הנדסת תוכנה

Software Engineering

תרגיל 4 להגשה במבני נתונים (קורס מס' 10117)

מרצים : ד"ר ראובן חוטובלי וד"ר דוד שטטר

תאריך הגשה : 16.1.2021 עד השעה 23:30 . **העבודה- בזוגות**.

חובה לענות על כל השאלות ולתת את כל התשובות על גבי גיליון התשובות המפורסמת בלבד.

שאלה 1

בשאלה זו 15 סעיפים (שאלות סגורות). עליך לענות על כל הסעיפים. בכל סעיף נתונות ארבע תשובות, שרק אחת מהן נכונה. בכל סעיף בחר את התשובה הנכונה, וסמן אותה ב- X בדף התשובות שבנספח.

סעיפים א' – ד' מתייחסים לבעיה שלהלן:

יש להציע מבנה נתונים המתחזק את S כאשר S הינה איחוד של m קבוצות זרות,

כלומר $S=\bigcup_{i=1}^{m} S_i$. כל S_i , $\forall i=1,2,...,m$, הינה קבוצה בת n מפתחות לכל היותר.

ברור שכל המפתחות ב-S שונים זה מזה. על המבנה נתונים לתמוך בפעולות הבאות :

פעולה	תיאור
Insert(i,k)	הוספת מפתח k לקבוצה S_i .
Delete(i,k)	מחיקת מפתח k מהקבוצה S_i . ניתן להניח כי k שייך לקבוצה S_i .
Min_Max()	מציאת המפתח המינימלי מבין המפתחות המקסימליים M_i $\forall i=1,2,....,m$, כאשר M_i הוא מפתח בעל ערך מקסימלי בקבוצה S_i $\forall i=1,2,....,m$.
GlobalMax()	מציאת המפתח המקסימלי מבין כל המפתחות שב - S.

סטודנט הציע את מבנה הנתונים הבא כפתרון לביצוע **יעיל** של כל הפעולות שלעיל .

הערה: המבנה יכול להיות מורכב מכמה מבנים פשוטים יותר.

להלן הצעת הסטודנט:

מבנה הנתונים המתחזק קבוצה S של מספרים ממשיים שונים מורכב מ :

1. עץ AVL1 שהוא עץ חיפוש בינארי מאוזן (עץ AVL) הממוין לפי האינדקסים של הקבוצות.

כל צומת x בעץ AVL1 מייצג את אחת הקבוצות S_i וצומת זה מכיל:

- שדה key אשר מכיל מספר, שהוא אינדקס של קבוצה מבין הערכים: $1, 2, ..., m$
- שדה left - מצביע (מחוון) לבן שמאלי .
- שדה right - מצביע (מחוון) לבן ימני.
- שדה point - מצביע (מחוון) לשורש של עץ AVL אחר שהוא עץ חיפוש בינארי מאוזן (עץ AVL) בו משוכנים איברים השייכים לאותה קבוצה. כלומר, אם השדה key מכיל ערך i אז שדה point יצביע לשורש של עץ AVL אחר בו משוכנים כל האיברים השייכים לקבוצה - S_i בלבד.

2. עץ AVL2 שהוא עץ חיפוש בינארי מאוזן (עץ AVL) המכיל רק האיברים המקסימליים בכל קבוצה. כלומר עץ AVL2 יכיל את: $M_1, M_2, ..., M_m$ כאשר M_i הוא מפתח בעל ערך מקסימלי בקבוצה S_i $\forall i = 1, 2, ..., m$.
3. משתנה MMax –מכיל את הערך המקסימלי מבין כל האיברים שנמצאים בעץ AVL2.
4. משתנה GMax –מכיל את הערך המינימלי מבין כל האיברים שנמצאים בעץ AVL2.

<u>א</u>. על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Insert(i,k) או Delete(i,k)?

1. $O(\log n)$ בלבד.
2. $O(\log m)$ בלבד.
3. $O(m \log n)$
4. $O(\log n + \log m)$

<u>ב</u>. על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Min_Max() ?

1. $O(\log n)$
2. $O(\log m)$
3. $O(1)$
4. $O(\log n + \log m)$

<u>ג</u>. על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה GlobalMax() ?

1. $O(\log n)$
2. $O(1)$
3. $O(\log m)$
4. $O(\log n + \log m)$

ד. בפעולת Insert(i, k) יש לבצע:

תלמיד א' מציע:

צעד 1: הכנס את i בעץ AVL1 והצומת שהתווסף זה עתה לעץ AVL1 נסמנו ב- p.

צעד 2: הכנס את k בעץ AVL שלשורשו מצביע השדה point של הצומת p **בלבד**.

תלמיד ב' מציע:

צעד 1: הכנס את i בעץ AVL1 והצומת שהתווסף זה עתה לעץ AVL1 נסמנו ב- p.

צעד 2: הכנס את k בעץ AVL שלשורשו מצביע השדה point של הצומת.

צעד 3: אם k הוא המקסימום בקבוצה S_i אז נמחק מהעץ AVL2 את M_i כאשר M_i הוא מפתח בעל ערך מקסימלי בקבוצה S_i.

איזו הצעה מבין שתי ההצעות מממשת בשלמות את הפעולה Insert(i, k)?

1. אף אחת מהן.
2. של תלמיד א' בלבד.
3. של תלמיד ב' בלבד.
4. שתי ההצעות.

סעיפים ה' – י' מתייחסים לבעיה שלהלן:

חברת הבניה "מבנים בע"מ" רוצה לשמור את נתוני משכורות העובדים ולבצע מספר פעולות עליהם.
נניח שנתוני העובדים נמצאים בקבוצה S והחברה מעסיקה n עובדים, כלומר |S|=n. יש להציע מבנה נתונים בגודל O(n) התומך בפעולות הבאות:

פעולה	תיאור
Init(employees,median_key3)	אתחול המבנה. Employees היא רשימה לא ממוינת של n העובדים והמשכורות שלהם. median_key3 הוא **השלישון** של משכורות העובדים – ידוע ונתון מראש ומשתנה ביחד עם העדכונים. (**ראו הגדרה של השלישון בהמשך**).
Insert(k)	הוספת עובד בעל משכורת k למבנה הנתונים.
Remove_Min()	הוצאת עובד בעל משכורת מינימלית. אם יש יותר מאחד, אז יש להוציא אחד מהם.
Remove_Max()	הוצאת עובד בעל משכורת מקסימלית. אם יש יותר מאחד, אז יש להוציא אחד מהם.
Average()	הדפסת המשכורת הממוצעת של העובדים בחברה.
Median3()	הדפסת **השלישון** של משכורת העובדים בחברה (**ראו הגדרה של השלישון בהמשך**).

הגדרה: **השלישון** שווה לערך המופיע במקום ה- $\left\lceil \frac{n}{3} \right\rceil$ לאחר מיון האיברים, כלומר האיבר ה- $\left\lceil \frac{n}{3} \right\rceil$ בגודלו (ערך המיקום ה- $\left\lceil \frac{n}{3} \right\rceil$).

סטודנט הציע את מבנה הנתונים הבא כפתרון לביצוע **יעיל** של כל הפעולות שלעיל:

- לשמור ארבע ערמות :

  - Heap1 - ערמת מקסימום אשר תכיל נתונים של כל העובדים.
  - Heap2 - ערמת מינימום אשר תכיל נתונים של כל העובדים.
  - Heap3 - ערימת מקסימום אשר תכיל נתונים של $\left\lceil \frac{n}{3} \right\rceil$ העובדים. ערמה זו תכיל את $\left\lceil \frac{n}{3} \right\rceil$ האיברים הקטנים ביותר.
  - Heap4 - ערמת מינימום אשר תכיל נתונים של $\left(n-\left\lceil \frac{n}{3} \right\rceil\right)$ העובדים.
  - כל איבר שנמצא ב- Heap3 קטן מכל איבר שנמצא ב- Heap4 .
- משתנה size אשר יכיל את מספר העובדים.
- Salary_sum אשר יכיל את סכום המשכורות של כלל העובדים.
- מבנה של עובד יכיל שדות בסיסיים ונוסף לכך הוא יכיל גם נתונים אודות :
  - מיקומו בערמת המקסימום - Heap1 .
  - מיקומו בערמת המינימום - Heap2 .
  - מידע באיזו ערמה הוא נמצא מבין Heap3 או Heap4 .
  - מיקומו ב- Heap3 או Heap4 , היכן שהוא שייך.

שימו לב : בפעולות ,Insert או Delete בהכנסה / הוצאה של איבר , אם יש צורך מעבירים איבר מערמה אחת לאחרת (מ- Heap3 ל- Heap4 או להיפך) כדי לשמור על דרישותיו של הסטודנט.

על סמך מבנה הנתונים שהציע הסטודנט , ענה על השאלות שבסעיפים ה׳- ח׳.

ה. בחר את ההיגד **הנכון** מבין ההיגדים הבאים :

1. **השלישון** יהיה תמיד האיבר המקסימלי, מבין האיברים שקטנים משורשו של הערמה Heap1.
2. **השלישון** יהיה תמיד האיבר המקסימלי, מבין האיברים שגדולים משורשו של הערמה Heap2.
3. **השלישון** יהיה תמיד האיבר המינימלי של הערמה Heap4.
4. **השלישון** יהיה תמיד האיבר המקסימלי של הערמה Heap3.

ו.

על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Median3() ?

1. $O(\log n)$

2. $O(\log\log n)$
3. $O(1)$
4. $O(\log^2 n)$

ז. על סמך מבנה הנתונים שהיציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Insert(k) ?
1. $O(\log\log n)$
2. $O(\log n)$
3. $O(1)$
4. $O(\log^2 n)$

ח. על סמך מבנה הנתונים שהיציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Remove_Min() ?
1. $O(\log\log n)$
2. $O(\log n)$
3. $O(1)$
4. $O(\log^2 n)$

ט. על סמך מבנה הנתונים שהיציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Init(employees,median_key3) ?
1. $O(n)$
2. $O(\log n)$
3. $O(1)$
4. $O(\log^2 n)$

י. על סמך מבנה הנתונים שהיציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Average() ?
1. $O(1)$
2. $O(\log n)$
3. $O(n)$
4. $O(\log^2 n)$

בסעיפים י"א' – ט"ו' התייחס למימוש הפעולה Init(employees,median_key3) שלהלן:

צעד 0 : קליטת הנתונים ל- S וקביעת ערכם של size ושל Salary_sum.
צעד 1 : העתק את אברי הקבוצה S למערך עזר A1 .
צעד 2 : העתק את אברי הקבוצה S למערך עזר A2 .
צעד 3 : __________(1)_______
צעד 4 : __________(2)_______
צעד 5 : temp=______(3)_______
צעד 6 : partition(S,temp,____(4)_____ ,______(5)_______)
צעד 7 : בנה ערמה Heap3 מהאיברים המשוכנים במערך A3 .
צעד 8 : בנה ערמה Heap4 מהאיברים המשוכנים במערך A4 .

במימוש הזה חסרים **חמישה** ביטויים המסומנים במספרים בין סוגריים עגולים. התשובה הנכונה עבור כל אחד מהביטויים החסרים מופיעה בסעיפים הבאים :

יא. התשובה הנכונה עבור ביטוי (1) לעיל היא:

1. בנה ערמה Heap1 מאברי המערך A1 ובעבור כל עובד שמור את מיקומו בערמה זו.
2. העתק את אברי הקבוצה S למערך עזר A3.
3. בצע פעולת partition ביחס לערך הראשון שבמערך עזר A1.
4. בצע פעולת partition ביחס לערך שמיקומו $\left\lceil \frac{n}{3} \right\rceil$ שבמערך עזר A1.

יב. התשובה הנכונה עבור ביטוי (2) לעיל היא:

1. בנה ערמה Heap2 ובעבור כל עובד שמור את מיקומו בערמה זו.
2. העתק את אברי הקבוצה S למערך עזר A4.
3. בצע פעולת partition ביחס לערך הראשון שבמערך עזר A2.
4. בצע פעולת partition ביחס לערך שמיקומו $\left\lceil \frac{n}{3} \right\rceil$ שבמערך עזר A2.

יג. התשובה הנכונה עבור ביטוי (3) לעיל היא:

1. $select(n, S)$
2. $select\left(\left\lceil \frac{n}{3} \right\rceil, S\right)$
3. $select\left(\left\lceil \frac{n}{2} \right\rceil, S\right)$
4. $select(1, S)$

יד. התשובה הנכונה עבור ביטוי (4) לעיל היא:

1. A2
2. A3
3. A1
4. S

טו. התשובה הנכונה עבור ביטוי (5) לעיל היא:

1. A1
2. A2
3. A4
4. S

<u>שאלה 2</u>

<u>בסעיפים א – ד' התייחס לבעיה שלהלן:</u>

במפעל לייצור מכוניות ניתן לכל מכונית היוצאת מקו הייצור **מספר מזהה ייחודי** . בנוסף לכל מכונית נרשם **תאריך הייצור שלה**.

הניחו שבכל רגע נתון יש n מכוניות במבנה.

יש להציע מבנה נתונים התומך בפעולות שלהלן :

שם פעולה	תיאור
Init()	אתחול מבנה ריק.
Insert(k , d)	הוספת מכונית חדשה למבנה עם מזהה k ותאריך ייצור d .
HowMany (d)	מחזירה את מספר המכוניות שתאריך הייצור שלהם הוא d .
Print(k1,k2)	הדפסת כל המכוניות שהמזהים שלהם גדולים או שווים ל- k1 וקטנים או שווים ל- k2 . הניחו ש : k1 < k2 .

סטודנט א' הציע את מבנה הנתונים הבא כפתרון לביצוע **יעיל** של כל הפעולות שלעיל .

הערה: המבנה יכול להיות מורכב מכמה מבנים פשוטים יותר וגם כל תאריך ניתן לייצגו כמספר שלם.

להלן הצעת הסטודנט:

מבנה הנתונים מורכב מ :

1. טבלת גיבוב (ערבול) (Hashing) . בטבלה זו נשמור מונים למספר המכוניות שנוצרו בתאריך מסוים.
2. עץ AVL1 שהוא עץ חיפוש בינארי מאוזן (עץ AVL) . **בעץ זה נשמור את תאריכי הייצור**.
   כל צומת בעץ AVL1 מייצג תאריך הייצור וצומת זה מכיל:
   - שדה key אשר מייצג תאריך הייצור.
   - שדה left - מצביע (מחוון) לבן שמאלי .
   - שדה right - מצביע (מחוון) לבן ימני.
   - שדה point - מצביע (מחוון) למונה בטבלת הגיבוב - **(Hashing)**, המציין מספר מכוניות שנוצרו בתאריך מסוים המיוצג על ידי צומת זה בעץ.
3. עץ AVL2 שהוא עץ חיפוש בינארי מאוזן (עץ AVL) . **בעץ זה נשמור את המזהים של המכוניות**.
4. רשימה דו-מקושרת המייצגת את סדר המכוניות לפי המזהים שלהן בעץ AVL2 של המזהים.

בנוסף, יש **הצבעה הדדית** בין המזהים ב- AVL2 לצמתים התואמים ברשימה הדו מקושרת.

א.

על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Insert(k , d) בממוצע?

1. $\Theta(\log n)$
2. $\Theta(\log\log n)$
3. $\Theta(1)$
4. $\Theta(\log^2 n)$

ב.

על סמך מבנה הנתונים שהציע הסטודנט , מהי **תוחלת** זמן הריצה למימוש הפעולה HowMany (d) ?

1. $\Theta(\log n)$
2. $\Theta(1)$
3. $\Theta(\log\log n)$
4. $\Theta(\log^2 n)$

ג.

על סמך מבנה הנתונים שהציע הסטודנט , מהי סיבוכיות זמן הריצה למימוש הפעולה Print(k1,k2) ?

1. $\Theta(d)$ בהכרח, כאשר d הוא מספר המכוניות להדפסה.
2. $\Theta(1)$
3. $\Theta(\log n + d)$, כאשר d הוא מספר המכוניות להדפסה.
4. $\Theta(\log\log n + d)$, כאשר d הוא מספר המכוניות להדפסה.

ד.

בעבור הפעולה Insert(k , d) ישנן כמה הצעות ואלו הן :

תלמיד א' מציע:

צעד 1: הכנס את k בעץ AVL2 .

צעד 2: הכנס את d בעץ AVL1 .

תלמיד ב' מציע:

צעד 1 : הכנס את k בעץ AVL2 ונעדכן את מקומו בתוך הרשימה הדו מקושרת.

צעד 2 : חפש את d בעץ AVL1 .

- אם d קיים בעץ AVL1 אז עדכן את המונה המתאים בטבלת הגיבוב **(Hashing)**
- אם d לא קיים בעץ AVL1 אז :
  - הוסף את d לעץ AVL1 ואתחל את המונה שלו ל- 1 בטבלת הגיבוב **(Hashing)**.
  - שמור מצביע מהתאריך בעץ AVL1 למונה בטבלת הגיבוב **(Hashing)**

איזו הצעה מבין שתי ההצעות מממשת בשלמות את הפעולה Insert(k , d) ?

1. אף אחת מהן.
2. של תלמיד א' בלבד.
3. של תלמיד ב' בלבד.
4. שתי ההצעות.

ה.

נתונה קבוצה S של מספרים שלמים ויש להציע מבני נתונים התומך בפעולות שלהלן:

שם פעולה		**זמן ריצה נדרש**
Insert(x)	הוספת x למבנה הנתונים.	$O(n)$ במקרה הגרוע
Search(x)	חיפוש אחר x במבנה הנתונים.	$O(1)$ זמן צפוי (ממוצע)
Delete(x)	מחיקת x ממבנה הנתונים.	$O(1)$ זמן צפוי (ממוצע)
Successor(x)	מחזירה את העוקב ל- x	$O(1)$ זמן צפוי (ממוצע)
Predecessor(x)	מחזירה את הקודם ל- x	$O(1)$ זמן צפוי (ממוצע)

מבני הנתונים הנדרש הוא:

1. טבלת גיבוב בלבד.
2. טבלת גיבוב ועץ AVL. בנוסף, בעבור $\forall x \in S$ נחזיק מצביעים הדדיים בין הצומת בטבלת גיבוב המכיל את x לצומת בעץ AVL המכיל את x.
3. טבלת גיבוב ורשימה דו מקושרת ממוינת. בנוסף, בעבור $\forall x \in S$ נחזיק מצביעים הדדיים בין הצומת בטבלת גיבוב המכיל את x לצומת ברשימה הדו מקושרת הממוינת המכיל את x.
4. טבלת גיבוב וערמה. בנוסף, בעבור $\forall x \in S$ נחזיק מצביעים הדדיים בין הצומת בטבלת גיבוב המכיל את x לצומת בערמה המכילה את x.

ו'.

מריצים את אלגוריתם QuickSort. בכל שלב איבר הציר נבחר להיות החציון. אבל, את החציון מוצאים על ידי פונקציה חדשה בשם NewSelect, שזמן הריצה שלה במקרה הגרוע על מערך בגודל n הוא $O\left(n^{3/2}\right)$.

מהי סיבוכיות זמן הריצה של QuickSort במקרה זה?

1. $O\left(n^{3/2}\lg n\right)$
2. $O\left(n\lg n\right)$
3. $O\left(n^{2}\right)$
4. $O\left(n^{3/2}\right)$

שאלה 3

בסעיפים א–ד' התייחס לבעיה שלהלן:

קלט: מערך A בגודל n.

שאלה: האם יש במערך A 2 ערכים **שמספר המופעים** של שניהם יחד הוא **בדיוק** 2009.

לפניך אלגוריתם שפותר את הבעיה הנתונה בזמן הקצר ביותר **בממוצע**.

אלגוריתם

1. עבור כל איבר $y \in Y$ בצע:
   1.1 הכנס את y לתוך מבנה הנתונים ___(1)_______
2. נעתיק את האיברים המשוכנים לתוך מבנה הנתונים ___(1)_______ למערך B.
   הנח כי כל תא במערך B הינו מבנה בעל 2 שדות, כאשר שדה אחד, בשם evar,
   מכיל אחד הערכים של המערך A הנתון, והשדה השני, בשם counter, מכיל
   את שכיחותו של הערך evar.
   ברור כי: במערך B – לא יהיו שני תאים **שונים** שהשדה evar שלהם יכילו את
   אותו ערך.
3. נמיין את המערך B ע"י ______(2)_______
4. נריץ את האלגוריתם ______(3)______ שבאמצעותו נדע האם יש במערך A 2
   ערכים שמספר המופעים של שניהם יחד הוא **בדיוק** 2009.

באלגוריתם הנ"ל חסרים ארבעה ביטויים, המסומנים במספרים בין סוגריים עגולים.
התשובה הנכונה עבור כל אחד מהביטויים החסרים מופיעה בשאלות הבאות:

<u>א.</u>

התשובה הנכונה עבור ביטוי (1) לעיל היא:

1. עץ חיפוש בינארי. ערכים כפולים נשמור באותו צומת בעץ על ידי מונה.
   כלומר, כל צומת מכיל ערך נתון במערך ואת שכיחותו (מספר המופעים שלו במערך). בעץ
   חיפוש בינארי – לא יהיו שני צמתים **שונים** שיכילו את אותו ערך.
2. טבלת Hash בגודל **n** (פתרון התנגשויות על ידי רשימות מקושרות). ערכים כפולים נשמור
   באותו איבר ברשימה על ידי מונה. ברשימה הזו – לא יהיו שני צמתים **שונים** שיכילו את
   אותו ערך. כלומר, כל צומת ברשימה מכיל ערך נתון במערך ואת שכיחותו (מספר המופעים
   שלו במערך).
3. רשימה דו-כיוונית ממויינת. ערכים כפולים נשמור באותו צומת ברשימה על ידי מונה.
   כלומר, כל צומת ברשימה הזו מכיל ערך נתון במערך ואת שכיחותו (מספר המופעים שלו
   במערך). ברשימה – לא יהיו שני צמתים **שונים** שיכילו את אותו ערך.
4. עץ מאוזן AVL. ערכים כפולים נשמור באותו צומת בעץ על ידי מונה.
   כלומר, כל צומת מכיל ערך נתון במערך ואת שכיחותו (מספר המופעים שלו במערך). בעץ
   המאוזן (AVL) – לא יהיו שני צמתים **שונים** שיכילו את אותו ערך.

<u>ב.</u>

התשובה הנכונה עבור ביטוי (2) לעיל היא:

1. מיון מהיר (Quick Sort) לפי השדה counter.
2. מיון מניה (Counting Sort) לפי השדה evar.

3. מיון מניה (Counting Sort) לפי השדה counter .
4. מיון ערמה (Heap Sort) לפי השדה counter .

ג.

התשובה הנכונה בעבור ביטוי (3) לעיל היא :

1. למציאת 2 מספרים שסכומם 2009 במערך ממוין B לפי השדה evar .
2. למציאת 2 מספרים שהממוצע שלהם 2009 במערך ממוין B לפי השדה counter.
3. למציאת 2 מספרים שסכומם 2009 במערך ממוין B לפי השדה evar , תוך שימוש בחיפוש בינארי.
4. למציאת 2 מספרים שסכומם 2009 במערך ממוין B לפי השדה counter .

ד.

סיבוכיות זמן הריצה של האלגוריתם הנתון היא בממוצע :

1. $\Theta(\lg n)$
2. $\Theta(n\lg n)$
3. $\Theta(n)$
4. $\Theta(n^2)$

בהצלחה