

## Introduction to Digital Speech Processing – Homework 2

資工三 B05902023 李澤諺

### Part 1. Run Baseline

```
jnfem112@jnfem112-VirtualBox:~/Introduction-to-Digital-Speech-Processing/Homework 2/dsp_hw2$ cat result/accuracy
===== HTK Results Analysis =====
Date: Wed Apr 10 12:04:03 2019
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]
=====
jnfem112@jnfem112-VirtualBox:~/Introduction-to-Digital-Speech-Processing/Homework 2/dsp_hw2$
```

### Part 2. Improve Accuracy

首先，我調整了 proto 檔案中的 NumStates，觀察到其對 accuracy 的影響如下：(在新增加的 state 中，皆有 39 個 Gaussian distribution，其 mean 和 variance 皆分別為 0.0 和 1.0，而 transition probability 皆設定為，第 1 列中僅有第 2 行為 1.0，其餘皆為 0.0，而最後一列的每一行皆為 0.0，至於剩下的其它列，第 n 列中僅有第 n 行和第 n+1 行為 0.5，其餘皆為 0.0)

NumStates	accuracy	NumStates	accuracy
5 (original)	74.34	13	95.57
6	81.47	14	95.80
7	87.80	15	95.91
8	89.64	16	95.97
9	90.91	17	95.57
10	93.67	18	95.11
11	94.25	19	95.05
12	95.22	20	94.42

由此可知，增加 NumStates 可顯著提升 accuracy，但是當 NumStates 超過 16 時，便會導致 overfitting，使得 accuracy 下降。因此在最後的結果中，我選擇使用 NumStates 為 16。

接著，調整 mix2\_10.hed 檔案和 mixAdd2\_10.hed 檔案中的參數。

(由於想觀察參數變化對 accuracy 的影響程度，擔心使用 NumStates 為 16 的 model，其 accuracy 已經很高，若參數可使得 accuracy 提高，其提升的幅度必定不大，很難看出其影響的幅度，因此以下皆是修改最剛開始的檔案，一次僅變動一個參數，觀察單一個參數變動對 accuracy 的影響)

首先，觀察到 MU 對 accuracy 的影響如下：

(以下為僅變更 sil，即 mix2\_10.hed 最後一行中的 MU 所得到的結果，另有觀察將每一行的 MU 皆一起改為以下數字，發現其所得的結果和下表差不多，因此推斷僅改變 sil 的 MU 和將全部的 MU 皆改為同一個數字，其影響程度差不多)

MU	accuracy	NumStates	accuracy
3 (original)	74.34	7	76.64
4	74.17	8	76.29
5	74.57	9	76.87
6	74.34	10	77.45

由此可知，增加 MU 會使得 accuracy 有上升的趨勢，不過影響程度不如 NumStates 顯著(也有可能是我調整參數的方式不好，有其它種調整方式可以使得 MU 對 accuracy 的影響程度大幅上升，但可能的調整方式太多種，過於麻煩，因此沒有嘗試)。

接著，我試著改變 mixAdd2\_10.hed 檔案中 MU+的值，以及 mix2\_10.hed 檔案和 mixAdd2\_10.hed 檔案中的 state[2-○]，發現不論如何調整，accuracy 皆沒有改變(也有可能其實 accuracy 有略為改變，但四捨五入輸出後數值才看起來像是沒有變化)，推測這些參數對 accuracy 的影響程度不大(也有可能是我調整參數的方式不好或是不對所致)。

最後，調整 03\_training.sh 中 iteration 的次數，發現其對 accuracy 的影響如下：

MU	accuracy	NumStates	accuracy
3、3、6 (original)	74.34	3、6、6	64.79
4、3、6	65.65	⋮	⋮
5、3、6	66.17	3、15、6	65.48
6、3、6	66.23	3、3、7	64.56
⋮	⋮	3、3、8	64.90
15、3、6	66.63	3、3、9	64.79
3、4、6	64.56	⋮	⋮
3、5、6	64.90	3、3、20	65.48

由此可知，一次僅改變一個 iteration，將 iteration 增加時，accuracy 沒有比最剛開始的檔案還要好(我也有嘗試將 3 個 iteration 一起增加，如 iteration 為 6、6、9 時，accuracy 為 66.63，仍沒有比最剛開始的 73.34 高)(有可能是我 iteration 還不夠大，或參數調整方式不好或是不對所致)。

綜合以上觀察，我最後選擇使用 NumStates 為 16，且 mix2\_10.hed 中的 MU 皆為 15，而其餘參數皆維持不變的 model，其所得到的 accuracy 為 96.32。

```
jnfem112@jnfem112-VirtualBox:~/Introduction-to-Digital-Speech-Processing/Homework 2/dsp_hw2$ cat result/accuracy
===== HTK Results Analysis =====
Date: Wed Apr 10 23:25:50 2019
Ref : labels/answer.mlf
Rec : result/result.mlf
----- Overall Results -----
SENT: %Correct=88.54 [H=425, S=55, N=480]
WORD: %Corr=96.43, Acc=96.32 [H=1676, D=42, S=20, I=2, N=1738]
=====
jnfem112@jnfem112-VirtualBox:~/Introduction-to-Digital-Speech-Processing/Homework 2/dsp_hw2$
```