

數位語音處理 HW1

R07922004 資工所 碩一 吳星耀

A. 環境設置：

OS: Ubuntu 16.04 LTS

Language: C++

Compiler: g++ 4:5.3.1 -1ubuntu1

B. 執行方式：

Train:

```
./train iteration model_init.txt seq_model_01.txt model_01.txt
```

Test:

```
./test modellist.txt testing_data.txt result.txt
```

C. 程式架構：

train.cpp:

首先先把 seq_model_0X.txt 中的內容讀取完畢後，對每段 sequence 都計算完 α 、 β 後計算並累加至 γ 與 ε ，並新增 gama_v 變數來累加每個 state 產生某個 o_t 的 γ 。所有 sequence 都計算完畢後再用 Baum welch Algorithm 來更新 hmm 的參數。

test.cpp:

程式使用 Viterbi Algorithm 來實作，但是由於答案並沒有使用到 state sequence，因此在程式中並沒有將機率最高的 state sequence 儲存下來。

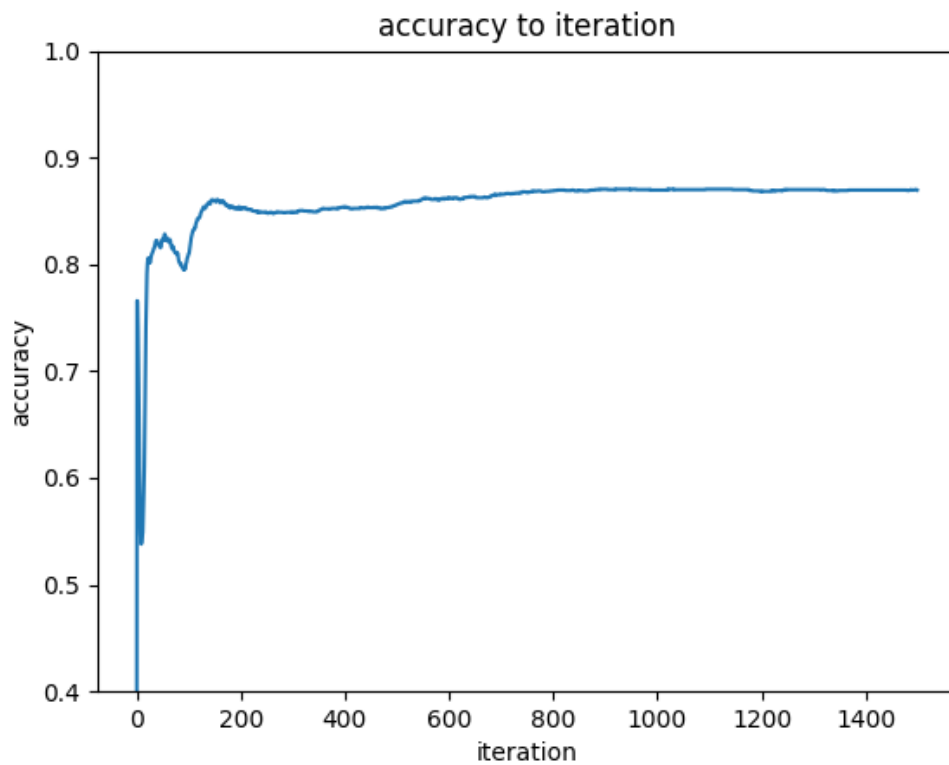
Makefile:

只是將 target 內容修改為 train 與 test，其餘並無太多更動。

另外，結果測試程式是以 python 撰寫。

D. 結果分析：

Iteration 為100次時的 testing data 1準確率為81%，而 iteration 為1000次時準確率為86.96%，下圖為 testing data 1 accuracy 對 iteration 作圖。



由結果可以發現 iteration > 200以後正確率已經趨近定值，甚至在1000以後有時會微微下降。由於 training data 跟 testing data 不可能分布完全相同，所以當參數越來接近 training data 的結果後，testing data 的測量結果下降大致合理。

因此我取 iteration = 887次作為 training 的次數，在 test1.txt 的估計結果為 87%。