

HW10: Longest Common Subsequence (LCS) 說明文件 v3

1. 輸入

程式的輸入為單一個文字檔；請設計能夠讓使用者輸入測資檔案名的功能，並根據使用者輸入的檔案名將測資讀入。**此次作業不需要實作防呆功能**，直接將使用者輸入的檔名用一個 string 存起來即可。一個文字檔包含有 2 行字串，表示 1 組（2 個）的 sequences。每個 sequence 由大小寫英文和數字所組成。

2. 輸出

此次作業的目標是將從先前讀入的 2 個 sequence 中之間「**所有**」的 **LCS**，以及 **LCS 的字串長度和總個數**。答案請輸出成文字檔，檔名須為「result_學號.txt」（例：result_E24089999.txt），並且檔案中的內容須遵循以下格式：

第 1 行 - LCS 的字串長度。（沒有 LCS 時視為 0）

第 2 行 - LCS 的總個數。（沒有 LCS 時視為 0）

第 3 行之後 - 所有的 LCS 字串；每行分別為一個 LCS。（沒有 LCS 時則保持空白行）

範例請參考投影片。

3. 評分&測資

此次作業評分主要將根據隱藏測資通過程度來評分。測資分為 3 個等級：

(1) 入門 (basic) - 輸入字串長小於 50 字；總共約有 1 千組 LCS。暴力解可能可以解出。總共有 1 組測資，全數通過即可得到 60 分。

(2) 進階 (advanced) - 輸入字串長小於 100 字；總共約有 1 萬組 LCS（**除了 advanced2_old.txt 之外**）。暴力解必定失敗，須使用較有效率的演算法。總共有 1 組測資，連同前面等級全數通過即可得到 100 分。

(3) 困難 (expert) - 輸入字串長超過 100 字；總共有 10 萬~30 萬組 LCS。除了較有效率的演算法外，須控制記憶體用量（Visual Studio 在 x86 環境記憶體上限為 2 GB，超過會造成 std::bad_alloc 的錯誤而使程式終止）。共有 1 組測資，連同前面等級全數通過即可得到 120 分。

入門以及進階測資須在 1 分鐘內得出結果，困難則須在 5 分鐘內得出結果。

附件中包含了 2 個範例、1 個入門、2 個進階等級的公開測資（非評分時測資）以及其對應的解答。解答中的 LCS 字串有經過排序（由 A 到 Z、由 0 到 9），同學可以將自己程式的結果進行比對，即可確認自己的答案是否正確。