**★★★☆☆**

**題組：Problem Set Archive** **with Online Judge**

**題號：11022: String Factoring**

**解題者：張辰瑜**

**解題日期：2021年5月4日**

**題意：**

輸入數個英文字符構成的字串，每一字串將相鄰重複的子字串表示$\left(子字串\right)^{n}$其中n為重複次數。

每一字符權重為1相鄰重複字串僅計一次，即其權重僅為該子字串的字符量，最後計算縮減過的字串最小權重。

**題意範例：**

$$PRATTATTATTIC⇒PR\left(A\left(T\right)^{2}\right)^{3}IC⇒6$$

$$G⇒\left(G\right)^{9}⇒1$$

$$PRIME⇒PRIME⇒5$$

$$BABBABABBABBA⇒\left(BAB\right)^{2}\left(A\left(B\right)^{2}\right)^{2}A⇒6$$

$$ARPARPARPARPAR⇒\left(ARP\right)^{4}AR⇒5$$

**解法：**

本題是dynamic program的類型，從一個字符開始增加字串長度求最小權重，因此取得字串總長為n宣告一$n×n$的矩陣，將對角線元素全設1表示單一字元的權重，以上全設極大值等程式運行時填充。

PRATTATTATTIC

取得權重的方式有三種，一是本就有如對角線或填充過的位址；二是由兩個較小子字串拼接，權重則為兩者相加取所有拼接方式中最小者。

三是由兩個較小子字串拼接後發現有相鄰重複的字符串，則計算將其縮減後是否小於前者。

$$A\left(T\right)^{2}$$

$$\left(A\left(T\right)^{2}\right)^{2}$$

$$A\left(T\right)^{2}\left(A\left(T\right)^{2}\right)^{2}$$

$$\left(A\left(T\right)^{2}\right)^{3}$$

2

綜合上述用遞迴的方式來解題，宣告函數並讓其知道現在處理的字串左右起點，討論dp矩陣內仍為極大值的情況。

先處理拆分方式，因遞迴性處理的字串長度會逐漸遞減，當dp矩陣內非最大值時不再遞迴，選擇最佳拆分此處對應第二種獲得權重的方式。

剛組合完的字串可能有重複元素應縮減，以字串的前i(當前字串長度的因數)個字元循環比對當前字串，若正確即可縮減，新權重取左起點+i-1位址的權重。

**解法範例：**

 以”APPLE”為例，將其轉換成$5×5$的矩陣，左矩陣為實際儲存的權重矩陣，右矩陣為字串拼接狀況的示意圖。

 長度為2的字串中僅”PP”需要特殊考慮其他可以直接加法運算，而”PP”經由解法的第三種取得權重方式也可以得到權重為1，再利用尋找最佳拼接方式進而得到”APP”與”APPL”的權重值。



 最後用同樣方式計算”PLE”與”PPLE”至此所有”APPLE”的拼接方式都有權重，尋找最佳權重並檢查”APPLE”有無相鄰重複字串，故答案輸出4。

**討論：**

該題屬於dynamic program難點偏向實作較不直觀的遞迴程式，將字串拆解成兩部分並尋找最小總和容易理解，較麻煩的是判斷重複或說因重新拼接後就得檢查，增加檢驗時間。

檢查重複能否縮減都是對當前完整字串，因在較小子串就已檢查過縮減與否，故較長的拼接字串也不會有原先小字串中未檢查到的縮減字串。

該演算法的時間複雜度應為${n\left(n+1\right)}/{2}$再乘上檢查所花費的$O\left(n\right)$即$O\left(n^{3}\right)$，而空間複雜度則是顯而易見的$O\left(n^{2}\right)$。