

Assignment 5 (2 ส.ค. 65) : String กำหนดส่งงาน : จ. 15 ส.ค. 65 (เวลา 23.59 น.)

ให้นักศึกษา

- เขียนคำตอบตามโจทย์กำหนดด้วยลายมือ แล้วถ่ายรูป (นามสกุล .jpg) หรือไฟล์ pdf ส่งที่เว็บส่งการบ้านภาควิชา
- ตั้งชื่อไฟล์ในรูปแบบ assign_x_id เมื่อ x คือหมายเลข Assignment และ id คือ รหัสนักศึกษา
(กรณีส่งหลายไฟล์ให้ตั้งชื่อเป็น assign_01_id_a.jpg โดย a หมายถึง ลำดับไฟล์ แล้วทำการ zip รวมทุกไฟล์ส่งในงาน Assignment เดียวกันด้วยชื่อ assign_01_id.zip แทน)
- ส่งงานภายในวันเวลาที่กำหนด หากส่งเลยกำหนดให้ชี้แจงเหตุผลกับอ. ประจำ section (พิจารณาคะแนนตามเหตุผล)

กำหนดข้อความ T และ pattern P ดังนี้

ข้อความ T : aaababaabaababab

pattern P : aabab

ให้แสดงวิธีการค้นหา pattern P ในข้อความ T ด้วยวิธี

- Knuth-Morris-Pratt's algorithm ดัง ตัวอย่างในสไลด์หน้า 22
- Boyer-Moore's algorithm ดัง ตัวอย่างในสไลด์หน้า 28

Knuth-Morris-Pratt's algorithmi.) Build a prefix table ; letting $F[j]$ be the prefix of $P[j]$, $0 \leq j < |P|$.

j	0	1	2	3	4
$P[j]$	a	a	b	a	b
$F[j]$	0	1	0	1	0

• At $j=0$, $F[j] = 0$ • At $j=1$, let $t = F[j-1] = 0 \rightarrow P[t] = P[0] = "a"$ Since $P[j] = P[t] \leftrightarrow "a" = "a"$, $\therefore F[j] = t+1 = 1$ • At $j=2$, let $t = F[j-1] = 1 \rightarrow P[t] = P[1] = "a"$ Since $P[j] \neq P[t] \leftrightarrow "b" \neq "a"$ but $t \neq 0$,reconsider $t = F[t-1] = 0 \rightarrow P[t] = P[0] = "a"$ Since $P[j] \neq P[t] \leftrightarrow "b" \neq "a"$ and $t=0$, $\therefore F[j] = 0$ • At $j=3$, let $t = F[j-1] = 0 \rightarrow P[t] = P[0] = "a"$ Since $P[j] = P[t] \leftrightarrow "a" = "a"$, $\therefore F[j] = t+1 = 1$ • At $j=4$, let $t = F[j-1] = 1 \rightarrow P[t] = P[1] = "a"$ Since $P[j] \neq P[t] \leftrightarrow "b" \neq "a"$, but $t \neq 0$,reconsider $t = F[t-1] = 0 \rightarrow P[t] = P[0] = "a"$ Since $P[j] \neq P[t] \leftrightarrow "b" \neq "a"$, and $t=0$, $\therefore F[j] = 0$

□

ii.) Search for the pattern P in T

assisted by comparison
Section 2 sat 630510600

- Compare $P[0...4]$ to $T[0...4]$

T: $\overset{0}{a} \overset{1}{a} \overset{2}{a} \overset{3}{b} \overset{4}{a} \dots$

P: $a a b a b$

$\begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix}$

$F[j] = 0 \ 1 \ 0 \ 1 \ 0$

\therefore Shift 1 character.

Compare $P[0...4]$ to $T[1...5]$

- T: $\dots a a b a b \dots$

P: $a a b a b$

$\begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix}$

Match exactly!

\therefore Found P at $T[1...5]$

□

Boyer-Moore's Algorithm

- T: $\overset{0}{a} \overset{1}{a} \overset{2}{a} \overset{3}{b} \overset{4}{a} \overset{5}{b} \overset{6}{a} \overset{7}{a} \overset{8}{b} \overset{9}{a} \overset{10}{a} \dots$

P: $a a b a b$

bad character: Shift 1 character so $T_4 = P_3$ ($T_4 \neq P_4$) (case a.)

good suffix: Shift 1 character (case b.)

So that $T[1...4] = 'a a b a' = P[0...3]$

both heuristical rules suggest 1 shift, Found P at $T[1...5]$

□