DSP HW2

R05922130 資工所 王瀚磊

**介紹：**

這次作業是使用HTK toolkit實作Hidden Markov Model，將一段wav音訊sample成許多39 dim的vector後進行辨識，然後嘗試修改一些參數將accuracy提高。

**環境建置：**

原本想在自己的電腦(Mac)完成這次作業，但發現需要自己compile，因此決定直接安裝虛擬機器使用Ubuntu，在第一次安裝的時候使用的版本為x64的Ubuntu，安裝完後無法執行compile好的htk tools，最後發現這些執行檔案是在32-bit環境下compile好的，因此重新安裝Ubuntu x86後即可正常執行。

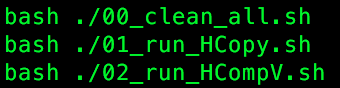
**程式執行：**

**Part 1. Run Baseline:**

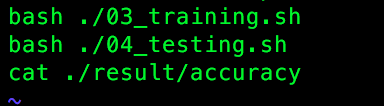
1. 首先將HTK toolkit的位置加入至環境變數，可以利用助教提供的Set\_htk\_path.sh設置。



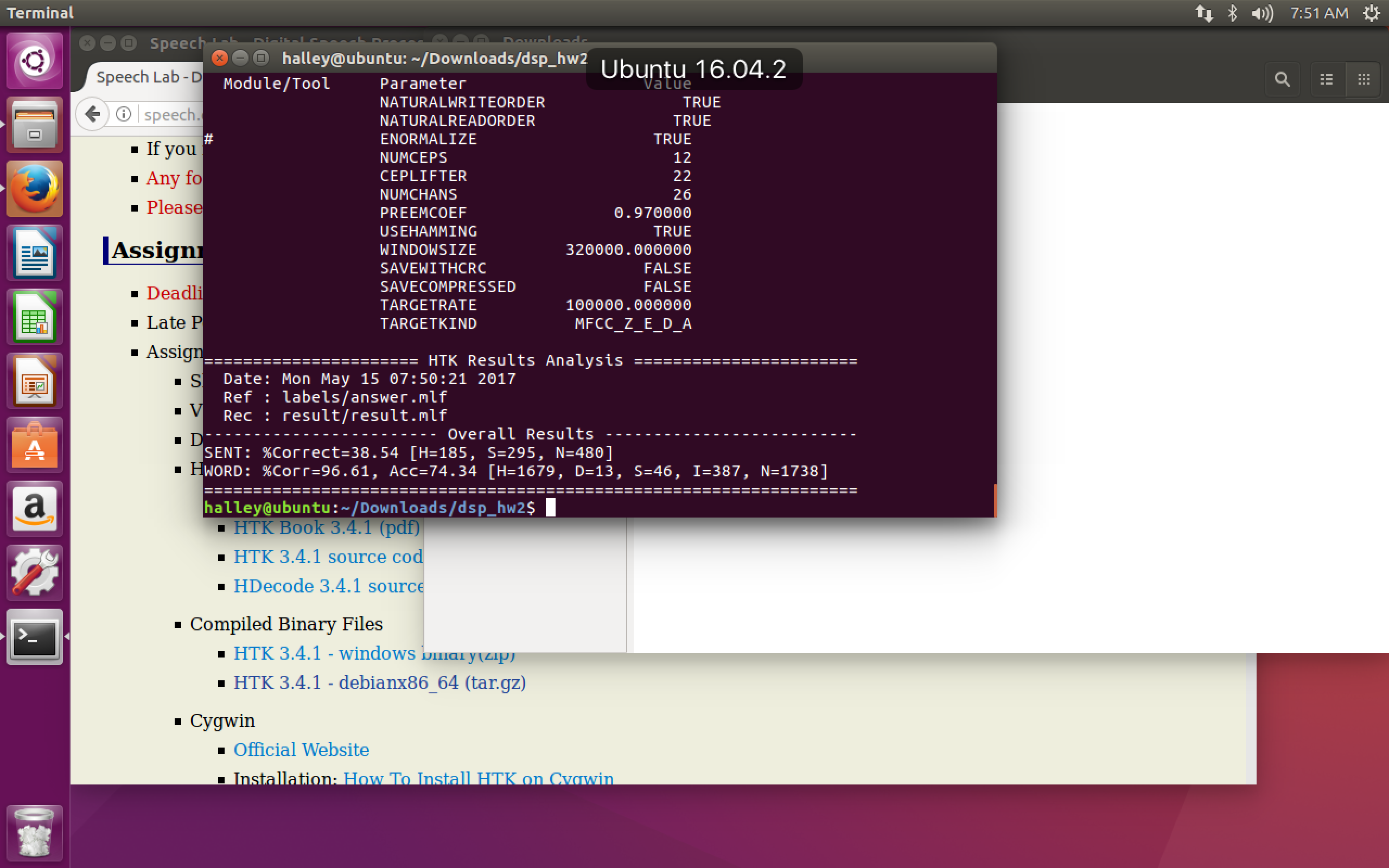
1. 接著先執行clean, HCopy以及HCompv，從wav中sample 39-dim的MFCC，並且做好HMM的初始化。



1. 接著利用自己寫的shell script執行train, test以及輸出等工作。

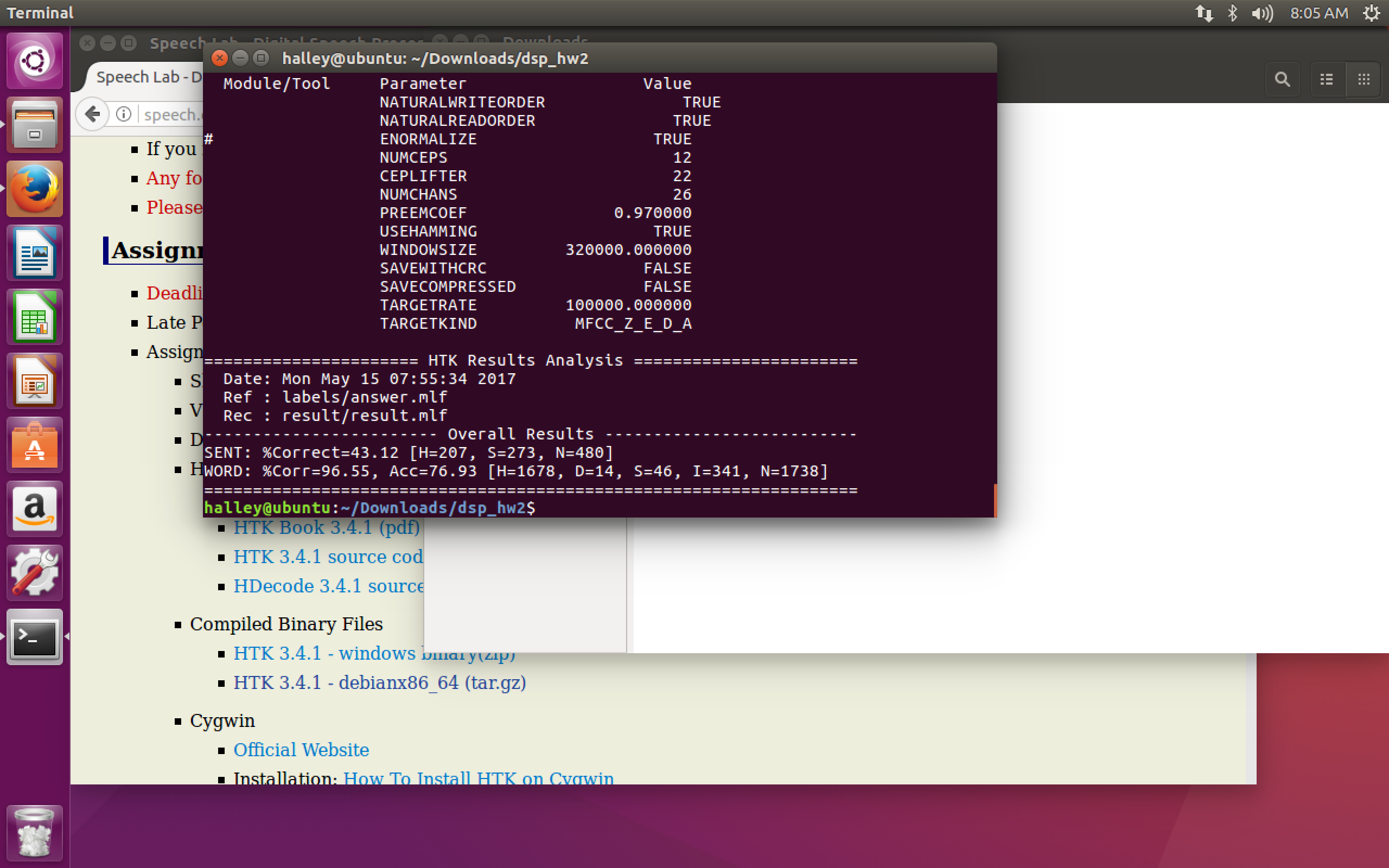


1. 執行結果可以得到助教投影片提到的default accuracy: 74.34%。



**Part 2. Improve Recognition Accuracy:**

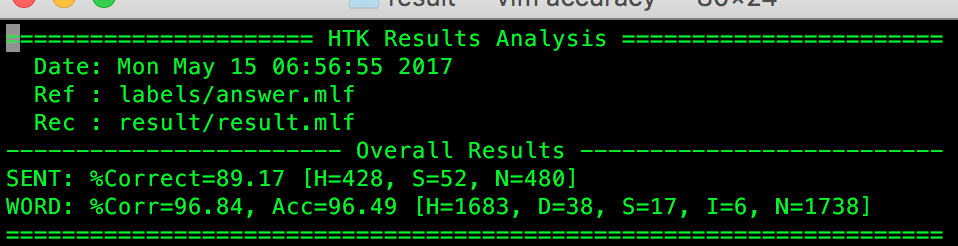
1. 接著我嘗試調高03\_training.sh中的iteration，調到10次後雖然accuracy有些微上升，但距離95%仍然有非常大的距離。



1. 之後開始調整lib/proto中HMM的state數量，將其從原本的5個調整到15個，並且將Transition Probability依照原先的方式定義，有0.5的機率停留在自己，以及0.5的機率跳到下一個state。



1. 經過這個調整，accuracy馬上衝到96%左右，接著再細條iteration至20後得到的accuracy。（可見state數目越高較容易提高其accuracy）



1. 接下來嘗試更改lib/mix2\_10.hed，將原本的state 2-9 or 2-4全部改為2-14，但結果對於Accuracy並沒有太多幫助。

