

מבני נתונים – פרויקט מספר 2 - ערמה d-ארית

דרישות

עליכם לממש ערימה d-ארית (d-ary heap). לכל צומת לכל היותר d ילדים. כל הרמות מלאות, חוץ אולי מהאחרונה, שמלאה מצד שמאל. הערימה מיוצגת באמצעות מערך. השורש של הערימה במקום 0. ילדים של צומת במקום i, נמצאים באינדקסים $d*i+1, \dots, d*i+d$. הערימה שקולה לערימה בינארית עבור $d=2$.

כל שאר הפעולות (insert, get-min, del-min, decrease-key, delete), הינן כמו אלו שנלמדו בכיתה עבור ערימה בינארית.

האיברים בערמה הם מסוג DHeap_Item ומכילים 3 שדות:

1. מפתח key, מספר שלם לא ייחודי (כלומר אותו מספר יכול לחזור מספר פעמים).
2. ערך name, מסוג מחרוזת.
3. מיקום pos מסוג int, שמצביע למיקום של האיבר במערך של הערימה.

שימו-לב, שבפעולות השונות תצטרכו לעדכן את שדה pos של DHeap_Item בהתאם לתזוזת האיבר בערימה. לא ניתן לשנות את ההגדרה של המחלקה DHeap_Item.

בערימה יש ארבעה data members (ניתן להוסיף עוד אם צריך):

- array, מערך שאיבריו מסוג DHeap_Item.
- size, שמחזיק את מספר האיברים הנוכחי בערימה.
- max_size, שהוא מספר האיברים המקסימלי המותר בערימה.
- d הוא הפרמטר של הערימה (מספר הילדים המקסימלי האפשרי של איבר בערימה).

ניתן להניח כי לא נעבור את מס' האיברים המקסימלי. הבנאי כבר ניתן לכם בקובץ השלד.

המימוש צריך להיות מבוסס על קובץ השלד המופיע באתר הקורס ולהסתמך על כל ההנחות שרשומות שם. הפעולות שמופיעות בקובץ הן כדלקמן:

getSize() - מחזירה את מספר האיברים בערימה.

arrayToHeap(array1) - הפונקציה מקבלת מערך מסוג DHeap_Item ובונה ערימה חדשה מהמערך (יש להתעלם ממה שהיה קודם בערימה). הפונקציה מחזירה את מספר ההשוואות שנעשו בתהליך.

isHeap() - מחזירה ערך TRUE אם ורק אם הערמה המיוצגת במערך חוקית.

parent(i,d), child(i,k,d) – מחזירות את מיקום ההורה, ואת מיקום הילד מספר k של האיבר שמיקומו i בערימה d-ארית.

Insert(item) - הכנסת האיבר item (מסוג DHeap_Item) לערימה. החזרת מספר ההשוואות שבוצעו במהלך ההכנסה.

Delete_Min() - מחיקת איבר המינימום של הערמה. החזרת מספר ההשוואות שבוצעו במהלך המחיקה.

Get_min() - החזרת איבר המינימום של הערימה.

Decrease_Key(item, delta) - הקטנת המפתח של האיבר item בשיעור של delta. הניחו כי delta אי-שלילי. החזרת מספר ההשוואות שהתבצעו במהלך הפעולה.

Delete(item) - מחיקת האיבר item מהערימה. ניתן להניח כי האיבר item נמצא בערימה (כלומר ערך pos שלו הוא חוקי ביחס לערימה). החזרת מספר ההשוואות שהתבצעו במהלך המחיקה.

DHeapSort(int[] array1, int d) – ממין את מערך הקלט array1 באמצעות ערימה d-ארי, כלומר HeapSort (כמו שנלמד בכיתה). לאחר ההרצה המערך array1 צריך להיות ממין. החזרת מספר ההשוואות שבוצעו במהלך המיין.

בקובץ השלד מופיעים ה-header-ים של כל הפונקציות. המימוש יבוצע על ידי מילוי קובץ השלד. במידת הצורך ניתן להרחיב את המימוש (למשל להוסיף פונקציות עזר שלא מופיעות בשלד או data members נוספים למחלקה), אך אסור לשנות את ההגדרות של הפונקציות לעיל. אסור לשנות את קובץ ה-DHeap_Item.

סיבוכיות

יש לתעד את סיבוכיות זמן הריצה (במקרה הגרוע) של כל פונקציה, כתלות במספר האיברים בעץ ובפרמטר d. עליכם להשיג סיבוכיות זמן ריצה הטובה ביותר מבחינת זמן ריצה אסימפטוטי.

פלט

אין צורך באף פלט למשתמש.

תיעוד

קובץ המקור ייבדק גם באופן ידני. חשוב להקפיד על תיעוד לכל פונקציה, וכמות סבירה של הערות. הקוד צריך להיות קריא, בפרט הקפידו על בחירת שמות משתנים ועל אורך השורות. לקוד המקור יצורף מסמך תיעוד חיצוני. המסמך יכלול את תיאור המחלקה שמומשה, ואת תפקידו של כל חבר במחלקה. עבור כל מתודה במחלקה יש לפרט מה היא עושה, כיצד היא פועלת ומה סיבוכיות זמן הריצה שלה. בפרט, אם פונקציה קוראת לפונקציית עזר, יש להתייחס גם לפונקציית העזר בניתוח.

בדיקות

התרגילים יבדקו באמצעות תוכנת טסטר שקוראת לפונקציות המפורטות למעלה, ומוודאת את נכונות התוצאות. קובץ הטסטר שלנו לא יפורסם לפני הבדיקות. עליכם לבדוק את המימוש בעצמכם! בפרט, כדאי מאד לממש טסטר, כדי לבדוק את נכונות ותקינות המימוש.

בקובץ שתגישו לא תהיה פונקציית main. אם הצלחתם לקמפל את הפרויקט לבדו, ללא טסטר, זה סימן שמשוה לא נכון במימוש.

הקוד ייבדק על מחשבי בית הספר על גרסה Java8 – ראו קישורים רלוונטיים בהנחיות לתרגיל מעשי 1.

מדידות

כתבו תוכנית (אין צורך להגיש אותה) שתפעיל את הפעולות שמימשתם, וענו בעזרתה על השאלה הבאה.

הכינו מערך עם m מספרים רנדומיים שלמים מפוזרים אחיד בין 0 ל-1000. הכניסו את הערכים לערימה בעזרת `arrayToHeap` (המחרוזת של ה-Item לא משנה). כעת, בצעו מיון על-ידי הפונקציה `DHeapSort`. ספרו כמה השוואות יתבצעו בכל המיון (עליכם להתחשב גם בפעולה של `arrayToHeap`), כלומר כמה פעמים היו השוואות בין שני איברים במיקומים שלהם במערך (תזכורת: ההשוואות מתבצעות בפעולת מחיקה במהלך `heapify-down`).

בצעו את הספירה עבור $m = 1,000, 10,000, 100,000$ ועל כל אחד עבור $d = 2, 3, 4$. הציגו את התוצאות. כמו כן, נתחו (באופן תיאורטי) מה מספר ההשוואות האסימפטוטי בחסם הדוק (חסם תטא) במקרה הגרוע כתלות במספר האיברים n ובפרמטר d .

כעת, בצעו את התהליך הבא: הכינו מערך רנדומי זהה כמו קודם (הפעם עם 100,000 איברים) והכניסו את האיברים לערימה. בצעו `Decrease-Key` על כל האיברים לפי סדר הכנסתם לערימה. בצעו `Decrease-Key` עם דלתא בערך x . ספרו כמה השוואות התבצעו בכל פעולות ה-`Decrease-Key`.

בצעו את הספירה עבור $x = 1, 100, 1000$ ועל כל אחד עבור $d = 2, 3, 4$. הציגו את התוצאות. כמו כן, נתחו מה מספר ההשוואות האסימפטוטי בחסם הדוק (חסם תטא) במקרה הגרוע כתלות במספר האיברים n ובפרמטר d .

עבור כל מדידה – בצעו אותה 10 פעמים. הציגו את הממוצע של 10 המדידות.

הגשה

הגשת התרגיל תתבצע באופן אלקטרוני באתר הקורס במודל.

הגשת התרגיל היא בזוגות בלבד!

כל זוג יבחר נציג אחד ויעלה תחת שם המשתמש שלו את קבצי התרגיל למודל בקובץ `zip` יחיד. על ההגשה לכלול שני קבצים: קובץ המקור (הרחבה של קובץ השלד `DHeap.java`), ומסמך תיעוד חיצוני, המכיל גם את תוצאות המדידות. את המסמך יש להגיש באחד הפורמטים הבאים: `txt`, `rtf`, `doc`, `docx` או `pdf`.

שם קובץ ה-`zip` צריך לכלול את תעודות הזהות של המגישים. 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.