

מבני נתונים – תרגיל 5

תאריך פרסום: 11.06.2023

תאריך הגשה: 02.07.2023 בשעה 23:59 (בשל השביתה ישנה הארכה כוללת)

מתרגל אחראי: תומר מאירמן

נושאים: Heaps, Sorts, Graphs

הנחיות:

- יש לקרוא הנחיות לגבי הגשת עבודות באתר הקורס ולהגיש עבודות על פי הנחיות אלו.
- אין צורך לפרט על הגדרות או משפטים שנלמדו בכיתה. עם זאת, יש להוכיח כל טענה שלא נלמדה בהרצאה או בתרגול.
- שאלות לגבי העבודה יש לשאול בפורום באתר הקורס או בשעות קבלה של המתרגל האחראי.
- בתיאור האלגוריתם יש להקפיד על השלבים הבאים:
 - הסבירו במילים את הרעיון הכללי.
 - כתבו את האלגוריתם בצורה ברורה, הקפידו על אינדנטציה ברורה, מספרו את השלבים השונים. לא חובה לצרף פסאודו-קוד, ניתן לתאר את השלבים בשפה פשוטה. יש לעשות זאת בצורה ברורה שאינה משתמעת לשני פנים. במידה ובחרתם לתאר פסאודו-קוד, ציינו מחוץ למסגרת של הפסאודו-קוד מה כל משתנה שהגדרתם מייצג.
 - הביאו הסבר/נימוק מדוע האלגוריתם אכן פותר את הבעיה.
 - נתחו את סיבוכיות האלגוריתם מבחינת זמן ריצה וזיכרון.
- בתיאור מבנה נתונים יש להקפיד על השלבים הבאים:
 - תארו את מבנה הנתונים במילים באופן מדויק.
 - יש להציג תרשים של מבנה הנתונים הכולל את החלקים השונים והקשרים ביניהם.
 - אם משתמשים במבנה נתונים שנלמד בכיתה "as is", ניתן להשתמש בו כקופסה שחורה ולהסביר באופן כללי כיצד הפעולות מתבצעות. אולם, שינויים במימוש הנלמד בכיתה מצריכים הסבר מדויק של השינויים שבוצעו.

שאלה 1 (20 נק')

עלינו ליצור מבנה נתונים יעיל לניהול איברים, כך שלכל אחד מהם יש מפתח ייחודי, ומאפשר את הפעולות שמופיעות בטבלה הבאה. יש לשים לב שמבנה הנתונים יהיה יעיל ככל הניתן בהתאם לתדירות השימוש. כלומר - הפעולה היעילה ביותר תהייה עבור השימוש התדיר ביותר, ופעולות תדירות פחות יכולות להיות ביעילות נמוכה יותר. היעילות הנדרשת אינה נתונה.

שם פעולה	תקציר	תדירות השימוש
FindMin()	מציאת האיבר בעל מפתח מינימלי	תדיר ביותר
FindMax()	מציאת האיבר בעל מפתח מקסימלי	תדיר ביותר
DeleteMin()	מחיקת האיבר עם המפתח המינימלי	תדירות בינונית
DeleteMax()	מחיקת האיבר עם המפתח המקסימלי	תדירות בינונית
Insert(k)	הוספת האיבר בעל מפתח k	תדירות בינונית
DeleteOldest()	מחיקת האיבר הישן ביותר במבנה הנתונים	תדירות בינונית
Search(k)	חיפוש האם האיבר בעל מפתח k נמצא במבנה הנתונים	תדירות בינונית

הניחו ש- n הוא מספר האיברים במבנה בכל רגע נתון.

יש לתת תיאור מילולי קצר של מבנה הנתונים (מומלץ להוסיף תרשים). לאחר מכן, יש לספק אלגוריתם (במילים או בפסאודו-קוד) לכל אחת מהשיטות ולנתח בקצרה את זמני הריצה של הפעולות במונחים של n .

שאלה 2 (20 נק')

- יהי $G = (V, E)$ גרף מכוון, חסר מעגלים (DAG - directed acyclic graph). יהיו s, t שני קודקודים בגרף כך שקיים מסלול מ- s לכל קודקוד אחר בגרף, וקיים מסלול מכל קודקוד בגרף לקודקוד t .
- א. אם נבצע מיון טופולוגי על הגרף הנתון G , מה ניתן לומר בוודאות על מיקום קודקודי s ו- t לאחר המיון? הוכיחו.
- ב. תארו אלגוריתם אשר מקבל גרף $G = (V, E)$ מכוון חסר מעגלים (DAG - directed acyclic graph), (בייצוג של רשימת סמיכויות (שכנויות) ומחזיר "כן" אם קיימים בגרף קודקודים s, t כמתואר לעיל. מה הוא זמן הריצה של האלגוריתם שתיארתם?

שאלה 3 (20 נק')

נתונה מטריצה A עם m שורות ו- n עמודות שבה הערכים בכל שורה ממוינים בסדר עולה וערכיהם שלמים וחסומים ב $-100m \leq k \leq 100m$.

כמו כן, ידוע כי הערכים בעמודה ה- i גדולים או שווים מכל הערכים בעמודה ה- $i-1$, לכל $2 \leq i \leq n$. הציעו אלגוריתם, בזמן ריצה של $O(nm)$ ו- $O(m)$ זיכרון נוסף, להדפסת איברי המטריצה בסדר עולה. לדוגמה: בהינתן המטריצה הבאה בעלת 3 שורות ו-4 עמודות

	1	2	3	4
	-5	8	15	20
	2	4	9	30
	3	9	15	16

עליכם להדפיס:

-5, 2, 3, 4, 8, 9, 9, 15, 15, 16, 20, 30

שאלה 4 (20 נק')

א. נתון T -ש- T היא מחרוזת של 100 תווים מעל אלפבית של 5 תווים, כאשר כל אות מהאלפבית מופיעה לפחות פעם אחת ב- T . מהו האורך הקצר ביותר שניתן לקידוד של T ע"י קוד הופמן? ציירו את העץ המתאים לקידוד קצר ביותר של T וציינו את אורך הקידוד.

ב. הוכיחו או הפריכו את הטענה הבאה:

נתון קוד הופמן עבור סדרה של שכיחויות ותווים. יהיו a, b שני תווים כך שעבור כל קידוד אופטימלי אורך מילת הקוד של a קטן ממש מאורך מילת הקוד של b . מבאן: השכיחות של a קטנה ממש מהשכיחות של b .

ג. הוכיחו או הפריכו את הטענה הבאה:

נתון קוד הופמן עבור סדרה של שכיחויות ותווים. יהיו a, b שני תווים עם אותו אורך של מילת קוד. אז השכיחות של a שווה לשכיחות של b .

שאלה 5 (20 נק')

במעבדות של אינטל מפתחים מעבד חדש.

במעבד קיימים n רכיבים לוגיים המחוברים אחד לשני.

במעבדות של אינטל החליטו כי הם רוצים לעקוב אחר החיבור בין כל הרכיבים, ולכן יצרו את הטבלה הבאה (לדוגמה עבור 4 רכיבים):

	C_1	C_2	C_3	C_4	קידוד
C_1	1	0	0	0	1000
C_2	0	0	1	0	0010
C_3	1	1	0	0	1100
C_4	1	0	1	0	1010

כל תא בטבלה C_i, C_j מציין:

1 – אם שני הרכיבים C_i ו- C_j מחוברים

0 – אם שני הרכיבים אינם מחוברים

כל רכיב מקבל קידוד ע"פ השורה שלו (העמודה האחרונה במטריצה)

כדי ללמוד על החיבוריות של הרכיבים, המדענים צריכים למיין את הרכיבים בסדר עולה הנקבע על פי החיבור שלהם עם שאר הרכיבים. עבור הדוגמה למעלה, מיון הרכיבים יהיה:

0010, 1000, 1010, 1100

נתון מערך של n מחרוזות המייצגות את הקידוד של n הרכיבים.

א) עזרו למדענים לבחור את המיון היעיל ביותר עבור הרכיבים מהאופציות הנתונות מטה. עבור כל אחד מהמיונים מטה, נתחו את זמן הריצה וסיבוכיות המקום, והסבירו איך יתבצע המיון.

1. מיון מנייה.

2. מיון בסיס.

3. מיון מבוסס השוואות.

ב) כעת המדענים רצו לבדוק האם k רכיבים בלבד מחוברים לכל שאר הרכיבים (n רכיבים).

1. כיצד תראה הטבלה המייצגת את הקידודים כעת (יש להציג דוגמה כללית לטבלה, הערכים בפועל אינם משנים)?

2. עבור ערכי k הבאים הציגו את זמן הריצה עבור המיונים מסעיף א':

$$k = O(n) \quad a.$$

$$k = O(\log n) \quad b.$$