

DS_project_01

0413354 魏嘉豪

- 演算法說明

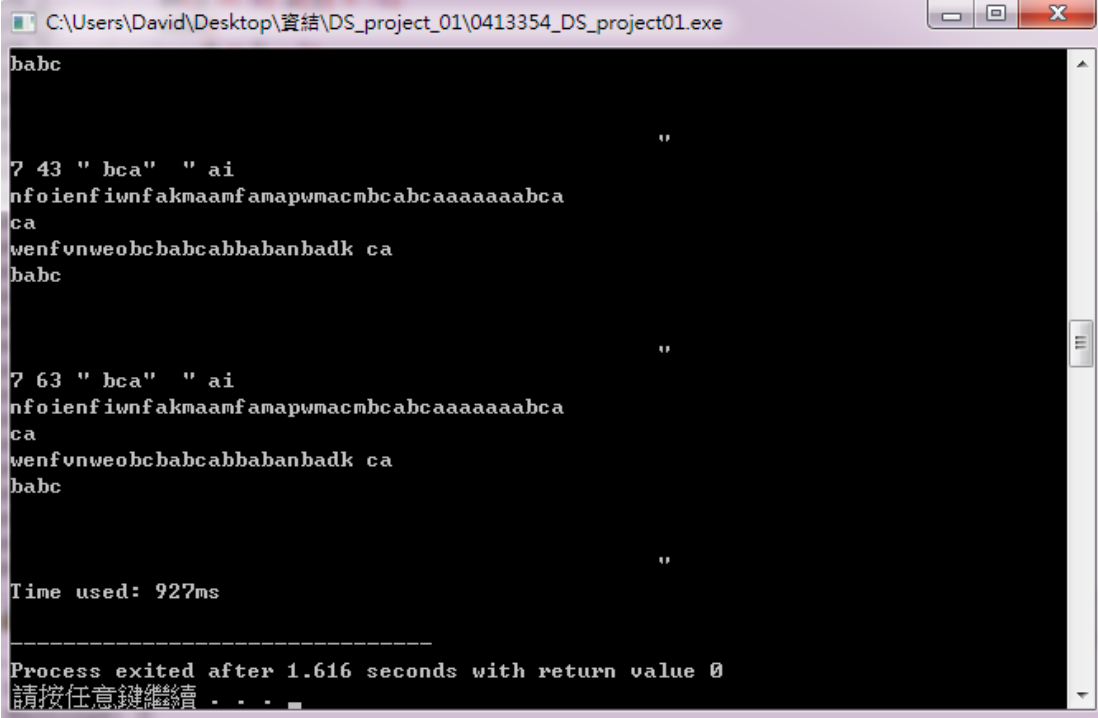
KMP 演算法的原理在當字元比對不符時，我們可以從比對字串 P 中得知某些字元可以略過不需檢查，如此就不必從文件 T 的下一個字元重新開始比對。

KMP 演算法的關鍵在於建構一個部分相符表格，以用來計算比對字串位移的格數，由於我們是在比對失敗時，才會參考這個表格，所以這個表格又被稱為失誤函數(failurefunction)。該函數定義一表格 next[n]，當比對失敗發生在P[j]時，可以拿P[next[n]]作為下一個重新比對的起點，以 P={a,b, c, d, a, d}為例，next 表格的內容為{-1, 0, 0, 0, 0, 1}。

在建構 failurefunction 時，當字元比對失敗在 P[j]，而且P[j-k]...P[j-1]正好等於P[0]...P[k-1]時，我們可以得到next[j]=k，

有了失誤表格，接著我們就可以繼續 KMP 搜尋演算法，接著我們將目的字串傳入，接著便可以傳入搜尋的 Pattern 字串 (P[]) 找尋目的字串是否含 Pattern 字串，如果有則傳回 Pattern 字串在目的字串中出現的第一個位置，否則傳回 -1。

kmp 實做:



```
C:\Users\David\Desktop\資料\DS_project_01\0413354_DS_project01.exe
babc

7 43 " bca" " ai
nfoienfiwnfakmaamfamapwmacnbcabcaaaaaabca
ca
wenfunweobchabcabbabanbadk ca
babc

7 63 " bca" " ai
nfoienfiwnfakmaamfamapwmacnbcabcaaaaaabca
ca
wenfunweobchabcabbabanbadk ca
babc

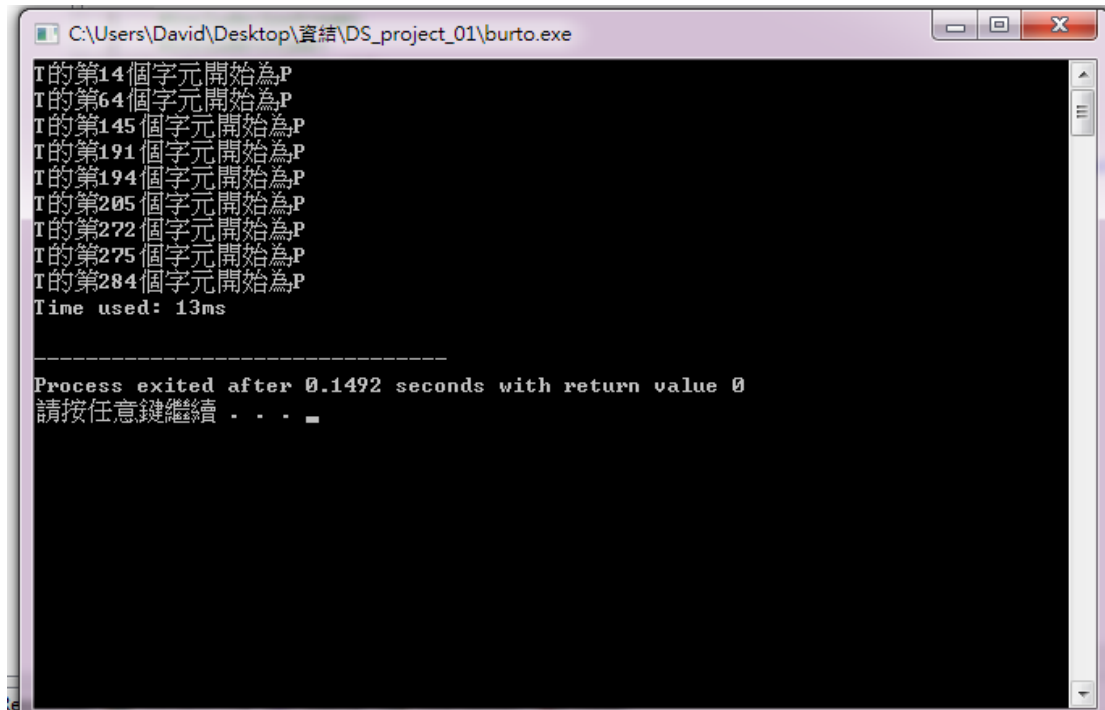
Time used: 927ms

-----
Process exited after 1.616 seconds with return value 0
請按任意鍵繼續 . . .
```

時間:927 ms

比較:

Brute-Force

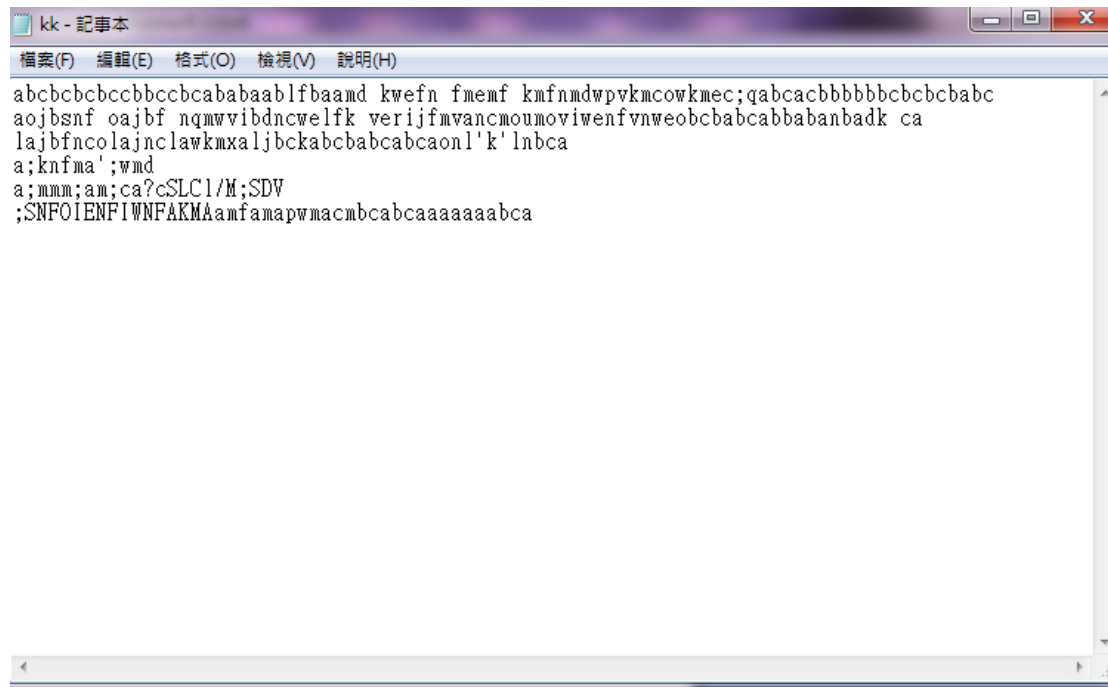


```
C:\Users\David\Desktop\資結\DS_project_01\burto.exe
T的第14個字元開始為P
T的第64個字元開始為P
T的第145個字元開始為P
T的第191個字元開始為P
T的第194個字元開始為P
T的第205個字元開始為P
T的第272個字元開始為P
T的第275個字元開始為P
T的第284個字元開始為P
Time used: 13ms

-----
Process exited after 0.1492 seconds with return value 0
請按任意鍵繼續 . . .
```

時間:13 ms

測資



```
kk - 記事本
檔案(F) 編輯(E) 格式(O) 檢視(V) 說明(H)
abcbcbcbccbbccbcabababablfbaamd kwefn fmemf kmfnmdwpvkmcowkmec;qabcacbbbbbbcbcbcbabc
aojbsnf oajbf nqmwvibdncwelfk verijfmvancmoumoviwenfvnweobcbabcabbabanbadk ca
lajbfncolajnclawkxaljbckabcbabcabcaonl'k'lnbca
a;knfma';wmd
a;mmn;am;ca?cSLC1/M;SDV
;SNFOIENFIWNFAKMAamfamapwmacmbcabcaaaaaaabca
```

Pattern

