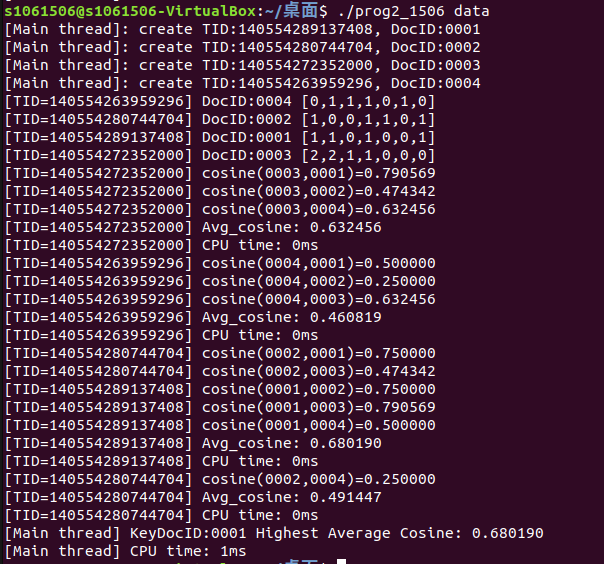
作業系統程式作業2

編譯方式：“g++ -o prog2\_1506 1061506\_02.cpp -std=c++17 -lpthread”

程式執行：“./prog2\_1506 (filename)”



Code由老師提供的thread-ex1.cpp修改而成，使用到了pthread的mutex來保持各個thread之間的同步。

Line 29~39: struct thdata

用來保存傳給各個thread資料的struct，main function後來判斷哪份文件是關鍵文件也是靠它還有它的比較運算子來判斷。

Line 49~67: function init

用來進行初始化的fucntion，在裡面初始化了所有的mutex，讀入了所有文件的資料並同時計算有幾份文件要處理。

Line 69~77: function create\_threads

幫助創建剛好數量thread的function。

Line 79~101: main function

* line 80~83: 檢查輸入的參數是否正確。
* line 89~90: 等待所有的thread計算完自己的平均餘弦相似係數。
* line 93~94: 尋找關鍵文件並印出。
* line 96~98: 印出Main thread使用的cpu time。

line 103~163: function article\_analyzer

用來讓每一條thread執行的任務。

* line 109~114: 分析文件並找出自己的詞頻(不完整)。
* line 116~122: 每個thread將自己出現過的字統一更新到Basis裡面。
* line 126~127: 從Basis中尋找，將沒出現在自己文章內的字加入詞頻向量並將出現次數設為0。
* line 128~142: 輸出自己處理的文件以及詞頻向量，並同時計算有多少thread完成了計算以及自身的向量長度。只有當所有thread都完成計算後，才去計算兩兩的相似度。
* line 146~155: 計算對其他文件的餘弦相似係數並輸出，同時也在計算對所有其他文件的平均餘弦相似係數。
* line 156~159: 輸出平均餘弦相似係數以及花費的cpu time。