

What I observed

如果 segment the test file 之前，執行 mydisambiguation file 的話，很多<unk>出現，但執行 SRILM 的 disambiguation file 的話，不會出現。SRILM 的 disambig 是考慮過處理沒有 segmented 的 file 的。

用 mydisambiguation file 或 SRILM 的 disambiguation file 來 disambiguate 助教提供的 example.txt 的結果是一樣

What I have done

按照助教提供的資料，使用 SRILM。

寫好 mapping.py。

寫好 mydisambig.cpp。