

20.7.2015

מבני נתונים
פתרון תרגיל מס' 10

1.

הצע/הציעי מבנה נתונים העונה לדרישות הבאות (כאשר n מציין את מספר האיברים במבנה הנתונים):

- הוספת איבר למבנה מבוצעת בזמן $O(n)$ במקרה הגרוע.
- בהינתן מפתח של האיבר (שמזהה את האיבר באופן יחיד) חיפוש האיבר והוצאתו ממבנה הנתונים מבוצעת בזמן $O(1)$ בממוצע.
- עבור k כלשהו מציאת k האיברים בעלי המפתח הגדול ביותר בזמן $O(k)$ במקרה הגרוע.

בנוסף להצעת מבנה הנתונים, תאר/י באופן מילולי איך מתבצעות כל הפעולות הנ"ל.

הערה: בפתרונות של שתי השאלות בתרגיל בית זה, כאשר מתוארת פעולה ללא פרוט הכוונה שהפרוט הוא לפי האלגוריתם שהוצג בכיתה. לדוגמה כאשר כתוב "מוסיפים את האיבר ל- hash" הכוונה שמוספים את האיבר ל- hash לפי האלגוריתם שהוצג בכיתה.

פתרון

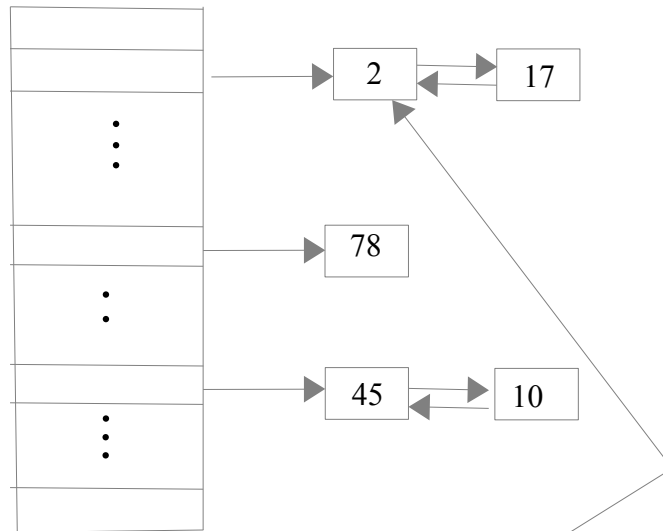
מבנה הנתונים המוצע הוא טבלת hash שמכילה את האיברים ובנוסף רשימה מקושרת דו כוונת ממוינת עם זנב של מיצגים של האיברים שבטבלת ה-hash. מכל איבר ברשימה המקושרת יש מצביע דו כוונת לאיבר המתאים לו ב-hash. הרשימה המקושרת ממוינת לפי המפתחות של האיברים בטבלת ה-hash בסדר עולה.

לדוגמה, נניח שבמבנה הנתונים יש 5 איברים שהמפתחות שלהם הם:

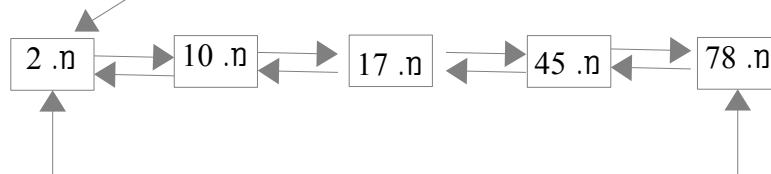
10, 45, 2, 17, 78

אזי מבנה הנתונים המוצע יראה כך:

טבלת hash



רשימה מקושרת ממוינת



הוספת איבר למבנה הנתונים מתבצעת באופן הבא, תחילה מוספים את האיבר ל- hash בזמן $O(1)$ במקרה הגרוע. ולאחר מכן מוסיפים את המיצג של האיבר לרשימה המקושרת הממוינת בזמן $O(n)$ במקרה הגרוע.

חיפוש איבר והוצאתו מתבצע באופן הבא: תחילה מחפשים את האיבר ב- hash בזמן $O(1)$ בממוצע בעזרת המצביע מגיעים למיצג של האיבר ברשימה הממוינת ומוציאים אותו מהרשימה בזמן $O(1)$ במקרה הגרוע. בסך הכל סיבוכיות הזמן של הפעולה היא $O(1)$ בממוצע.

בהינתן k כלשהו, מציאת k האיברים בעלי המפתח הגדול ביותר מתבצעת באופן הבא: בעזרת הזנב ניגשים לסוף הרשימה המקושרת ומדפיסים את k האיברים שנמצאים בסוף הרשימה המקושרת. זמן הפעולה $O(k)$ במקרה הגרוע.

2. שאלה זו הופיעה במבחן מועד א בקיץ 2006

זמני הסיבוכיות בשאלה זו מתייחסים לסיבוכיות זמן במקרה גרוע ביותר.

בשאלה זו מניחים שהאיברים במבנה הנתונים צבועים בצבע אדום או לבן. לכל איבר x במבנה הנתונים ישנו שדה בשם $color(x)$ שערכו הוא red או $white$ אם האיבר צבוע בצבע אדום או לבן בהתאמה. כמו כן, ניתן להניח שלכל איבר מפתח ייחודי, כלומר אין שני איברים בעלי מפתח זהה.

הצע/הציעי מבנה נתונים שתומך בפעולות הבאות (מניחים שמספר האיברים הנוכחי הוא n):

- הכנסת איבר חדש בזמן $O(\log n)$;
- חיפוש איבר בזמן $O(\log n)$;
- הוצאת איבר בזמן $O(\log n)$ (כשמקבלים מצביע לאיבר שרוצים להוציא);
- בהינתן מספר k מציאת האיבר בעל המפתח הכי קטן כך שקיימים במבנה הנתונים k איברים אדומים עם מפתחות קטנים יותר ממנו בזמן $O(\log n)$. אם לא קיים איבר כזה, כמו למשל במקרה שאין איברים אדומים במבנה הנתונים יוחזר NULL.

יש לתאר (מילולית) כיצד מתבצעת כל אחת מהפעולות הנ"ל. במידה והנך מתבסס/ת על מבנה נתונים שנלמד בכיתה, יש לציין זאת ולהסביר את השינויים שהוכנסו בו.

פתרון

מבנה הנתונים המוצע הוא עץ 2-3 ראשי שמכיל את כל האיברים (גם האדומים וגם הלבנים) בתוספת רשימה מקושרת בין העלים של העץ. בנוסף עץ 2-3 משני שהאיברים שלו (שנמצאים בעלים) מכילים מיצגים של כל האיברים האדומים שנמצאים בעץ הראשי. בנוסף לכל צומת x בעץ המשני ישנו שדה $size(x)$ שמכיל את מספר העלים שנמצאים בתת העץ ששורשו x .

לדוגמה, נניח שבמבנה הנתונים יש 5 איברים אדומים שהמפתחות שלהם הם:

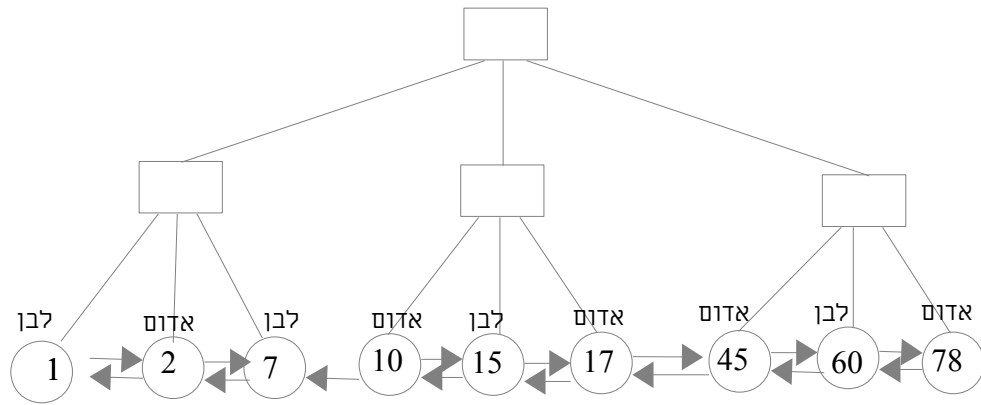
10, 45, 2, 17, 78

ובנוסף יש 4 איברים לבנים שהמפתחות שלהם הם:

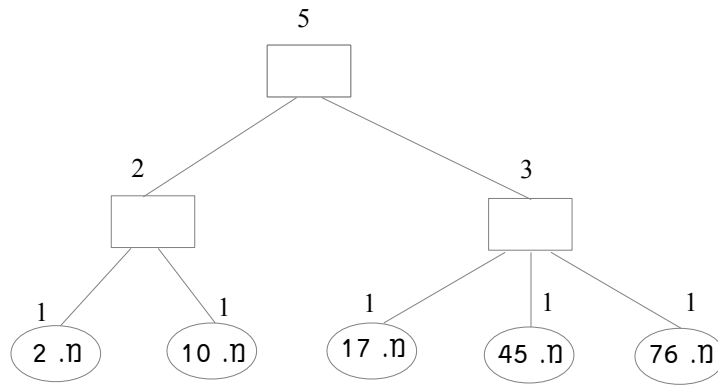
7, 15, 1, 60

אזי מבנה הנתונים המוצע יראה כך:

עץ 2-3 ראשי



עץ 2-3 חשני



הוספת איבר למבנה הנתונים מתבצעת באופן הבא, מוסיפים את האיבר לעץ הראשי ומחברים אותו לרשימה המקושרת בין העלים (זמן הפעולה $O(\log n)$) אם האיבר אדום מוסיפים מיצג שלו לעץ המשני ומעדכנים את שדה ה-size בכל המסלול שמתחיל מהמיצג שלו ועולה למעלה עד לשורש (זמן הפעולה $O(\log n)$).

חיפוש איבר למבנה הנתונים מתבצע באופן הבא: מחפשים את האיבר בעץ הראשי (זמן הפעולה $O(\log n)$).

הוצאת איבר ממבנה הנתונים מתבצעת באופן הבא: מוציאים את האיבר מהעץ הראשי ומעדכנים את הרשימה המקושרת בין העלים (זמן הפעולה $O(\log n)$). אם האיבר אדום מחפשים אותו בעץ המשני ומוציאים אותו ממנו, ומעדכנים את השדה size של כל הצמתים שנמצאים במסלול החל מהאבא שלו ועד לשורש. (זמן הפעולה $O(\log n)$).

בהינתן מספר k מציאת האיבר בעל המפתח הכי קטן כך שקיימים במבנה הנתונים k איברים אדומים עם מפתחות קטנים יותר ממנו, מתבצעת באופן הבא: בעזרת השדה size מוצאים את האיבר ה- k בעץ המשני, אם אין איבר כזה (כי בעץ המשני יש פחות מ- k איברים) מחזירים null. (זמן הפעולה $O(\log n)$). אם יש איבר כזה מחפשים אותו בעץ הראשי (זמן הפעולה $O(\log n)$), ומחזירים את האיבר שמימין לו ברשימה המקושרת בין העלים שבעץ הראשי (זמן הפעולה $O(1)$). סך הכל זמן הפעולה הוא $O(\log n)$.