**תיעוד פרוייקט מעשי 2 בקורס מבני נתונים – מימוש ערימת פיבונאצ'י**

לירון כהן, lironcohen3, 207481268  
יובל מור, yuvalmor, 209011543

**המחלקה AVLTree**

המחלקה מממשת עץ AVL, ובה שדה יחיד מסוג IAVLNode המצביע על שורש העץ.

כל אחד מצמתי העץ מממש מופע ממחלקת AVLNode, אודותיה נפרט בהמשך.

**מתודות AVLTree**

**AVLTree()**

המתודה בונה עץ ריק, כלומר מאתחלת את שורש העץ להיות null.

סיבוכיות המתודה היא כיוון שהיא קובעת ערכים של שדה.

**המחלקה AVLNode**

המחלקה מממשת צומת בעץ AVL לפי המנשק IAVLNode.

**שדות AVLNode**

* Int Key – המפתח של הצומת
* String info – הערך של הצומת
* IAVLNode parent – מצביע להורה של הצומת
* IAVLNode left – מצביע לבן השמאלי של הצומת
* IAVLNode right – מצביע לבן הימני של הצומת
* Boolean isReal – האם הצומת אמיתי או וירטואלי
* Int height – גובה / דרגת הצומת
* Int size – גודל תת העץ שתלוי מהצומת

**מתודות AVLNode**

**AVLNode(int key, String info)**

המתודה בונה עצם מסוג AVLNode, מאתחלת את המפתח והערך שלו למפתח והערך שהתקבלו כארגומנטים.

בנוסף, אם המפתח שהתקבל אינו -1 (כלומר צומת אמיתי), משנה את ערך ה-isReal לאמת (כיוון שבאופן דיפולטי משתנים בוליאניים מאותחלים לשקר), מגדירה את שני הבנים של הצומת לצמתים וירטואליים חדשים ואת ערך שדה ה-size שלו ל-1.

אם המפתח שהתקבל הוא -1, מגדירה את דרגתו כ-1 ואת שדה ה-size שלו ל-0.

סיבוכיות המתודה היא כיוון שהיא מבצעת השמות ובדיקות שנעשות בזמן קבוע.

**חלק המדידות**

**שאלה 1 סעיף א'**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| מספר סידורי | גודל המערך | כמות החילופים במיון רגיל עבור מערך אקראי | כמות החילופים במיון רגיל עבור מערך ממוין הפוך | עלות החיפושים במיון AVL עבור מערך אקראי | עלות החיפושים במיון AVL עבור מערך ממוין הפוך |
| 1 | 10,000 | 25,063,646 | 49,995,000 | 221,782 | 231,358 |
| 2 | 20,000 | 100,023,653 | 199,990,000 | 480,520 | 502,688 |
| 3 | 30,000 | 224,839,071 | 449,985,000 | 777,310 | 788,113 |
| 4 | 40,000 | 401,139,380 | 799,980,000 | 1,056,505 | 1,085,346 |
| 5 | 50,000 | 625,021,906 | 1,249,975,000 | 1,334,497 | 1,386,195 |
| 6 | 60,000 | 900,795,981 | 1,799,970,000 | 1,614,641 | 1,696,195 |
| 7 | 70,000 | 1,225,062,394 | 2,449,965,000 | 1,986,235 | 2,010,660 |
| 8 | 80,000 | 1,599,255,957 | 3,199,960,000 | 2,233,915 | 2,330,660 |
| 9 | 90,000 | 2,032,584,245 | 4,049,955,000 | 2,522,302 | 2,650,660 |
| 10 | 100,000 | 2,504,360,436 | 4,999,950,000 | 2,885,898 | 2,972,357 |

ציפיות ומסקנות מהטבלה

מיון רגיל - השוואה של מערך ממוין הפוך ומערך אקראי

מבחינת קלטי המערכים האפשריים לאלגוריתם המיון, מערך ממוין בסדר הפוך הוא הגרוע יותר מבחינת כמות הפעולות. מערך זה הכי רחוק ממערך ממוין בסדר עולה ולכן נצפה לכמות חילופים גדולה מאוד ביחס למערך אקראי, אשר יכול להגריל סדר ערכים (באמצעות פעולת shuffle) היוצר מערך הקרוב יותר למערך הממוין הדרוש ולכן ידרוש מספר נמוך יותר של חילופים.

מיון AVL – השוואה של מערך ממוין הפוך ומערך אקראי

ראשית נדגיש שעקב אופן בניית המערכים, האיברים בעץ ה-AVL לאחר כל סבב הכנסות יהיו זהים (כלומר איברים מ-0 ועד גודל המערך פחות 1, אשר הוכנסו לעץ בסדר שונה). עקב העובדה שעץ AVL הוא עץ חיפוש, אלגוריתם ההכנסה שלו מסדר את העץ כך שישמור על כללי עץ חיפוש וכן מאזן את העץ. לכן, נצפה שלא יהיה הבדל משמעותי בין מערכים שונים שהוכנסו למיון העץ. נציין שעקב סדר ההכנסה במערך ממוין הפוך, ועקב העובדה שככל שמתקדמים במערך האיברים קטנים יותר וגם העץ גדל, נוכל לשער שעלות החיפושים תגדל מעט יותר בסדר הכנסה זה.

השוואה של כמות החילופים במיון רגיל אל מול עלות החיפושים במיון AVL

מכיוון שעץ AVL הוא עץ מאוזן, עלות חיפוש תהיה ב- (סדר גודל של גובה העץ), לעומת מיון רגיל שבכל איטרציה עובר על סדר גודל של כל המערך כדי לבצע את ההשוואות, כלומר . לכן, נצפה שעלות החיפושים במיון AVL תהיה קטנה משמעותית מכמות החילופים במיון רגיל.

כמות החילופים כתלות בגודל המערך

מכיוון שכמות החילופים במיון רגיל תלויה בגודל המערך וכן גם עלות החיפושים במיון AVL, נצפה שככל שנגדיל את גודל המערך כך יתבצעו יותר חילופים ועלות החיפושים תעלה.

ניתן לראות שהתוצאות שהתקבלו תואמות את הציפיות.

ניתוח כמות החילופים ב-Insertion Sort עבור מערך אקראי

מכיוון שמדובר במערך אקראי, נחשב את תוחלת כמות החילופים הממוצעת, על פני כל המערכים האפשריים. נסמן את המערך ב-.

עבור כל איטרציה של הלולאה באלגוריתם המיון, נגדיר עבור   
אינדיקטור .

נשים לב שעבור מתקיים:

נקבל:

*מסכום סדרה חשבונית נקבל:*

*ובסך הכל נקבל:*

*מהגדרת אלגוריתם המיון, מכיוון שמבצעים איטרציה על כל איברי המערך ומבצעים חילופים סך הכל, סיבוכיות האלגוריתם תהיה:*

ניתוח כמות החילופים ב-Insertion Sort עבור מערך ממוין הפוך

נשים לב שמדובר בקלט ספציפי ולכן נחשב באופן ישיר כמה חילופים מתבצעים עבורו.

נניח שהמערך בגודל .

* *עבור האיבר הראשון במערך, נבצע 0 חילופים.*
* *עבור האיבר השני במערך נשווה אותו לראשון, הם יהיו בסדר הפוך ולכן נבצע חילוף אחד.*
* *עבור האיבר השלישי הקטן מהשניים שלפניו, נבצע שני חילופים.*
* *באופן כללי, עבור האיבר ה- נבצע חילופים.*

*נסכום את מספר החילופים ונקבל סדרה חשבונית:*

*ולכן מספר החילופים עבור מערך זה הינו חילופים, כלומר חילופים.*

ניתוח עלות החיפושים ב-AVL insertion sort עבור מערך אקראי

ראשית ננתח את עלות החיפושים כתלות ב- וב- (כמות החילופים שהייתה מתבצעת אם המיון היה רגיל).

נסמן בתור את כמות החילופים שנעשו לאיבר במקום ה- במערך, כך שמתקיים:

*עבור חיפוש של איבר במערך נעבור על של מספר האיברים הנמצאים לפני האיבר במערך וגדולים ממנו (כלומר הוכנסו לפניו לעץ ונעבור דרכם בחיפוש), כלומר קשתות. נסמן את כמות הקשתות הכוללת בתור , ונקבל:*

*כאשר אי השיוויון האחרון נובע ממניפולציה אלגברית על אי שיוויון הממוצעים והשיוויון האחרון נובע מחוקי לוגריתמים.*

*לבסוף נזכור שיש לבצע סריקת של העץ ולכן נקבל .*

*קיבלנו חסם כללי על עלות החיפושים הכוללת ולכן זהו גם חסם על התוחלת המבוקשת.*

*בבירור (מכיוון שזוהי הכמות המקסימלית של ההחלפות עבור איבר במערך), ולכן:*

*ומכאן נקבל .*

ניתוח עלות החיפושים ב-AVL insertion sort עבור מערך ממוין הפוך

נשים לב שמדובר בקלט ספציפי ולכן נחשב באופן ישיר כמה חילופים מתבצעים עבורו.

ראינו בהרצאה שחיפוש איבר עם מפתח בעץ נעשה בסיבוכיות של , ובמקרה הגרוע נקבל קשתות לכל חיפוש.

מחוקי לוגריתמים נקבל:

**שאלה 1 סעיף ב'**

כפי שהראינו בסעיף הקודם, החסם העליון על כמות החיפושים הוא . מלבד עלות החיפושים, עבור חסם על הסיבוכיות נתחשב גם בפעולות איזון שמתבצעות על עץ ה-.

במקרה הגרוע (עבור חסם עליון), נקבל שכל פעולת איזון חסומה ע"י (שמתקבלת עבור שרשרת ובמימוש שלנו מתקבלת תמיד עקב פעולת ). מתבצעת פעולת איזון עבור כל הכנסה, כלומר פעולות איזון ולכן נקבל:

**שאלה 2 סעיף א'**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי | עלות join ממוצע עבור split אקראי | עלות join מקסימלי עבור split אקראי | עלות join ממוצע עבור split של איבר מקסימלי בתת העץ השמאלי | עלות join מקסימלי עבור split של איבר מקסימלי בתת העץ השמאלי |
| 1 | 2.833 | 6 | 2.769 | 16 |
| 2 | 2.769 | 7 | 3.083 | 17 |
| 3 | 2.750 | 10 | 2.929 | 18 |
| 4 | 2.286 | 6 | 2.923 | 18 |
| 5 | 2.700 | 6 | 2.846 | 18 |
| 6 | 2.800 | 5 | 2.667 | 19 |
| 7 | 2.714 | 5 | 2.688 | 19 |
| 8 | 2.733 | 7 | 2.600 | 19 |
| 9 | 3.067 | 7 | 2.867 | 20 |
| 10 | 3.200 | 6 | 3.000 | 20 |

ניתוח עלות ממוצעת ל-join עבור split של האיבר המקסימלי בתת העץ השמאלי

נשים לב שאם ישנם צמתים בעץ, גובה העץ יהיה כ-. עלות ממוצע תהיה:

*כאשר מבצעים על האיבר המקסימלי בתת העץ השמאלי של העץ, כמות פעולות ה- שנבצע תהיה כגובה העץ כולו, שהוא . בנוסף, ראינו בהרצאה שמחיר פעולת היא כסכום מחירי פעולות ה- המתבצעות בתוך פעולת ה- ומכאן נקבל את העלות המבוקשת.*

ניתוח עלות ממוצעת ל-join עבור split אקראי

כפי שראינו, במקרה הגרוע (האיבר המקסימלי בתת העץ השמאלי) העלות הממוצעת תהיה ובפרט חסם קבוע זה הוא גם חסם גם על התוחלת עבור מקרה אקראי.

ניתוח עלות מקסימלית ל-join עבור split אקראי

בהנחה שהעץ מאוזן, ראינו בתרגול שכמות הצמתים ברמה שבגובה היא (כלומר ברמת העלים יש מהצמתים, ברמה שמעליה יש מהצמתים וכו').

עבור לפי צומת ברמה ה-, ה- המקסימלי יעלה כמספר המקסימלי של הקשתות שעולות "באותו הכיוון" (כלומר שרשרת של בנים ימניים או שרשרת של בנים שמאליים).   
עלות פעולת ה- המקסימלי תהיה כגובה העץ פחות גובה הרמה, כלומר .

נחשב לפי משפט התוחלת השלמה את תוחלת כמות פעולות ה- בחלוקה לפי רמות בהן ה- מבוצע:

ניתוח עלות מקסימלית ל-join עבור split של האיבר המקסימלי בתת העץ השמאלי

עקב המיקום הייחודי של האיבר המקסימלי בתת העץ השמאלי, נשים לב שה- המקסימלי יתקבל   
ב- האחרון.

בהתחלה יבוצעו (גובה העץ פחות 1) פעולות המחברות "מכיוון עלייה שמאלה" (בנים ימניים) ולכן הפרשי גבהי העצים בפעולות אלו יהיו נמוכים.

לעומת זאת, פעולת ה- האחרונה תבצע בין עץ ריק (תת העץ הימני של האיבר המקסימלי בתת העץ השמאלי, שהוא עלה ולכן תת העץ ריק) לבין כל תת העץ הימני של העץ כולו.

הפרש הגבהים במקרה זה הוא וזו תהיה העלות המקסימלית ל- כחלק מ- זה. נקבל:

**שאלת בונוס**

נשים לב שעבור של איבר אקראי, עלות ה- המקסימלי היא אורך הרצף המקסימלי של עליות מהאיבר לשורש "מאותו הכיוון" (כלומר שרשרת של בנים שמאליים או שרשרת של בנים ימניים).

*נשים לב, שאיבר הוא בן שמאלי בהסתברות ובן ימני בהסתברות .*

*לכן, אם נתבונן באיבר האקראי הנבחר וברצף האבות ממנו עד השורש, נרצה לספור את כמות הבנים הרצופים מאותו הכיוון, מה שתואם לצורת התפלגות גיאומטרית אותה ראינו בקורס מבוא להסתברות.*

*ובפרט:*

* *רצף של שני איברים מאותו הכיוון יתקבל בהסתברות ועלות ה- תהיה .*
* *רצף של שלושה איברים מאותו הכיוון יתקבל בהסתברות ועלות ה- תהיה 3.*
* *רצף של כל האיברים מאותו הכיוון יתקבל בהסתברות ועלות ה- תהיה .*

*ממשפט התוחלת השלמה נקבל את הסכום:*

*מחיפוש ברשת האינטרנט נמצא כי אורך הרצף המקסימלי (בדומה לבעיית רצפים בהטלת מטבע) הוא . אם אורך הרצף המקסימלי הוא , עלות ה- תהיה ומכאן נקבל את התוצאה .*