

程式設計 (106-1)

作業十

作業設計：孔令傑
國立臺灣大學資訊管理學系

繳交作業時，請至 PDOGS (<http://pdogs.ntu.im/judge/>) 為第一題上傳一個 PDF 檔、為第二題做同儕互評，再為第三、四題各上傳一份 C++ 原始碼 (以複製貼上原始碼的方式上傳)。第四題是 bonus 加分題。每位學生都要上傳自己寫的解答。不接受紙本繳交；不接受遲交。請以英文或中文作答。

這份作業的截止時間是 **2017 年 12 月 5 日凌晨一點**。在你開始前，請閱讀課本的第 7.10 和 22.9–22.13 節¹。為這份作業設計測試資料並且提供解答的助教是林敬傑。

第一題

(20 分) 請看下面的程式：

```
#include <iostream>
#include <cstring>
using namespace std;

void reverse(char p[]);
int main()
{
    char s[100] = "12345";
    reverse(&s[1]);
    cout << s << "\n";
    return 0;
}

void reverse(char p[])
{
    int n = strlen(p);
    char* temp = new char[n];
    for(int i = 0; i < n; i++)
        temp[i] = p[n - i - 1];
    for(int i = 0; i < n; i++)
        p[i] = temp[i];
    delete [] temp;
}
```

¹課本是 Deitel and Deitel 著的 *C++ How to Program: Late Objects Version* 第七版。

- (a) (10 分) 如果 `reverse(&s[1])` 被改成 `reverse(&s[2])`，會印出什麼？為什麼？如果是被改成 `reverse(s)`，又會印出什麼？為什麼？為什麼要傳入 `s` 的時候不是寫 `&s`？寫 `&s` 為什麼會錯？
- (b) (10 分) 如果把 `char s[100] = "12345"` 改成 `char* s = "12345"`，這個程式是對是錯？有 `syntax error` 嗎？是否可能或一定會出現 `run-time error`？如果會，是出現在哪一句或哪幾句程式碼？為什麼？

第二題

(0 分) 請在 PDOGS 上批改你被隨機分配到的作業九第三題的程式碼，根據它在正確性以外的部份給它 1 至 5 分的評分，並且說明你給分的依據。建議在評分時參考以下六個面向。在前五個面向上，一個面向上做得好就得一分，還不錯則半分，不好則零分；在第六個面向上則在有必要時扣分。六個面向的分數合計後無條件進入即為你最後給的總分。

- 可讀性：變數與函數名稱是否具有合適的資訊量？程式碼排版是否良好且具有前後一致性？是否有合適的註解？關於註解，當然不需要每一行都有註解，但若你發現在某一大段落裡都沒有註解，或某個你感覺很不易看懂的部份沒有註解，你可以指出來；不要直接說「註解太少」但沒有說是哪邊缺乏註解。
- 模組化程度：是否有宣告合適的函數、`structure` 或 `class`²？是否有避免將非常類似的程式片段寫複數次而非寫成函數？是否有避免一個函數做非常多事情？函數間是否有合適的 `decoupling`？直接閱讀 `main function` 是否能很快地理解程式在大方向上的運算邏輯？
- 效率：程式運算是否有合理的運算效率？當然我們不要求每個同學都寫出超級有效率的精妙演算法，但至少一個程式不應該進行過多不必要的運算，也不應該耗用過多不必要的記憶體空間。如果你看不出這個程式的效率有明顯的問題，我們建議你直接給一分。
- 擴充性：當要解的問題變得更複雜的時候，我們能不能簡單地修改這個程式以解決新的問題，而不是寧可砍掉重練？這個議題當然也很主觀，所以如果你不能明確地指出在怎樣的新問題上，這個程式會有擴充性問題，我們建議你直接給一分；如果你不能指出很嚴重的問題，我們建議你至少給半分。但對批改者來說，這個關於擴充性的思考其實是很好的訓練。試試看吧！
- 其他：如果有任何其他令你想扣分的理由，請明確地寫出來並且在這個面向上扣分；沒有的話就給一分。
- 題目規範：你應該檢查那份程式碼有沒有違反題目的規範，如果有（例如題目說不可以用上課沒教過的東西，但他用了，或者題目說一定要用指標和動態記憶體配置，但他沒用），就扣他三分。當然，請明確地指出他哪邊違反了題目的規範。

本題其中 10 分取決於檢視你的程式碼的同學給你的分數總和（必要時助教會出來主持公道，請不用緊張），另外 10 分取決於你對同學的程式碼的評語和評分的合理性和建設性（原則上除非被申訴，且助教檢視後發現你確實評得很不公平，否則只要有評就會得到 10 分）。

²我們還沒教 `structure` 和 `class`，所以本次批改請看函數就好。

第三題

(60 分) 在許多 e-mail 系統中，若我們輸入一個字串去搜尋，系統會用許多條件來回傳符合條件的 e-mail。在本題中，我們將請你按照給定的規則實做一個搜尋功能。給定已經存在的既有字串 x_1 、 x_2 直到 x_N 以及使用者輸入的字串 y ，我們要幫既有字串依照與 y 的適配程度排序。假設既有字串和搜尋字串的頭尾都不含空白，我們排序的規則如下：

1. 如果 x_i 和 y 一模一樣，則適配度為 10000 分。
2. 如果 x_i 完整地包含 y (亦即 y 是 x_i 的子字串)，但 x_i 和 y 不完全相同，則適配度至少為 1000 分。若 y 出現在 x_i 中 n_i 次，則適配度為 $1000 + n_i$ 分。請注意 x_i 出現的複數次可以互相重疊，例如在 abcdecdecdefg 中，我們說 cdecdec 出現兩次，即使那兩次互相重疊。
3. 如果 x_i 不完整地包含 y ，但包含 y 的一個子字串 y_{sub} ，則適配度為第三等級。此時我們應計算最長的 y_{sub} 的字串長度，此字串長度乘以 10 即為適配度。

根據以上規則，請幫每個既有字串計算適配度分數。請注意大寫和小寫英文字元是不一樣的。

輸入輸出格式

系統會提供一共 20 組測試資料，每組測試資料裝在一個檔案裡。在每個檔案中會有 $N + 2$ 行，第一行包含一個整數 N ，第二行起的第 $i + 1$ 行是一個字串 x_i 。最後一行是搜尋字串 y 。已知 $1 \leq N \leq 1000$ 、所有字串都只含有最多 100 個英文字元、數字字元、空白字元和「, . ? ! : ;」這幾個標點符號。

讀入資料後，請根據題目規則幫每一個既有字串計算適配度分數，然後在同一行一一印出。兩個分數之間用一個逗點隔開。舉例來說，如果輸入是

```
5
This is
This is an apple
This is an apple. This is not a book.
Oh yeah!!!!
Today is his birthday
This is
```

則輸出應該是

```
10000,1001,1002,10,40
```

第一個字串和搜尋字串完全符合；第二個字串完整地包含了搜尋字串一次；第三個字串完整地包含了搜尋字串兩次；第四個字串包含的搜尋字串的子字串中最長的是「h」，長度為 1；第五個字串包含的搜尋字串的子字串中最長的是「his」，長度為 4 (最後面有個空白)。

你上傳的原始碼裡應該包含什麼

你的.cpp 原始碼檔案裡面應該包含讀取測試資料、做運算，以及輸出答案的 C++ 程式碼。當然，你應該寫適當的註解。針對這個題目，你**不可以**使用上課沒有教過的方法，但 `<cString>` 函式庫中所有的

功能你都可以用。

評分原則

- 這一題的其中 40 分會根據程式運算的正確性給分。PDOGS 會編譯並執行你的程式、輸入測試資料，並檢查輸出的答案的正確性。一筆測試資料佔 2 分。
- 這一題的其中 20 分會在作業十中被評定。屆時我們會讓同學們互相檢視彼此的本題程式碼，並且就可讀性、易維護性、模組化程度、排版等面向寫評語和給評分（當然一切都是匿名的）。該任務在本題中會佔 20 分，其中 10 分取決於檢視你的程式碼的同學給你的分數（必要時助教會出來主持公道，請不用緊張），另外 10 分取決於你對同學的程式碼的評語和評分的合理性和建設性。若你在本次作業中完全沒有寫這一題，那屆時自然沒有人能檢視你的程式碼，你也就得要損失這 10 分了。

第四題 (bonus)

(20 分) 承上題，現在有另一個適配度分數的來源。我們將既有字串 x_i 依照標點符號和空白鍵切開後，會得到一組詞（可以重複）；將搜尋字串 y 也同樣地切開後，會得到另一組詞（可以重複）。對第二組詞中的每一個詞，我們計算它出現在第一組詞的次數。將所有的這些次數加總並加進第三題計算之適配度分數，我們就能得到新的適配度分數。

本題的輸入、輸出格式和第三題一模一樣。舉例來說，如果輸入是

```
5
This is
This is an apple
This is an apple. This is not a book.
Oh yeah!!!!
Today is his birthday
This is
```

則輸出應該是

```
10002,1003,1006,10,41
```

請注意搜尋字串中的 **This** 出現在第三個字串中兩次，**is** 也出現兩次，所以第三個字串的適配度分數是 $1002 + 4 = 1006$ 。

針對這個題目，你**可以**使用任何方法。這一題的 20 分都根據程式運算的正確性給分，一筆測試資料佔 2 分。