

DSP HW3 report

B04705036 電機三 許天鈞

Environment

OS: macosx sierra

MACHINE_TYPE: macosx

processor: Intel Core i5

How to compile

在 dsp_hw3_b04705036 根目錄下

make all (mydisambig.cpp 會被 compile 為 disambig.exe)

How to execute

在 dsp_hw3_b04705036 根目錄下

make map (mapping.cpp 被 compile 為 mapping.exe 且被執行，因此 ZhuYin-Big5.map 應被產出在根目錄)

make run (disambig.exe 被執行，應在 result2 子目錄中看到產出的 \$\$i.txt)

What I have done

Makefile:

- 改 SRIPATH 為自己電腦 srilm-1.5.10 所在的資料夾路徑
- 新增 SRILIB 參數，表示 srilm library 所在路徑
- 改 MACHINE_TYPE 為 macosx
- 改 CXX 為 /usr/bin/g++ (macosx 系統適用)
- 在 all 中加入 -I 及 -L 指令讓他抓到 srilm 資料位置，因為雖然 vpath 跟 CXXFLAGS 已經有寫 lib 位置跟 -I(header 位置)，可是我用原本助教寫好的 makefile 一直 compile 出 error，所以才在 all 的地方再寫一次才 compile 成功

Train character-based bigram:

- 根據 ppt 上的 command 下指令用 corpus_seg.txt 訓練 bigram.lm

ZhuYin-Big5.map:

- 寫 mapping.cpp 把 Big5-ZhuYin.map 轉成 ZhuYin-Big5.map

Decode testdata(using srilm disambig):

- 根據 ppt 上的 command 下指令，用自己寫好的 ZhuYin-Big5.map 將斷句好的 testdata decode 至 result1 子目錄中

Viterbi decoding(using mydisambig.cpp):

- 寫 mydisambig.cpp 把斷句好的 testdata（含有注音文）轉換成文意合理的中文語句，產出在 result2 子目錄中