1

Εστω string $S$ μεγεθους $n$ και υποψηφιο string $u$

Ενα υποψηφιο string $u$ είναι valid αν, για $C = \underbrace{u + u + \dots + u}_{X\ \text{φορες}}$ τα string $S$ και $C$ εχουν το ιδιο μεγεθος και $S_i \neq C_i$ για το πολυ ενα $i$.

## 1: Προταση

Αν το $u$ είναι valid, τοτε $n \,\%\, u.size == 0$

### Αποδειξη

Εστω $u$ είναι valid και $n \,\%\, u.size \neq 0 \Rightarrow$
$S.size() \neq C.size() \Rightarrow$
$u$ ΔΕΝ ειναι valid $\Rightarrow$ Ατοπο

Αρα $u$ ειναι valid $\Rightarrow n \,\%\, u.size == 0$
$\rightarrow$ αν για $u$ ισχυει $n \,\%\, u.size()$ θα το λεμε υποψηφιο" 4

## 2: Ορισμος

Εστω ενα string $S$ μεγεθους $n$ και εστω ενας αριθμος $h \leq n$. Ορισουμε το i'οστο h-Τμημα του $S$ ($T_h(S)_i$) την υποακολουθια του $S$ που περιλαμβάνει τους χαρακτήρες

$$[S[i \cdot h], S[i \cdot h + 1], \dots S[i \cdot h + h - 1]]$$

Παραδειγμα $S = $ "αααbbbccc", $n = 9$, $h = 3$

$T_3(S) = [$"ααα", "bbb", "ccc"$]$
και $T_3(S)_1 = $ "ααα"

Συνεπως το $T_h(S)_{ij}$ είναι ο $j$ χαρακτηρας του ιτή h-Τμηματος του string $S$.

## 3: Συνεπεια

για καθε $i \leq \frac{n}{h}$ ισχυει $(T_h(S)_i).size() == h$
i.e καθε h-Τμήμα έχει μέγεθος $h$.

## 4: Πρόταση

Αν το $u$ είναι valid, τοτε $u$ είναι υποακολουθία του $S$

### Απόδειξη για $h > 1$ ($u.size() = h$)

Εστω το $u$ είναι valid, και $u$ ΔΕΝ είναι υποακολουθία του $S$
και $C = u + u + u \dots + u$ και $S$ μεγεθους $n$ $\Rightarrow$
για οποιοδηποτε $i, j$ ισχυει $S_h(S)_{ij} \neq S_h(C)_{ij} \Rightarrow$

Αρα το $u$ δεν είναι valid (παραπανω απο 1 conflict)

### Αποδειξη για $h = 1$ ($u.size() = h$)

Εστω το $u$ είναι valid και $u$ ΔΕΝ είναι υποακολουθία του $S \Rightarrow$
αφου $h = 1$ τοτε ο χαρακτήρας εντος του $u$ δεν βρίσκεται στο $S$, $\Rightarrow$ αρα το $u$ δεν είναι valid.

## 5: Αλγοριθμος για να καθορισουμε αν ενα υποψηφιο $u$ ειναι valid

### 5.1: Υποθέσεις

1. Το υποψηφιο string $u$ είναι μεγεθους τετοιου ωστε $n \,\%\, u.size() == 0$ — Προταση 1
2. Το υποψήφιο string $u$ είναι υποακολουθια του $S$ — Προταση 4
3. καθε h-Τμήμα έχει μέγεθος $h$. — Συνεπεια 3

### 5.2: Διαδικασια

Εστω string $S$ μεγεθους $n$ και εστω $h \leq n$

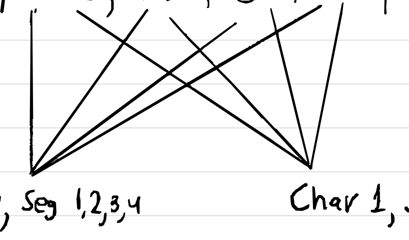
$$T_h(S) = [\text{"αα..."}, \text{"ab..."}, \text{"...."}, \dots, \text{"αα..."}]$$

για κάθε χαρακτηρα $i: 0 \rightarrow h$ και για καθε h-Τμήμα $j: 0 \rightarrow \frac{n}{h}$ υπάρχουν το πολυ 1 ζευγαρι $i$'s $(i_1, i_2)$ ετσι ωστε...

$$T_h(S)_{ji_1} \neq T_h(S)_{ji_2}$$

αν και μονο αν υπάρχει ενα valid $u$ εντος των h-Τμημάτων

$$T_h(S) = [\text{"αα"}, \text{"bb"}, \text{"cc"}, \text{"dd"}]$$



Char 0, Seg 1,2,3,4 Char 1, Seg 1,2,3,4

### Αποδειξη $\Leftarrow$

Εστω οτι υπαρχει ενα valid $u$ εντος των h-Τμηματων και
εστω οτι υπάρχουν παραπανω απο 1 ζευγαρια $i$'s $(i_A, i_B)$
ετσι ωστε $S_h(S)_{ji_1} \neq S_h(S)_{ji_2} \Rightarrow$
αρα υπάρχουν παραπάνω απο 1 conflicts μεταξυ $S$ και $C \Rightarrow$
Αρα δεν υπαρχει valid $u$ εντος των segments $\Rightarrow$ Ατοπο

### Αποδειξη $\Rightarrow$

Εστω οτι:
για κάθε χαρακτηρα $i: 0 \rightarrow h$ και για καθε h-Τμημα $j: 0 \rightarrow \frac{n}{h}$ υπάρχουν το πολυ 1 ζευγαρι $i$'s $(i_1, i_2)$ ετσι ωστε

$$T_h(S)_{ji_1} \neq T_h(S)_{ji_2}$$

και εστω οτι ΔΕΝ υπάρχει ενα valid $u$ εντος των h-Τμημάτων

$\Rightarrow$ Αρα υπαρχει ακριβως ενα conflict μεταξυ $S$ και $C$
$\Rightarrow$ Αρα υπαρχει ενα valid $u$ εντος των h-Τμημάτων
$\Rightarrow$ Ατοπο

### 5.3: Σημειωση

αν ισχυει το 2, αρκεί να επιστρεψουμε το $h$, καθως μας ενδιαφερει το μεγεθος του $u$ και οχι το $u$ καθ' αυτο

### 5.4: Ελαχιστοποιηση $h$

Αποδειξαμε οτι για γνωστο $h$ μπορουμε να ξερουμε αν υπαρχει valid $u$ of size $h$. Ποιο είναι το ελάχιστο $h$;

Τρεχουμε τον παραπάνω αλγοριθμο για καθε $i: 1 \rightarrow n$ και εξεταζουμε το $i$ μονο αν είναι valid ($n \% i == 0$)