

# מבני נתונים ואלגוריתמים – תרגיל 9

## שאלה 1 (25 נקודות)

דני מעוניין לחפש תבנית את כל ההופעות של תבנית  $P$  במחרוזת  $T$  המכילה תווים מיוחדים שישומונו בסימן שאלה (ב- $P$  אין סימני שאלה). סימן שאלה נחשב כמו אות כרצוננו באלף-בית של המחרוזות שלנו. לדוגמא, במחרוזת  $T=aba?aab?a$  מופיעה התבנית  $P=abaa$  באינדקס 0 כי אם ה-? הראשון היה  $a$  היינו מקבלים  $abaa$ . באותו אופן  $P$  מופיעה באינדקסים 2 (ה-? הראשון יכול להיות  $b$ ) ו-5 (ה-? השני יכול להיות  $a$ ).

לדני יש רעיון לאלגוריתם למצוא את כל ההופעות של  $P$  בתוך  $T$ . הוא יריץ אלגוריתם KMP עם שינוי קל: כל פעם כאשר הוא יצטרך להשוות בין  $T[i]$  ל- $P[q]$ , אם  $T[i] = '?'$  הוא יתייחס אל התווים כאילו הם שווים ואחרת הוא ישווה ביניהם כרגיל.

הראו על ידי דוגמא נגדית שהאלגוריתם של דני לא עובד. (כלומר, הוא לא מוצא את כל ההופעות של  $P$  בתוך  $T$ ).<sup>1</sup>

## שאלה 2 (25 נקודות)

חשבו את הטבלה  $C$  של אלגוריתם KMP ואת הטבלאות GSP ו-LSP של אלגוריתם BM עבור המחרוזת  $abaaabaaba$ .

## שאלה 3 (50 נקודות)

עבור מחרוזת  $X$  נגדיר את  $X^n$  להיות המחרוזת המתקבלת משרשור  $n$  עותקים של  $X$ . לכל מחרוזת  $Y$  נגדיר את  $\rho(Y)$  להיות ה- $n$  המקסימלי עבורו קיים  $X$  כך ש- $Y = X^n$ . (לדוגמא:  $\rho(abab) = 2$ ,  $\rho(abaa) = 4$  ו- $\rho(aba) = 1$ . שימו לב ש- $\rho(Y) \geq 1$  לכל  $Y$ ).

**סימון:** עבור מחרוזת  $T$  נגדיר את  $T_i$  להיות  $T[0]T[1] \dots T[i-1]$  (הרישא באורך  $i$  של  $T$ ).

א. נניח כי  $X$  מחרוזת באורך  $k$  ויהי  $n|k$  כך ש- $X_{k-\frac{k}{n}}$  סיפא של  $X$ . הוכיחו שקיימת  $Y$  כך ש-

$$Y^n = X$$

ב. תהי  $X$  מחרוזת באורך  $k$  ונניח ש- $\rho(X) > 1$ . הראו כי ה- $0 \leq i < k$  המקסימלי עבורו

$X_i$  הוא סיפא של  $X$  הוא  $i = k - \frac{k}{n}$ . [הדרכה: כתבו  $X = Y^n$ . הראו שאם הטענה שגויה אז

$$\rho(Y) > 1 \text{ כדי לעשות זאת העזרו בלמה בסוף.}]$$

ג. כתבו אלגוריתם המקבל מחרוזת  $T$  ומדפיס את  $\rho(T_i)$  לכל  $0 < i \leq \text{Length}(T)$ . על

האלגוריתם לעבוד ב- $O(\text{Length}(T))$  פעולות. הוכיחו כי האלגוריתם עובד והצדיקו את סיבוכיות הזמן שלו.

בפתרון אתם רשאים להעזר בלמה הבאה:

**למה:** תהי  $T$  מחרוזת ונניח כי יש מחרוזות לא ריקות  $A, B$  כך ש- $T = AB = BA$ , אזי  $\rho(T) > 1$ .

<sup>1</sup> פסאודו קוד של KMP נמצא באתר.

### שאלת בונים (15 נקודות)

מחרוזת  $T$  נקראת כפולה אם  $T = XX$  עבור מחרוזת  $X$  כלשהי. כתבו אלגוריתם המקבל מחרוזת  $T$  ומחזיר את ה- $i$  המקסימלי כך שהרישא  $T_i := T[0]T[1] \dots T[i-1]$  היא מחרוזת כפולה. על האלגוריתם לעבוד בסיבוכיות זמן  $O(\text{Length}(T))$  בממוצע ועם  $O(1)$  זיכרון בלבד (אם מתעלמים מהזיכרון הדרוש לאחסון הקלט). הסבירו בקצרה מדוע האלגוריתם עובד, והצדיקו את סיבוכיות הזמן והזיכרון שלו.