Digital Speech Processing Hw#2

B04501095 黃平瑋

- **baseline accuracy**

```
====================== HTK Results Analysis =======================
  Date: Fri May 18 02:13:18 2018
  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]
===================================================================
~
```

- **Improving accuracy**

  1. increase iteration
  the original script in **03_training.sh** only run HEREST 3 times, I increase the iteration time to 15 so as to converge to better accuracy.

  2. add more states to model
  I increase the states from 5 to 12 in model of liN(零) to jiou(九), since that more states can slice sound to smaller part and each part will contain less feature. Then, each state can describe these feature more properly.

  3. increase Gaussian mixtures
  Since the observation is a continuous random variable, adding more Gaussian model can approach to real distribution.

  final result

```
====================== HTK Results Analysis =======================
  Date: Fri May 18 15:57:26 2018
  Ref : labels/answer.mlf
  Rec : result/result.mlf
------------------------ Overall Results --------------------------
SENT: %Correct=90.00 [H=432, S=48, N=480]
WORD: %Corr=97.58, Acc=96.84 [H=1696, D=26, S=16, I=13, N=1738]
===================================================================
~
```

- **Training Process**

  number of state: 12
  Gaussian mixtures: 4
  training iteration: 15
  Training process is based on the slides and scripts provided by TAs.

I also found that increasing iteration time in training only slightly improve performance, however, increasing the number of state can significantly improve performance.