

數位語音處理概論 Homework_3 Report

B04902028 資工三 洪浩翔

1. Environment: CSIE workstation.

2. How to compile:

輸入 `make clean` 然後 `make (MACHINE_TYPE) (SRIPATH)`即可編譯。

3. How to execute:

1. 若要執行 `bigram`，先生成 `char based` 的 `file`，再用 `srilm` 產生 `bigram.lm`，然後輸入 `make map` 產生 `mapping file`，最後輸入 `make run` 後即會生成 `result2/1.txt ~ 10.txt`。

2. 若要執行 `trigram`，一樣先生成 `char based` 的 `file`，然後再用 `srilm` 產生 `trigram.lm`，再修改 `Makefile` 中的 `LM ?= bigram.lm` 改為 `LM ?= trigram.lm`，以及 `./mydisambig` 的地方 `-order` 改為 `3`，然後 `make map` 產生 `mapping file`，最後輸入 `make run` 後即會生成 `result2/1.txt ~ 10.txt`。

4. What I have done:

`Mapping.py` 生成 `mapping file`，`mydisambig.cpp` 支援 `bigram` 跟 `trigram` 的運算，但輸入 `order 2` 跟 `3` 以外的東西會停止並輸出錯誤訊息。另外，由於 `mapping` 中若使用 `big5` 會無法顯示如“裏”等字體，因

此改為使用 **big5hkscs** 來解決此問題。若 **mydisambig** 出現不符合原本設定的輸入也會噴 **error** 然後程式錯誤。最後，**mydisambig** 只支援 **char based**，若是輸入未分割的 **file** 結果會有蠻嚴重的錯誤。