

DSP HW2 Report

R09944072 鍾宜樺

➤ 遇到的困難：mac M1 晶片不支援這次作業，需用 intel 晶片的電腦。

在自己的電腦 train 時，完全沒有調參數下 training 約 6 mins、testing 約 2 mins、accuracy 約為 -26%。看到 spec 上限時 5 mins 猜測是不是助教希望我們做平行化加速 (Kaldi 有釋出平行化方式)，一開始覺得很驚恐。參數調了三天後，根據 Kaldi 官網敘述跟範例程式，我卻發現當前的參數調整方向不合理。

在我 train 出的最好的 model accuracy = 61.66% (training 7 分鐘) 中參數設定為：transition-scale \approx 0.1、acoustic-scale \approx 0.7、self-loop-scale \approx 0.2。根據官網：參數設定為 0.1 時，代表該參數較不重要，意即 HMM model 的 transition-scale 其實是不重要的參數。當 transition-scale 被設定很小的時候，HMM 容易卡在一個 state 中出不來，這並不合理。

因此向朋友借了電腦跑一次，發現在參數完全相同的狀況下，accuracy 完全不同、程式運行的速度快了將近 10x；也跟借了另一台 M1 的電腦跑一次作業，數據和速度都和在我的電腦上運行結果一樣。故認定或許 Kaldi 並沒有支援 M1 晶片作業，但整個運行過程卻都能正常作業，讓人很難看出來結果是錯的。最後我上去 AWS 借了一個免費的 server，終於完成了這次的作業。

➤ 作業觀察 (以下為在寫作業的時候觀察到的參數設定，以及一點討論)

1. test_beam：用在 beam search 中決定最後輸出幾個機率最高的序列，基本上這會有一個上限值 (根據實驗觀察在=12 以上輸出結果幾乎都相同，這和 #states 有關係)，再往上不會比較好。
2. opt_acwt：這個值主要設定在 0.1 附近 ([ref](#)) 作微調，代表聲學模型的 log-probability 相對語言模型的權重低。

實驗結果：

參數設定 (numiters, maxiterinc, numgauss, totgauss, realign_iters, transition-scale, acoustic-scale, self-loop-scale, opt_acwt, test_beam) = (18, 7, 41, 400, 1-12, 0.8, 0.1, 0.3, 0.19, 14)，accuracy=95.34%

```
((kaldi) root@939e6d3eb526:/opt/kaldi/dsp-hw2-1# bash 4-test.sh
Converting acoustic models to HTK format
viterbi/mono/final.mmf viterbi/mono/tiedlist exist, skipping ...
Generating results for test set with acoustic weight = [ 0.19 ]
output -> viterbi/mono/test.mlf
log -> viterbi/mono/log/latgen.test.log
result -> viterbi/mono/test.rec
accuracy -> [ 95.34 ] %

Execution time for whole script = 00 hours 00 mins 06 secs
```

(preprocess = 0 secs, extract-feat = 19 secs, training time = 3 mins 25 secs; testing time = 06 secs;

Total: 3 min 50 sec)