

# DSP HW1

資工三 B05902001 廖彥綸

## 一、架設環境

資工系工作站( linux 系統)

gcc 編譯器

## 二、執行方式

### make

make 後產生兩個執行檔 train 和 test

**./train times model\_init.txt seq\_model\_?.txt model\_?.txt**

times 為大於 0 的整數，表示疊代次數

model\_init.txt 為初始 model 狀態

seq\_model\_?.txt 為訓練資料

model\_?.txt 為訓練完成後輸出 model 的名字

? 帶入 01 至 05

**./test modellist.txt testing\_data.txt result.txt**

modellist.txt 訓練產生所有 model 的名字

testing\_data.txt 測資

result.txt 輸出結果(每一筆測資對應一個 model 的名字)

## 三、總結

讀進 seq\_model 的資料後先將英文字母轉成數字(A = 0, B = 1...)以便儲存和計算。根據上課和投影片內容依序做出 alpha、beta、gamma 等函式。依序更新 model 中 initial、transition、observation 的值，此為一次疊代。進行數次疊代後，輸出模型。

testing 的部分同樣先對測資做字母轉成數字的處理，之後使用 Viterb Algorithm，以最大可能機率的 model 為答案。

對 testing\_data01 的正確率約為 70% ( 51 次疊代)。當疊代次數太少時，新加入的資料無法足夠影響原本的模型，準確度較低，而疊代次數過多時，新的資料將會使原本模型逐漸降低功用，只能以多次嘗試找出最佳的疊代次數。而疊代次

數非常大時改變數值的起伏相當有限，模型也趨於穩定。

本次作業所有（5 個）聲音模型的初始條線都相同( `model_init.txt` )在實際的情形是不太可能的，所以初始模型的值意義會很小，新的資料可能可以完全主導模型的訓練。