מסמך תיעוד וניסויים

שם מגיש: מתן סומך | ת"ז: 213120744

שם מגיש: שון סירוטה | ת"ז: 325027589

מסמך תיעוד

תיאור המחלקה המומשה

מימשנו את המחלקה Heap שמתארת אובייקט של ערימה.

הבנאי של הערמה מקבל שני פרמטרים בוליאנים: lazyMelds, lazyDecreaseKeys וממש לפיהם ערימה מהסוגים:

ערימה בינומית, ערימה בינומית עצלה, ערימת פיבונאצ'י ערימה בינומית עם ניתוקים.

בניגוד למימוש הרגיל של ערימה בינומית, בה פעולת Meld ממומשת באמצעות "סכום בינארי" של הערימות, פעולת הMeld מומשה על ידי successive linking.

בנוסף לתיאור ערמה כללית השתמשנו בשדות: מצביע למינימום שמאפשר לנו גישה לערימה, כמות הצמתים בערמה (size), עלות של heapifyCost, כמות צמתים מסומנים וכמות חיתוכים.

על מנת לתאר צומת בערמה השתמשנו בשדות: מצביע לילד, מצביע לאיבר הבא, מצביע לאיבר קודם, מצביע להורה, דרגה של צומת ושדה בוליאני שמעיד האם האיבר מסומן. את המפתח והערך של הצומת שמרנו במחלקה HeapItem כאשר היא מופיע כשדה של צומת ומקשורת לצומת.

ניתוח פעולות מרכזיות

Heap(lazyMelds,DecreaseKeys)

סיבוכיות: $O(1)$

הבנאי של הערמה מקבל שני פרמטרים בוליאנים: lazyMelds, lazyDecreaseKeys וממש לפיהם ערימה מהסוגים:

ערימה בינומית, ערימה בינומית עצלה, ערימת פיבונאצ'י ערימה בינומית עם ניתוקים.

בבנאי אנחנו מאתחילים את השדות:

$$\text{size, numTrees, totalMarkedNodes, totalLinks, totalCuts, totalHeapifyCosts}$$

להיות 0.

Insert(k,info)

סיבוכיות: בהתאם למימוש של הפונקציה meld.

הכנסת איבר לערמה מתבצעת באמצעות הפונקציה meld ולכן הסיבוכיות שלה היא בהתאם לסיבוכיות לסיבוכיות של meld.

- אם lazyMelds = true אזי הסיבוכיות הינה $O(1)$
- אם lazyMelds = false אזי הסיבוכיות הינה $O(\log n)$ מכיוון שאנו מבצעים successive linking.

findMin()

סיבוכיות: $O(1)$

אנחנו שומרים מצביע למינימום.

deleteMin()

סיבוכיות:

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
WC	$O(\log n)$	$O(n)$	$O(n)$	$O(n)$
Amortized	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$

על מנת למחוק את איבר המינימום מהערמה אנחנו מבצעים את השלבים הבאים:

(1) יצירה ערמה של ילדיו של המינימום - יכול לקחת כדרגה של הצומת המינמלית, שזה $O(\log n)$.

(2) מחיקה של המינימום מהערמה - מבוצע על ידי החלפת מצביעים ולכן לוקח $O(1)$.

(3) חיבור של הערמה של הילדים לערמה שנשארה - מבוצע על ידי החלפת צביעים ולכן לוקח $O(1)$.

(4) עושים successive linking - יכול לקחת ככמות הצמתים שיש בערמה. במקרה הגרוע ראינו כי בערמת פיבונאצי וערמה בינומית עצלה פעולה זו יכולה לקחת $O(n)$ אך בזמן לשיעורין עדיין ישאר $O(\log n)$.

cut(x,y)

סיבוכיות:

אם lazymeld=True - $O(1)$, ואם אם lazymeld=False - $O\log n)$.

פעולה זו מקבלת שני צמתים ומנתקת את צומת x מהצומת y ואז מוסיפה את x כשורש לערמה באמצעות meld.

הניתוק של הצומת x מהצומת y נעשה על ידי שינוי מצביעים ולכן העלות של הפעולה היא כעלות של פעולת הmeld. נשים לב שאם lazymeld=True אזי הפעולה נעשה ב$O(1)$ ואם lazymeld=False אזי אנחנו נבצע successive linking ולכן העלות תהיה $O(\log n)$ מכיוון שכאשר lazymeld=False אזי אנחנו שומרים על האינווריאנטה בה כמות השורשים חסומה על ידי $\log n$.

cascadingCut(y)

סיבוכיות:

	סיבוכיות
WC	$O(\log n)$
Amortized	$O(1)$

פעולה זו מבצעת cascadingCut כפי שראינו בכיתה. אנחנו מתחילים מצומת y וכל עוד ההורה של y הוא צומת מסומן אז מנתקים את y מהערמה וממשיכים באופן רקורסבי להורה של y.

במקרה הגרוע הסיבוכיות של הפעולה חסומה על ידי עומק הצומת y, כלומר $O(\log n)$ ובכיתה הראנו כי העלות Amortized הינה $O(1)$.

decreaseKey(x,d)

סיבוכיות:

- lazyDecreaseKey=false - $O(\log n)$
- lazyDecreaseKey=True - העלות במקרה הגרוע הינה $O(\log n)$ וAmortized העלות הינה $O(1)$.

אם lazyDecreaseKey=false אז אנחנו מורידים את ערך המפתח ומבצעים את פעולת הHeapifyUp שהעלות שלה הינה $O(\log n)$.

אם lazyDecreaseKey=True אז אנחנו מורידים את ערך המפתח ואז משווים אותו להורה. אם הוא קטן מההורה אז מבצעים cut וְ cascadingCut ולבסוף מעדכנים את המינימום בהתאם לצורך.

delete(x)

סיבוכיות:

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
WC	$O(\log n)$	$O(n)$	$O(n)$	$O(n)$
Amortized	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$

אנו מבצעים את הפעולה delete באמצעות decreseKey וְ deleteMin ולכן הסיבוכיות בהתאם לסיבוכיות של פעולות אלה.

meld(heap2)

סיבוכיות:

- lazyMelds=false - טבלת סיבוכיות מופיעה תחת הניתוח של consolidate (ראה למטה)
- lazyMelds=True - $O(1)$

הפעולה ממומשת על ידי concat (שרשור הרשימות) ובמידה וְ lazyMelds=false אזי מבצעים גם consolidate.

פעולת הconcat לוקחת זמן קבוע ולכן הסיבוכיות בהתאם לפעולה consolidate.

removeNode()

סיבוכיות: $O(1)$

פעולה זו מוחקת מסירה מהערמה את השורש של המינימום ומחזירה ערמה חדשה רק עם המינימום. פעולה זו היא פעולת עזר לפונקציה consolidate. הסרת המינימום נעשה על ידי שינוי מצביעים ולכן הסיבוכיות של פעולה זו קבועה.

concat(H)

סיבוכיות: $O(1)$

פעולה זו מקבלת ערמה נוספת ומשרשת את שתי הערמות לערמה אחת. השרשור נעשה על ידי שינוי מצביעים ולאחר מכן השוואת המינימום לקביעת המינימום של הערמה המשורשת. שינוי מבצביעים והשוואת מינימוסמים לוקחים זמן קבוע ולכן סיבוכיות הפעולה הינה $O(1)$.

consolidate()

סיבוכיות:

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
WC	$O(\log n)$	$O(n)$	$O(n)$	$O(n)$
Amortized	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$

(1) תחילה אנחנו מאתחילים מערך עזר D. אנו יודעים כי הדרגה המקסימלית של צומת בערימת פיבונאצי עם n צמתים חסומה על ידי $D(n) \leq \lfloor \log_{\Phi} n \rfloor$ ולכן מערך בגודל $\log_{\Phi} n + 1$ מספיק לנו להבטיח שלא נחרוג מגבולות המערך. נשים לב שחסם זה נכון גם עבור הערמות.

(2) אנחנו עוברים על שורשי הערמה ובכל שלב עוברים על השורש x, מסירים אותו מהערמה בעזרת המתודה removeNode (שלוקחת זמן קבוע) ושמים אותו במערך D לפי גודלו דרגתו. כאשר x הוא עם דרגה d כך שכבר יש צומת בתא זה במערך אז אנחנו מבצעים linking וממשיכים לבדוק במיקום החדש במערך עד שמוצאים תא ריק.

עלות שלב זה הינה $O(T_0 + D(n))$ כאשר T_0 הינו מספר השורשים לפני האיחוד.

(3) בשלב זה אנחנו מקבלים שהערמה שהתחלנו איתה ריקה ובידנו מערך D כך שבכל תא לא ריק במערך יש שורש מדרגה ייחודית. אזי אנחנו עוברים על המערך ומאחדים את כלל השורשים לערמה אחת. במהלך המעבר, מעדכנים את מצביע ה-min של הערמה.

עלות שלב זה הינה כגודל המערך כלומר $O(D(n))$.

נשים לב שבכל אחת מהערמות מלב הערימה הבינומית יכולים להיות לנו $O(n)$ שורשים בשלב 2 שנצטרך לעבור עליהם ולכן הWC של כלל הערמות מלב הערימה הבינומית הינו $O(n)$. ערימה בינומית אוכפת סדר בכל רגע ולכן בכל רגע אין לנו יותר מ$O(\log n)$ שורשים בערמה.

heapifyUp(x)

סיבוכיות: $O(\log n)$

המתודה מקבלת מצביע לצומת x ומבצעת עליו heapifyUp, כלומר היא בודקת את ערך המפתח שלו ומפעפת אותו למעלה עד למקום המתאים לו. הפעפוע למעלה נעשה באמצעות החלפת מציעים של הHeapItem ולכן זו עלותזו קבועה אך כמות הפעפוע חסומה על ידי עומק של הצומת x ולכן העלות הינה $O(\log n)$.

פעולות "שמירת נתונים"

את המתודות:

size(), numTrees(), totalMarkedNodes(), totalLinks(), totalCuts(), totalHeapifyCosts()

ערכים אלה מופיעים כשדות של הערמה. תחילה אנחנו מעדכנים את כל הערכים להיות 0 ובמהלך הפעולות השונות מעדכנים ערכים אלה בהתאם. כך שאנחנו מחזירים את כל הפעולות האלה ב$O(1)$.

הערה. הוספנו נספח שמאתר כיצד אנחנו שומרים על שדות אלה במהלך ביצוע הפעולות השונות.

מסמך ניסוי

סעיף 1

נסמן בT_0 את כמות השורשים בערמה לפני ביצוע פעולה.
נסמן בT_1 את כמות השורשים בערמה לאחר ביצוע פעולה.
נסמן בL את כמות הLinkים
נסמן בm את כמות הצמתים המסומנים.
נסמן בm' את כמות הצמתים המסומנים לאחר הפעולה.

נגדיר פונקציית פוטנציאל:

$$\Phi = 2\cdot\#(\text{number of Trees})+3\cdot\#(\text{number of mark nodes})$$

ניתוח Insert:
בפעולת Insert אנו מבצעים meld ולכן מתבצע consolidate.
אזי מתקיים כי $T_1 = T_0 + 1 - L$ ולכן:

$$\widehat{c} = c + \Phi_{After} - \Phi_{before} = (L+1) + (2T_1 + 3m) - (2T_0 + 3m) =$$

$$L + 1 + 2T_0 + 2 - 2L + 3m - 2T_0 - 3m \leq 3 - L \leq 3$$

ולכן פעולת Insert לוקחת זמן קבוע לשיעורין.

ניתוח decreaseKey:
נסמן בc את כמות החיתוכים שבוצעו. אזי:
נקבל שלאחר הפעולה נוספו c עצים חדשים לערמה, לכן $T_1 = T_0 + c$.
נסמן בm' את כמות הצמתים המסומנים לאחר הפעולה. אזי $m' = m - c$ או $m' = m - c + 1$ חיתוכים ובכל מקרה $m' \leq m - c + 1$. אזי:

$$\widehat{c} = c + \Phi_{After} - \Phi_{before} = (c) + (2T_1 + 3m') - (2T_0 + 3m) \leq$$

$$(c) + (2T_0 + 2c + 3(m - c + 1)) - (2T_0 + 3m) = 3$$

ולכן פעולת decreaseKey לוקחת זמן קבוע לשיעורין.

ניתוח deleteMin:
בפעולה זו אנחנו מוחקים את המינימום, מוסיפים את ילדיו כשורשים לעץ ואז עושים פעולת consolidate.
נשים לב שלאחר הפעולה מתקיים $T_1 \leq \log n$. אזי:

$$\widehat{c} = c + \Phi_{After} - \Phi_{before} = (T_0 + \log n - 1 + L) + (2T_1 + 3m) - (2T_0 + 3m)$$

נשים לב כי $L \leq T_0 + \log n - 1$ שהרי כמות הלינקים יכולה להיות לכל היותר ככמות השורשים. אזי:

$$\leq \log n - 1 + \log n - 1 + 2T_1 =$$

$$2\log n + 2T_1 - 2 \leq 5\log n$$

ולכן קיבלנו כי העלות לשיעורין של פעולה זו הינה $O(\log n)$.

ניתוח findMin:
אנחנו שומרים מצביע למינימום ולכן העלות היא $O(1)$.

ניתוח delete:
אנו מבצעים את הפעולה delete על ידי decreaseKey ו deleteMin ולכן העלות לשיעורין שלו מקסימלית מבינהם, כלומר $O(\log n)$.

לסיכום:

פעולה	עלות לשיעורין
Insert	$O(1)$
decreaseKey	$O(1)$
deleteMin	$O(\log n)$
findMin	$O(1)$
delete	$O(\log n)$

סעיף 2

פעולה	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
Insert	$O(\log n)$	$O(1)$	$O(1)$	$O(1)$
decreaseKey	$O(\log n)$	$O(\log n)$	$O(1)$	$O(1)$
deleteMin	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$
findMin	$O(1)$	$O(1)$	$O(1)$	$O(1)$
delete	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$

סעיף 3

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
ניסוי 1	$O(n \log n)$	$O(n)$	$O(n)$	$O(n)$
ניסוי 2	$O(n \log n)$	$O(n)$	$O(n)$	$O(n)$
ניסוי 3	$O(n \log n)$	$O(n)$	$O(n)$	$O(n)$

סעיף 4

בחרנו להשאיר את ערכים אינם מעוגלים מכיוון שערכים אשינם מעוגלים יכולים גם להראות לנו מתי ערך מסויים נשאר קבוע בין הריצות השונות (לדוגמה גודל הערמה בסוף) למתי ערך מסויים תלוי באופן הכנסת האיברים לערמה ובהוצאתם מהערמה. לדוגמה בניסוי השני בערימת פיבואנצי קיבלנו שכמות העצים הממוצעת הייתה 2.95, דבר שמעיד שהייתה קריאה בה מספר העצים בסיום היה 2 ואינו קבוע 3.

ניסוי 1

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
זמן ריצה (מילישניות)	87.95	12.4	11.75	86.6
גודל הערמה בסיום	464645	464645	464645	464645
מספר העצים בסיום	9	9	9	9
מספר חיבורים	464653.4	464636	464636	464653.3
מספר חיתוכים	0	0	0	0
סך עלויות heapify up	0	0	0	0
עלות מקסימלית לפעולה	18	464636	464636	18

ניסוי 2

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
זמן ריצה (מילישניות)	737.95	618.45	151.8	354.35
גודל הערמה בסיום	46	46	46	46
מספר העצים בסיום	4	4	2.95	2.9
מספר חיבורים	7552642.65	7553000	748852.55	1406696.3
מספר חיתוכים	0	0	748809.5	748749.8
סך עלויות heapify up	4149614.25	4149773.3	0	0
עלות מקסימלית לפעולה	35.7	464636	464636	39.25

ניסוי 3

	ערימה בינומית	ערימה בינומית עצלה	ערימת פיבונאצ'י	ערימה בינומית עם ניתוקים
זמן ריצה (מילישניות)	113.25	36.25	78.9	249.4
גודל הערמה בסיום	464644	464644	464644	464644
מספר העצים בסיום	8	8	8	8.95
מספר חיבורים	464671.15	464653.35	929272	993349.7
מספר חיתוכים	0	0	464636	528696.15
סך עלויות heapify up	373198.05	373306.25	0	0
עלות מקסימלית לפעולה	18	464636	464636	29.9

סעיף 5

סעיף א

לפי התוצאות ניתן לראות הבדלים מהותיים בין זמני הריצה, מספרי חיבורים ועלות מקסימלית לפעולה.

סעיף ב

סעיף ג

ניתן לראות שהעלות המקסימלית לפעולה קטנה בערימה בינומית לעומת ערימה בינומית עצלה וערימת פיבונאצי, דבר שמאשש את הניתוח לשיעורין בו יש לנו פעולה יקרה ש"מכסה" על הפעולות הזולות.

נספח

שמירה על ערכים במבנה נתונים

להלן פירוט של אופן שמירה של הערכים הבאים:

size, numTrees, totalMarkedNodes, totalLinks, totalCuts, totalHeapifyCosts

אנחנו שומרים ערכים אלה מבלי לפגוע בסיבוכיות של שאר הפעולות.

שמירת הערךsize

- כאשר מבצעים insert בעת יצירת ערמת הסינגלטון שלנו נעדכן את הsize שלו להיות 1 על מנת שבמהלך פעולת הconcat נעלה את כמות הצמתים ב1.
- בסוף פעולת הdeleteMin נקטין את כמות הצמתים ב1.
- נשים לב שפעולת הdelete קוראת למתודה deleteMin ולכן אין צורך להוריד גם שם את גודל המערך.
- במהלך ביצוע concat נאחד את כמות הצמתים של שתי הערמות.

שמירת הערךnumTrees

- בפעולת concat מאחדים את כמות השורשים של שני העצים
- בפעולת insert כאשר יוצרים את הערמה עם הצומת הבודדת מעדכנים את כמות השורשים של הערמה זו להיות 1 - כך מובטח לנו שבמהלך הmeld כאשר מבצעים concat אכן כמות השורשים תגדל ב1.
- במהלך פעולת deleteMin אנחנו מוחקים את המינימום ומוסיפים לעץ את כל הילדים של המינימום. אזי לפני קריאה לפעולת consolidate נעדכן את כמות השורשים להיות:

- אם יש רק שורש אחד אז נעדכן את numOfTrees להיות MinchildNumOfTrees שזה פשוט ה$rank$ שלה מינימום.

- אם יש יותר משורש אחד אז נעדכן את numOfTrees להיות numOfTrees+MinchildNumOfTrees-1.

- כאשר מבצעים linking במהלך הconsolidate אז נוריד את כמות השורשים ב1. בכל מקרה כאשר אנחנו בונים בחזרה את הערמה נספור את כמות השורשים (שזה כמות התאים שאינם ריקים במערך).
- כאשר אנחנו מבצעים cut אז אנחנו יוצרים ערמה מהצומת שאותו מנתקים ואז עושים meld, נעדכן בערמה זו את כמות השורשים להיות 1 כדי שבעת ביצוע concat נוסיף 1 לכמות השורישם במערך.

שמירת הערךtotalMarkdNodes

- כאשר אנחנו מבצעים cut נוריד את כמות הצמתים המסומנים ב1.
- כאשר אנחנו מבצעים cascadingCut ומגיעים לצומת שאינו מסומן ולפני השורש נסמן אותו ונעלה את הכמות הצמתים המסומנים ב1.

שמירת הערךtotalLinks

- אנו מבצעים linking רק במהלך המתודה cosolidate ולכן לאחר ביצוע פעולת link נעלה את כמות הlinkים ב1.
- במהלך הפעולה concat נחבר את כמות הlinkים של שתי הערמות על מנת לשמור את ההיסטוריה של כמות הlinkים.

הערה. מבחינת design ייתכן שהיה יותר נכון לעדכן את כמות הlinkים בתחילת המתודה link אך מתודה זו נמצאת תחת המחלקה HeapNode ולכן בחרנו לעקוב אחר כמות הlinkים בדרך עקיפה.

שמירת הערךtotalCuts

- במתודה cut נעלה את כמות החיתוכים ב1.
- במהלך הפעולה concat נחבר את כמות החיתוכים של שתי הערמות על מנת לשמור את ההיסטוריה של כמות החיתוכים.

שמירת הערךtotalHeapifyCosts

- במתודה heapify, הגדרנו משתנה count שסופר את כמות הheapifyUp שביצענו והמתודה heapify מחזירה את המונה הזה. אזי, בעת קריאה למתודה heapify (שקוראת בdecreaseKey כדי לבצע heapifyUp נעשה:

```
this.heapifyCosts += heapifyUp(x.node);
```

וכך נשמור בכל קריאה לheapifyUp אנחנו מעדכנים את המונה הכללי.