• Environment:

- 在 win10 使用 atom 撰寫 code, 並使用 LF(Unix based)編碼儲存。
- ubuntu on win 10 compile
- gcc version5.4.0 (Ubuntu 5.4.0-6 ubuntu1~16.04.5)

Execute

- ./mydisambig -text \$(txtfile) -map \$(mapfile) -lm \$(bigramfile) -order 2 > \$(result_dir)
- Bigram.lm Big5-ZhuYin.map (放在與 mydisambig, mapping.py 同資料夾即可), testdata 設定路徑為 (./testdata)

• Other:

本次實作會需要用到 map,雖然 c++本身就有 map 了,但 SRILM 已經有實作較高階的 map (VocabMap),但官方沒有文件,但 header 檔內的內容勉強還可以讀懂,裏頭 friend class VocabMapIter,支援 iter 配合迴圈很實用。

另外在做 dp 時,如果一開始的 2D array 的記憶體配置太小的話,輸出結果會有一些字元會壞掉,後來將 initial 的記憶體大小設大一點問題就解決了。