אוניברסיטת תל אביב

סמסטר ב' תשע"ט

**מבני נתונים - פרויקט מספר 2 – בעיית הבחירה**

**בעיית הבחירה-תזכורת**

בהינתן n פריטים עם מפתחות עם יחס סדר מלא. יש למצוא את האיבר עם המפתח הk בגודלו (כלומר, המפתח הקטן ביותר כך שיש עוד k-1 פריטים בעלי מפתחות הקטנים ממנו(

מטרת הפרויקט היא לממש מספר פתרונות לבעיית הבחירה, לחקור ולנתח את ההבדלים בין הפתרונות השונים ביחס לn,k ולהציג את המסקנות.

1. :selectRandQuickSort(int [] ,int k) מיון n הפריטים ע"י האלגוריתם rand-quicksort והחזרת האיבר עם האינדקס הk.
2. :selectInsertionSort(int [] ,int k)מיון n הפריטים ע"י האלגוריתם Insertion sort והחזרת האיבר עם האינדקס הk.
3. :selectHeap(int [] ,int k) יצירת ערמת מינימום בינארית מn הפריטים וביצוע k פעמים deleteMin .
4. :selectDoubleHeap(int [] ,int k) יצירת ערמת מינימום בינארית מn הפריטים + ערמת עזר שגודלה אינו עולה על k בסיבוכיות ,בדומה לתרגיל בית 3 שאלה 7.
5. : randQuickSelect(int [] ,int k) ע"י האלגוריתם Randomized-QuickSelect כפי שנלמד בהרצאה.
6. medOfMedQuickSelect(int [] ,int k) ע"י האלגוריתם MedOfMed-QuickSelect כפי שנלמד בהרצאה.

**דרישות**

1. עליכם לממש עבור כל בעיה פונקציה המקבלת מערך בגודל n ואינדקס k . בנוסף לפתרון הבעיה הפונקציות צריכות לספור את כמות ההשוואות שמבצע כל אלגוריתם על המערך(השוואה נספרת רק בין 2 איברי קלט. השוואות אחרות , כמו השוואה בין אינדקסים, לא נחשבת) .   
   על הפונקציה להחזיר pair (מספר השוואות שבוצעו, האיבר הk בגודלו) . כלומר, ברכיב הראשון את האיבר הk בגודלו וברכיב השני את מספר ההשוואות שבוצעו.  
   שמות הפונקציות הם השמות שהוגדרו.
2. **ניסויים:** נגדיר n=i\*10,000 k=1+j\*n/10,   
   כתבו תוכנית (אין צורך להגיש אותה) שתבצע את הפעולות הבאות:
   1. לכל i מ1 ל 10:
      1. הגרל מערך A של מספרים טבעיים בתחום [0,10,000] בגודל n .
      2. לכל j מ0 עד 10 בצע:
         1. ספור עבור כל אחת מהפונקציות\* את מספר פעולות ההשוואה שבוצעו עבור A ו k .

\* randQuickSelect ו selectRandQuickSort הם אינם אלגוריתמים דטרמיניסטיים ולכן עבורם במקום הרצה אחת בצעו מספר הרצות בכל פעם. מספר פעולות ההשוואה יהיה הממוצע על כל אחת מההרצות האלו.

כלומר, n רץ מ 10,000 עד ל 100,000 בקפיצות של 10,000. עבור כל n הערך של k ירוץ מ 1 ועד ל n בקפיצות של n/10

1. הסבירו את התוצאות שקיבלתם, אל מול הניתוח התיאורטי של זמן הריצה של כל האלגוריתמים.  
   בפרט, התייחסו לשתי השאלות הבאות:

שאלה 1: מהי סיבוכיות זמן הריצה של כל אחד מהאלגוריתמים כתלות ב- n וב- k ? התייחסו הן

למקרה הגרוע הן למקרה הטוב. התייחסו אף למקרה הממוצע, כאשר עבור אלגוריתמים דטרמיניסטיים הממוצע הוא ממוצע על הקלט ועבור אלגוריתמים אקראיים הממוצע הוא ממוצע על ההגרלות. ציינו תוצאות שמתייחסות לתוחלת זמן הריצה שנלמדו בכיתה, ובמקרים שלא נלמדו, הציגו הסבר משלכם (אין צורך בניתוח הסתברותי מדויק).

שאלה 2: כיצד הניתוח התיאורטי מתיישב עם התוצאות האמפיריות (ההרצות)? נסו להסביר את

המקורות האפשריים להבדלים, אם קיימים.

באופן כללי, נסו להסביר את התוצאות שלכם ככל יכולתכם. מותר להעלות השערות, גם אם אתם

לא בטוחים בנכונותן, כל עוד הן הגיוניות ומנומקות היטב.

אם לדעתכם המסקנות אינן חד משמעיות, או שברצונכם לשפר את מהימנותן, בצעו הרצות נוספות . שיכולות לסייע לכם.

**סיבוכיות**

יש לתעד בקוד ובמסמך נפרד (ביותר פירוט) את סיבוכיות זמן הריצה במקרה הגרוע ובממוצע (האסימפטוטית, במונחי O הדוקים) של כל בעיה, כתלות בn וk. עליכם להשיג סיבוכיות זמן ריצה (במקרה הגרוע ביותר) נמוכה ככל הניתן עבור כל אחת מהפונקציות.

**פלט**

אין צורך בפלט למשתמש.

**תיעוד**

בנוסף לבדיקות אוטומטיות של הקוד שלכם, קובץ המקור ייבדק גם באופן ידני. חשוב להקפיד על תיעוד לכל פונקציה ,אם פונקציה קוראת לפונקציית עזר, יש להתייחס גם לפונקציית העזר בניתוח.

וכמות סבירה של הערות. הקוד צריך להיות קריא, בפרט הקפידו על בחירת שמות משתנים ועל אורך השורות.

**בדיקות**

התרגילים ייבדקו באמצעות תוכנת טסטר שקוראת לפונקציות המפורטות מעלה, ומוודאת את נכונות התוצאות. קובץ הטסטר שלנו **לא יפורסם** לפני הבדיקות. עליכם לבדוק את המימוש בעצמכם! בפרט, כדאי מאוד לממש טסטר, כדי לבדוק את תקינות ונכונות המימוש.

בקובץ שתגישו לא תהיה פונקציית main (דבר זה יפגע בטסטר שיבדוק לכם את התרגילים). אם הצלחתם לקמפל את הפרוייקט לבדו (ללא טסטר), זה סימן שמשהו לא נכון במימוש שלכם.

הקוד ייבדק על מחשבי בית הספר על גירסא Java8.

הנחיות להשמשת סביבת העבודה בבית (ג'אווה+אקליפס):

<http://courses.cs.tau.ac.il/software1/1415b/misc/workenv.pdf>

מדריך לעבודה עם Eclipse (סעיפים 5-9, 15(:

[http://www.vogella.com/](http://www.vogella.com/tutorials/Eclipse/article.html)

הנחיות לפתיחת חשבון מחשב, למי שמעוניינ/ת לעבוד במעבדת בית הספר:

<http://cs.tau.ac.il/system/accounts0>

שימוש בג'אווה 8 במעבדות האוניברסיטה:

<http://courses.cs.tau.ac.il/software1/1415b/misc/lab-eclipse.pdf>

**הגשה**

הגשת התרגיל תתבצע באופן אלקטרוני באתר הקורס במודל.

**הגשת התרגיל היא בזוגות בלבד!**

כל זוג ייבחר נציג **אחד** ויעלה תחת שם המשתמש שלו את קבצי התרגיל (תחת קובץ zip) למודל. על ההגשה לכלול שלושה קבצים:  
קובץ בו ממומשות הפונקציות כפי שתוארו תחת השם selectProblems.java.  
קובץ טקסט info.txt המכיל את פרטי המגישים הבאים: תז, שמות ושמות משתמש.  
מסמך המציג את תוצאות הניסויים ומסקנותיהם כפי שמפורט בהמשך. את המסמך יש להגיש בפורמט pdf.

שמות קובץ תוצאות הניסויים וקובץ הzip צריכים לכלול את שמות המשתמש האוניברסיטאיים של **שני המגישים** לפי הפורמט Select\_username1\_username2.pdf/doc/zip/…. בתוכן הקבצים יש לציין את שמות המשתמש, תעודות הזהות ושמות המגישים (בכותרת המסמך ובשורת הערה בקובץ המקור).

הגשת שיעורי הבית באיחור - באישור מראש בלבד. הגשה באיחור ללא אישור תגרור הורדת נקודות מהציון.

הגשת התרגיל היא חובה לשם קבלת ציון בקורס.

**מסמך הצגת תוצאות הניסויים**

אורכו של הקובץ לא יעלה על 4 עמודים.

הציון על חלק זה תלוי באיכות הכתוב, ולא באורכו. עשו מאמץ להציג את הדברים בצורה מדויקת וקולעת,

מבלי להשמיט פרטים או הסברים חשובים מצד אחד, אך מבלי לכלול פירוט מיותר או לא רלוונטי מצד

שני.

חלק זה יכלול:

- כותרת מתאימה

- תקציר (abstract) - פסקה או שתיים בתחילת המסמך, המסכמות את מטרת הפרויקט ותוצאותיו,

מבלי להיכנס לפרטים. על התקציר לאפשר לקורא, שאיננו מכיר כלל את הפרויקט, להבין

בכלליות ובמהירות מה מטרת הפרויקט ומה המסקנות העיקריות שלכם.

- הצגת תוצאות ההרצות בצורה ברורה ונוחה, שמאפשרת הבנה והסקת מסקנות בקלות (בחרו

את הדרכים שנראות לכם מתאימות – טבלאות, עקומות, שילוב כלשהו שלהן וכו’).

- התייחסות לשתי השאלות בעמוד הקודם.

- התייחסות להרצות נוספות, אם ביצעתם כאלו, והתובנות שעלו מהן, או כל התייחסות נוספת

מעבר לדרישות המינימום, שנראית לכם נחוצה (חלק זה אינו חובה).

**בהצלחה!**