

Environment: 本機

How to execute:

(一) 先打 make 編譯 train.c 跟 test.c

(二) 打 ./train \$iteration(大於 1000 最穩) ../model\_init.txt

../seq\_model\_01.txt model\_01.txt

./train \$iteration(大於 1000 最穩) ../model\_init.txt

../seq\_model\_02.txt model\_02.txt

./train \$iteration(大於 1000 最穩) ../model\_init.txt

../seq\_model\_03.txt model\_03.txt

./train \$iteration(大於 1000 最穩) ../model\_init.txt

../seq\_model\_04.txt model\_04.txt

跟

./train \$iteration(大於 1000 最穩) ../model\_init.txt

../seq\_model\_05.txt model\_05.txt

(打”../”是因為我不是把 model\_init 跟 seq\_model 放在資料夾，是放在前一個目錄)

最後再打

./test ../modellist.txt ../testing\_data1.txt result1.txt

./test ../modellist.txt ../testing\_data2.txt result2.txt

(一樣 modellist 跟 testing\_data 放在前一個目錄)

Summary of my program:

train.c:

把 sequence 從檔案讀進來，做 argv[1] 次 iteration，每一次 iteration 做的動作是：先把參數歸零，然後計算每個 alpha, beta, gamma, 跟 epsilon，因為更新 hmm 需要把所有 sequence 的狀況都記錄下來，所以我開變數去存 gamma 跟 epsilon 的和，每個 sequence 做完一次就加一次。最後再用那些存 gamma 跟 epsilon 和的變數算新的 pi, a 跟 b。

test.c:

讀取 sequence，然後每個 sequence 都拿去算 model 1 到 5 的機率，機率是用 delta 算的，哪個 model 機率最大就是哪個，如果是做 testing\_data1 的話生成一個 acc.txt 算正確率。