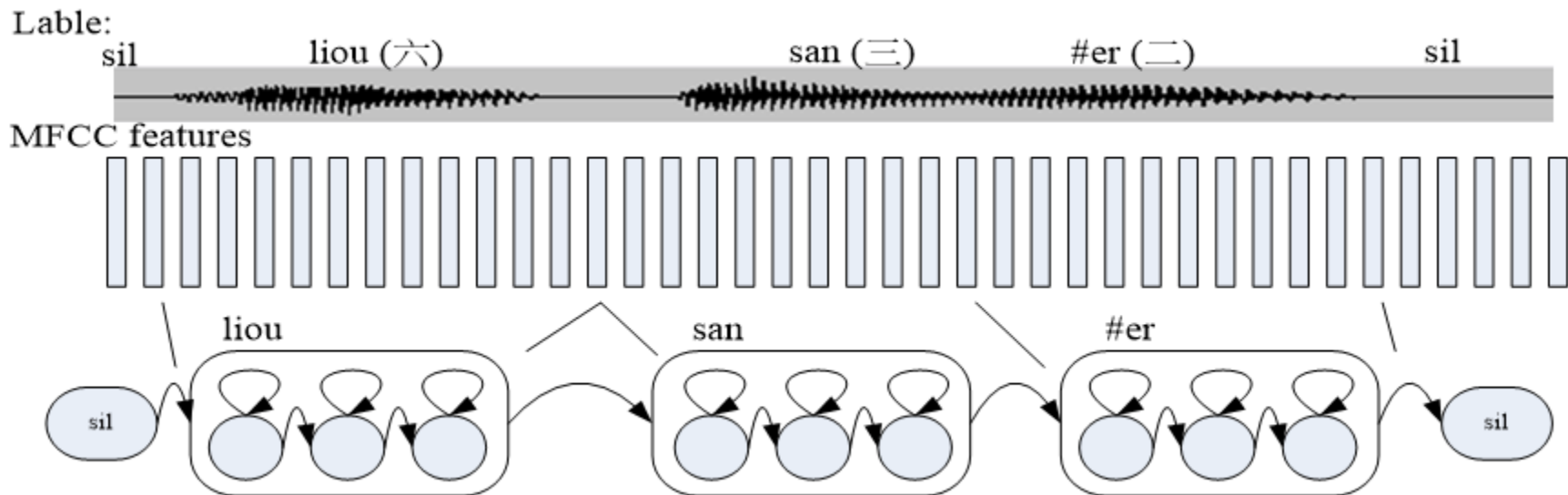
Training Flowchart



Adjust HMMs - HERest

Basic problem 3 for HMM

- Given O and an initial model $\lambda=(A,B, \pi)$, adjust λ to maximize $P(O|\lambda)$



Adjust HMMs - HERest

```
HERest -C lib/config.cfg -S scripts/training.scp -I labels/Clean08TR.mlf  
-H hmm/macros -H hmm/models -M hmm lib/models.lst
```

Adjust parameters λ to maximize $P(O|\lambda)$

- one iteration of EM algorithm
- run this command three times => three iterations

-I labels/Clean08TR.mlf

- set label file to "labels/Clean08TR.mlf"

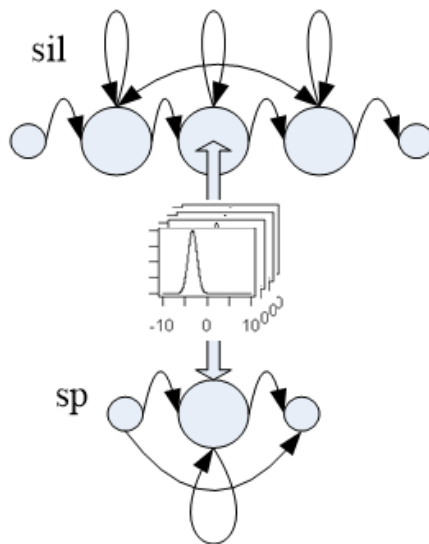
-o lib/models.lst

- a list of word models (liN (零), #i (一), #er (二),... jiou (九), sil)

Add SP Model

```
bin/spmodel_gen hmm/models hmm/models
```

Add "sp"(short pause) HMM definition to MMF file "hmm/hmmdef"



Modify HMMs - HHEd

```
HHEd -H hmm/macros -H hmm/models  
-M hmm lib/sil1.hed lib/models_sp.lst
```

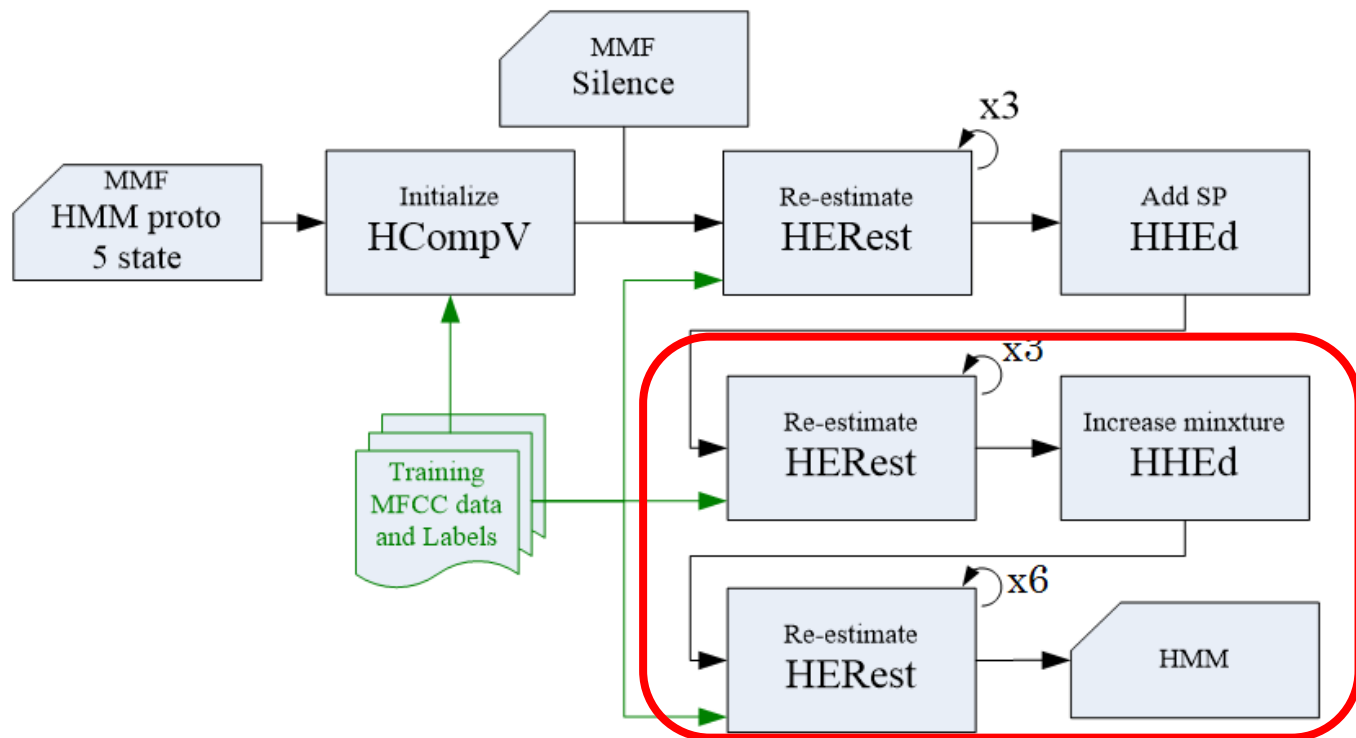
lib/sil1.hed

- a list of command to modify HMM definitions

lib/models_sp.lst

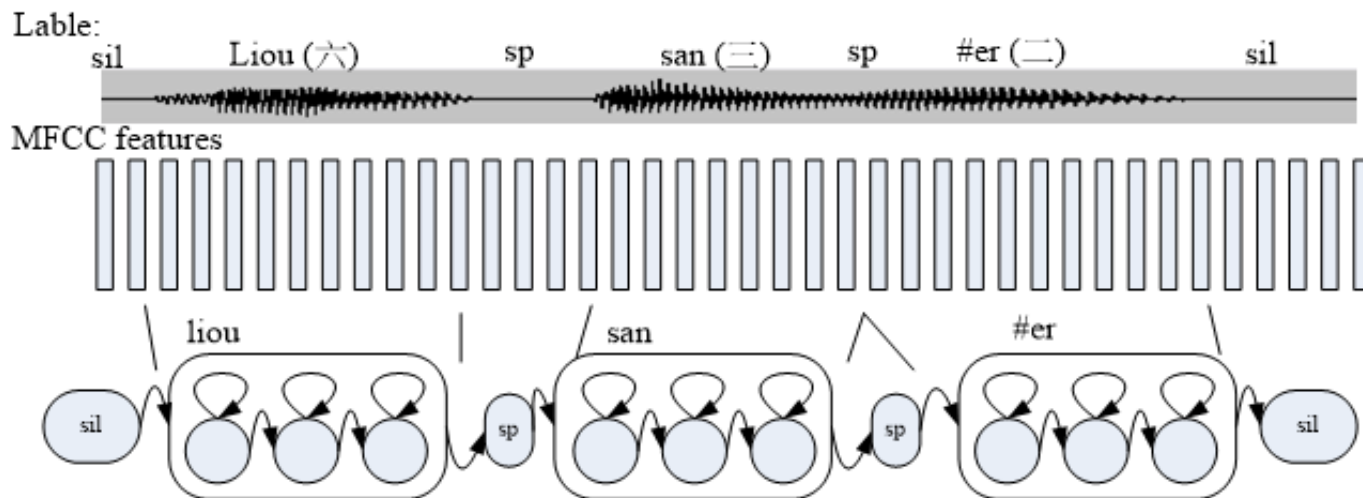
- a new list of model (liN (零), #i (一), #er (二),... jiou (九), sil, sp)

Training Flowchart



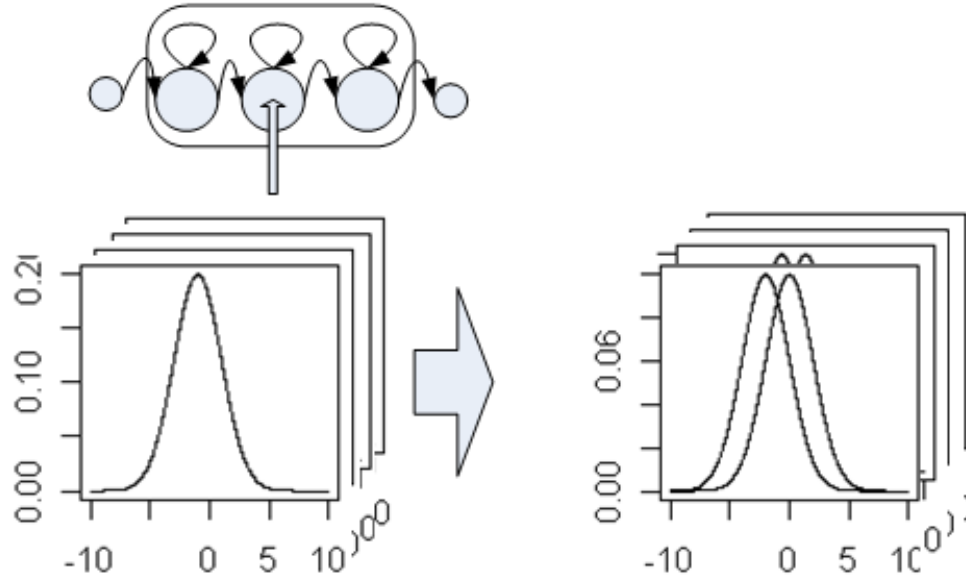
Adjust HMMs Again - HERest

```
HERest -C lib/config.cfg -S scripts/training.scp  
-I labels/CleanO8TR_sp.mlf -H hmm/macros  
-H hmm/models -M hmm lib/models_sp.lst
```



Increase Number of Mixtures - HHEd

```
HHEd -H hmm/macros -H hmm/models  
-M hmm lib/mix2_10.hed lib/models_sp.lst
```



Modification of Models

lib/mix2_10.hed

MU 2 {liN.state[2-4].mix}

MU 2 {#i.state[2-4].mix}

MU 2 {#er.state[2-4].mix}

MU 2 {san.state[2-4].mix}

MU 2 {sy.state[2-4].mix}

...

MU 3 {sil.state[2-4].mix}

You can modify # of Gaussian mixture here.

This value tells HTK to change the mixture number from state 2 to state 4. If you want to change # state, check lib/proto.

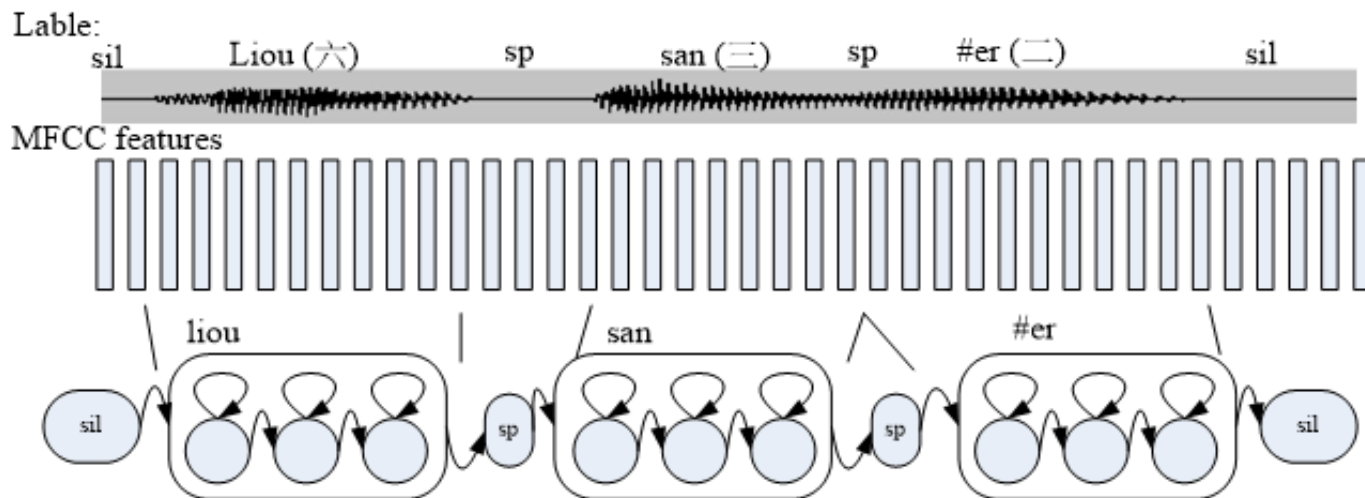
MU +2 {san.state[2-9].mix}

You can increase # Gaussian mixture here.

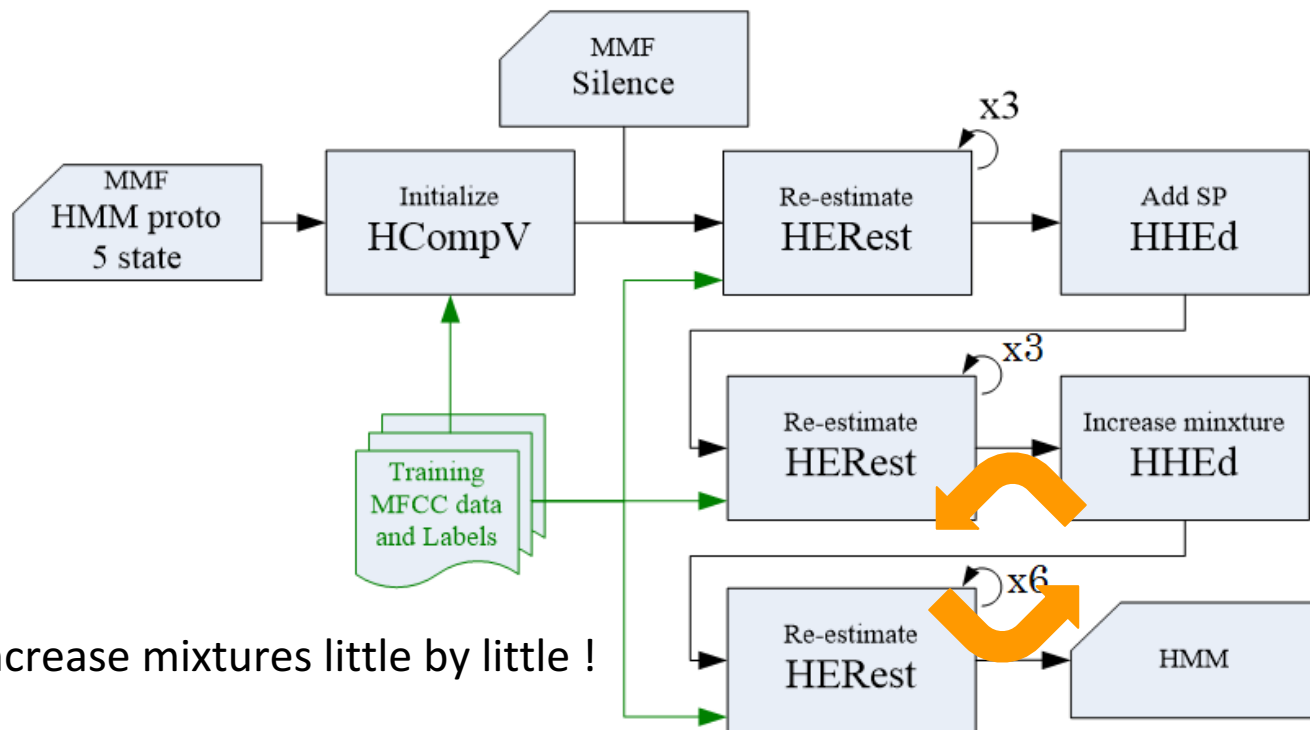
Check HTKBook 17.8 HHed for more details

Adjust HMMs Again - HERest

```
HERest -C lib/config.cfg -S scripts/training.scp  
-I labels/CleanO8TR_sp.mlf -H hmm/macros  
-H hmm/models -M hmm lib/models_sp.lst
```

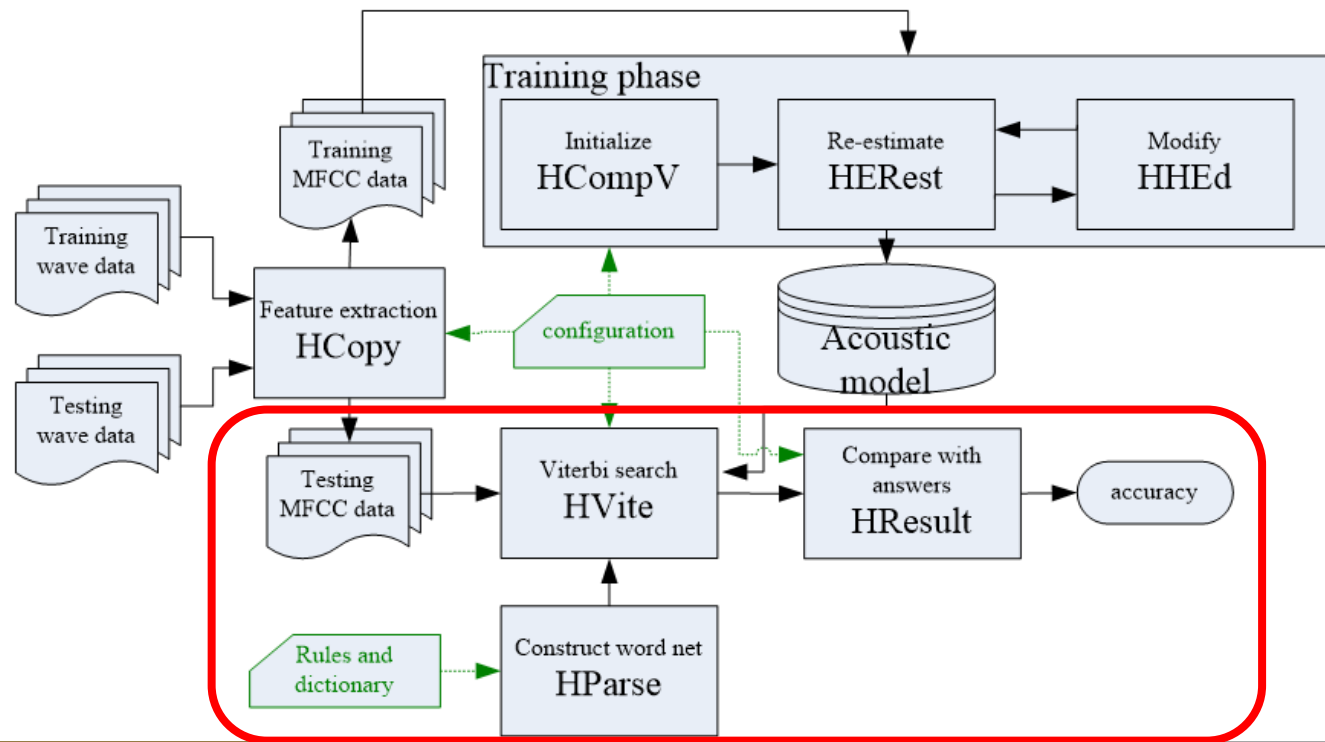


Training Flowchart



Hint : Increase mixtures little by little !

Testing Flowchart



Construct Word Net - HParse

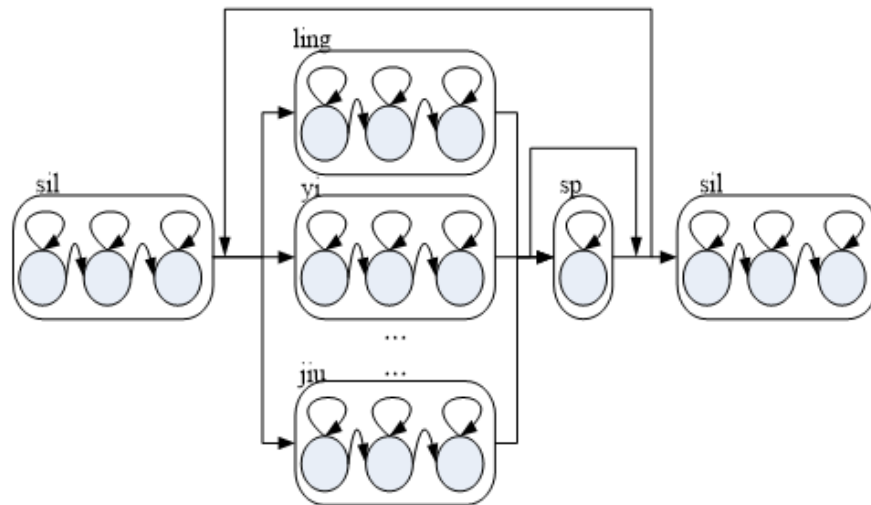
HParse `lib/grammar_sp` `lib/wdnet_sp`

lib/grammar_sp

- regular expression
- easy for user to construct

lib/wdnet_sp

- output word net
- the format that HTK understand



Viterbi Search - HVite

```
HVite -H hmm/macros -H hmm/models -S scripts/testing.scp  
-C lib/config.cfg -w lib/wdnet_sp -l '*' -i result/result.mlf  
-p o.o -s o.o lib/dict lib/models_sp.lst
```

-w lib/wdnet_sp

- input word net

-i result/result.mlf

- output MLF file

lib/dict

- dictionary: a mapping from word to phone sequences
ling -> liN, er -> #er, 一 -> sic_i i, 七 -> chi_i i

Don't worry `WARNING [-8232] ExpandWordNet: Pronunciation 1 of sp is 'tee' word in HVite`

Compared With Answer - HResults

```
HResults -e "???" sil -e "???" sp  
-I labels/answer.mlf lib/models_sp.lst result/result.mlf
```

Longest Common Subsequence (LCS)

===== HTK Results Analysis =====

Date: Wed Apr 17 00:26:54 2013

Ref : labels/answer.mlf

Rec : result/result.mlf

----- Overall Results -----

SENT: %Correct=38.54 [H=185, S=295, N=480]

WORD: %Corr=96.61, Acc=74.34 [H=1679, D=13, S=46, I=387, N=1738]

=====

Report - Part 1 (40%) - Run Baseline

1. Download HTK tools (recommend: compiled binary) and homework package
2. Set PATH for HTK tools : *set_htk_path.sh*
3. Execute (bash shell script)

01_run_HCopy.sh

02_run_HCompV.sh

03_training.sh

04_testing.sh

Report - Part 1 (40%) - Run Baseline (cont.)

3. You can find accuracy in “result/accuracy”
the baseline accuracy is 74.34%
4. Put the screenshot of your result on the report.

Useful tips

1. To unzip files

unzip XXXX.zip

tar -zxvf XXXX.tar.gz

2. To set path in “set_htk_path.sh”

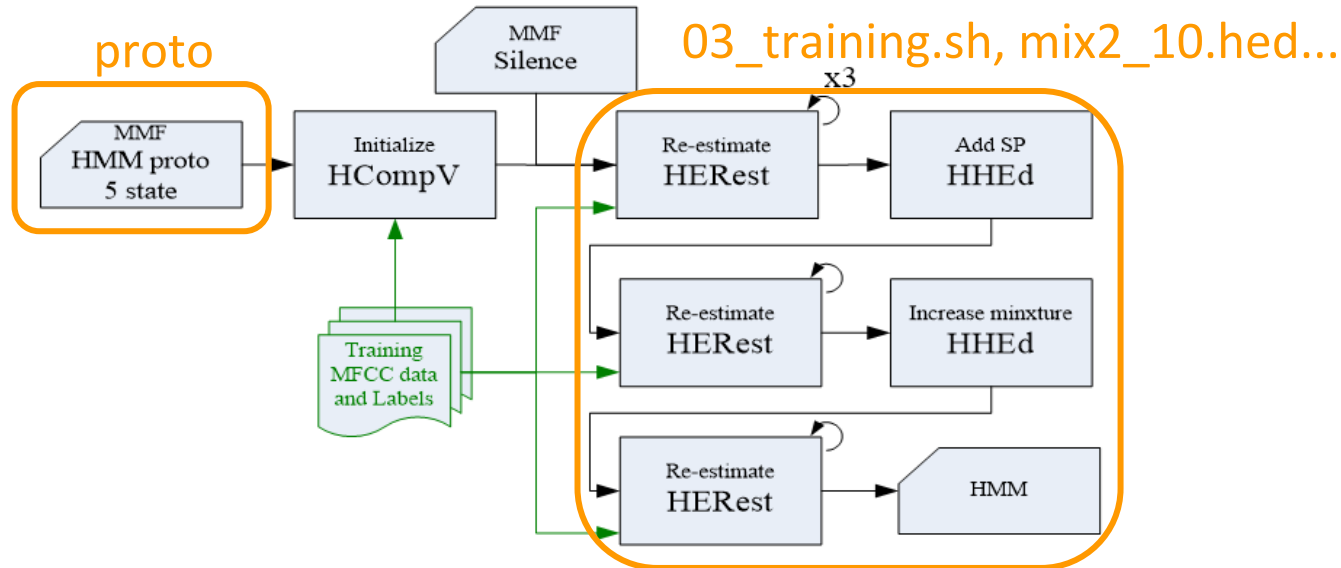
PATH=\$PATH:“~/XXXX/XXXX”

3. In case shell script is not permitted to run...

chmod 744 XXXX.sh

Report - Part 2 (40%) - Improve Accuracy

- Acc > 95% for full credit ; 90~95% for partial credit and put the screenshot of your result on the report.



Part 2 - Attention 1

- Executing 03_training.sh twice is different from doubling the number of training iterations.

To increase the number of training iterations, please modify the script, rather than run it many times.

```
for i in 0 1 2 ;  
do  
    echo "iteration $i"  
    HERest -C $config -I $label \  
        -t 250.0 150.0 1000.0 -S $data_list \  
        -H $macro -H $model -M $mmf_dir $model_list  
done
```


Part 2 - Attention 2

- Every time you modified ***any parameter or file***, you should run ***00_clean_all.sh*** to remove all the files that were produced before, and restart all the procedures. If not, the new settings will be performed on the previous files, and hence you will be not able to analyze the new results.

(Of course, you should record your current results before starting the next experiment.)

Report - Part 3 (30%)

- Write a report describing your training process and accuracy.

Number of states, Gaussian mixtures, iterations, ...

How some changes effect the performance

Other interesting discoveries

- Well-written report may get +10% bonus.

Submission Requirements

1. 4 shell scripts

your modified 01~04_XXXX.sh

2. 1 accuracy file

with only your best accuracy (The baseline result is not needed.)

3. proto, mix2_10.hed

your modified hmm prototype and file which specifies the number of GMMs of each state

4. 1 report (in PDF format)

Submission Requirements (cont.)

5. Put those **8 files in a folder**, compress the folder to 1 zip file and upload it to **CEIBA**.

- Folder name should be bXXXXXXXX (e.g. b04901000)
- .zip only
- **20%** of the final score will be taken off for wrong format

6. Deadline: 11/23 23:59:59

- Late Penalty: 10% off every 24 hours after deadline
(less than 24 hours will be viewed as 24 hours).
- Submission after 3 days will get zero point.

If you have any problem...

- Check for hints in the linux and shell scripts. ex: 鳥哥
- Check the HTK book.
- Ask friends who are familiar with Linux commands or Cygwin. (link : [how to HTK on Cygwin](#))

Contact TA

- email : ntudigitalspeechprocessingta@gmail.com
title: [HW2-1] Problem Description
- Office Hour: Thursday 14:30-15:30 電二531 王君璇
(Please send an email before coming!)