מטלה מעשית 1 מבני נתונים

מתן טלוי 318903028

ירון ראובן איטח 208935312

מחלקת :

מחלקה זו מכילה את השדות הבאים:

* - ערך הצומת.
* - מצביע לבן השמאלי, מאותחל בתור צומת וירטואלי, (מפורט בהמשך).
* - מצביע לבן הימני, מאותחל בתור צומת וירטואלי, .
* - מצביע לאב הצומת, מאותחל בתור .
* - גובה הצומת (גובה של עלה מוגדר להיות , של צומת וירטואלי ).
* - מספר הצמתים בתת העץ כולל הצומת הנוכחי.

מחלקה זו מכילה את הפעולות הבאות אשר **כולן בסיבוכיות**  ומשתמשות בפונקציית לבדיקה אם הצומת היא וירטואלית או אמיתית:

* – מגדירה את שדה להיות ה של הקלט.
* - מחזירה את שדה של הצומת או אם הצומת הוא וירטואלי.
* - מגדירה את השדה להיות ה שהוכנס.
* - מחזירה את שדה של הצומת, או אם אין כזה.
* - מגדירה את השדה להיות ה שהוכנס.
* - מחזירה את שדה של הצומת, או אם אין כזה.
* - מגדירה את השדה להיות ה שהוכנס.
* - מחזירה את השדה של הצומת, או אם אין כזה.
* - מחזירה אם הצומת מייצג צומת אמיתי בעץ. כלומר צומת שאינו וירטואלי וזאת כאשר .
* - מגדירה את שדה של הצומת לפי הערך שהוכנס.
* - מחזירה את שדה של הצומת. עבור צומת וירטואלי הגובה .
* - מגדירה את שדה של הצומת. מאותחלת עם הערך 0.
* - מחזירה את שדה של הצומת. עבור צומת וירטואלי .
* – מחזירה את הbalance factor של הצומת. חישוב: גובה הבן השמאלי פחות גובה הבן הימני.
* – מעדכן את שדה של ה בחישוב ה של הבנים שלו כפי שנלמד בכיתה.
* – מעדכן את שדה של ה בחישוב ה של הבנים שלו כפי שנלמד בכיתה.

מחלקת :

מחלקה זו מכילה את השדות הבאים:

* - מצביע על שורש העץ, על איבר מסוג .
* - מחזיק את כמות הצמתים בעץ.
* – מצביע לאיבר הראשון ברשימת האיברים הממוינת בסדר עולה.
* – מצביע לאיבר האחרון ברשימת האיברים הממוינת בסדר עולה.

מחלקה זו מכילה את הפעולות הבאות:

* - פעולה אשר מחזירה אם שורש העץ הינו , כלומר העץ ריק, ו- אחרת. פעולה זו בסיבוכיות שכן היא פונה לשדה של העץ.
* - מחזירה את השדה של העץ , אלא אם מדובר בצומת וירטואלי עבורו היא מחזירה את הערך . גישה לשדה בסיבוכיות .
* - פעולה המחזירה את ערך של הצומת בעלת המפתח , אם קיימת. פעולה זו משתמשת בפונקציה בסיבוכיות ולכן זוהי גם סיבוכיותה של פונקציית .
* - פונקציה אשר מקבלת צומת וערך מפתח ובודקת על-ידי "טיול" על גבי העץ מהשורש ועד לעלה העץ אם הגענו בהתאם לתכונות של עץ חיפוש בינארי לערך המפתח המבוקש. כלומר התקדמות לעומק העץ מתבצעת לפי הלוגיקה של אם ערך המפתח גדול מערך הצומת אותה אנו מבקרים, נמשיך לטייל ימינה לערכים גדולים יותר ושמאלה אחרת. אם הגענו לעלה וירטואלי סימן שלא מצאנו את הצומת המבוקשת כי היא אינה קיימת ונחזיר . מכיוון שעומק עץ הינו וכל הפעולות שאנו מבצעים הינן פעולות בזמן קבוע הרי שמדובר בסיבוכיות לרקורסיה כולה.
* - תחילה נחפש את הצומת באמצעות פעולת ואם היא כן נמצאת נחזיר שכן אסור ב להחזיק שני צמתים עם אותו . אחרת, נאתחל משתנה הבנוי מ שהגיעו כקלט. נבדוק אם הצומת החדש הוא המקסימום או המינימום החדש בעץ ואם כן נעדכן את השדות המתאימים. אם העץ ריק, נבדוק באמצעות שימוש בפונקציית , נעדכן את השדות המתאימים ונחזיר 0 (כמספר פעולות קידום/סיבוב). נגדיר משתנה שמתעד את מספר פעולות הקידום והסיבוב. באמצעות פונקציית , בעלת סיבוכיות נמצא את אביו של הצומת המיועד להכנסה ונכניסו לצד המתאים של אביו באמצעות ערכו ונעדכן את השדות המתאימים. כעת נבדוק אם הכנסנו את הצומת כבן שמאלי/ימני של צומת אונארי או של עלה. בהתאם, נעדכן את השדות המתאימים עד השורש. נבחין כי אם הכנסנו את הצומת לעלה, אז עלינו לבצע את פעולת בסיבוכיות של כדי לאזן את העץ מחדש. תוך כדי הפעולה נשמור את מספר הקידומים/סיבובים שבוצעו ונחזיר ערך זה. נבחין כי כל הפעולות הינן בלתי תלויות ולא מקוננות ולכן בסך הכל מדובר בסיבוכיות של .
* - פונקציה איטרטיבית אשר תכליתה למצוא משורש העץ על-פי כללי החיפוש את הצומת עבורה . אם זה לא קיים בעץ, אזי נחזיר את הצומת שהוא הכי קרוב מבחינת תכונות החיפוש ל כפי שנלמד בהרצאה. נבחין כי במקרה הגרוע אנו נבצע ירידה אחת עד לעלה וירטואלי כלשהו ולכן סיבוכיות פעולה זו הינה .
* - פונקציה אשר מקבלת שני צמתים, אב ובן. מבצעת החל מהבן ועד לשורש העץ פעולות איזון בהתאם לנלמד בהרצאה, וסופרת באמצעות את מספר פעולות הקידום/סיבוב שבוצעו במהלך הטיול. במידה ויש לבצע סיבוב או סיבוב כפול, נעזר בפעולות למיניהן. לאחר פעולות האיזון נעדכן את שורש העץ במידת הצורך, ונחזיר את מספר פעולות הקידום/סיבוב אשר התבצעו במשך פונקציה זו. במידה ולא הגענו לשורש העץ מפעולות איזון (כלומר העץ התאזן לפני כן), נמשיך לטייל עד שורש העץ כדי לעדכן את השדות המתאימים. מכיוון שאנו מטיילים רק במעלה העץ, במקרה הגרוע נטייל מעלה ועד לשורש וכן נבחין כי הסיבוכיות של פונקציות העזר הינן , הסיבוכיות כולה הינה .
* - פונקציה אשר מבצעת סיבוב שמאלה כפי שנלמד בהרצאה בהינתן שני צמתים, אב ובן. סיבוכיות .
* - פונקציה אשר מבצעת סיבוב ימינה כפי שנלמד בהרצאה בהינתן שני צמתים, אב ובן. סיבוכיות .
* - פונקציה אשר מבצעת סיבוב כפול, ימין ואז שמאל, כפי שנלמד בהרצאה בהינתן שלושה צמתים, אב, בן ונכד. סיבוכיות .
* - פונקציה אשר מבצעת סיבוב כפול, שמאל ואז ימין, כפי שנלמד בהרצאה בהינתן שלושה צמתים, אב, בן ונכד. סיבוכיות .
* - תחילה נבדוק שהאיבר נמצא בעץ. נמצא את האיבר שצריך למחוק באמצעות קריאה ל- ונכנו . אם גודל העץ הוא אזי אנו מוחקים את השורש ולכן נאפס את העץ ונסיים. אם לצומת שאנחנו רוצים למחוק יש שני ילדים, נמצא את ה שלה, נמחק את הצומת ונשים במקומו את ה שלו. במקרה זה נקרא לפונקציית ה- ונאזן את מעלה העץ החל ממיקום איבר ה- אחרי המחיקה. במקרה קצה שבו האיבר שנמחק הינו האבא של ה-, נבודד נקודתית את . אם הינו עלה, פשוט נמחק את העלה. במקרה האחרון, הינו צומת אונארי ולכן נעקוף את . נפחית את גודל של העץ באחד, ותוך כדי נעדכן את השדות המתאימים (גבהים וגדלים) של האיברים המתאימים. נאתחל משתנה אשר מונה את מספר הקידומים והסיבובים שנעשים תוך כדי האיזון. נטייל החל מאביו של המיקום של הצומת שמחקנו ועד לשורש ונבצע פעולות איזון בהתאם לנלמד בכיתה. לבסוף, נבדוק אם מחקנו את האיבר המינימלי או המקסימלי של העץ ואם כן נעדכן את השדות בהתאמה על-ידי שימוש ב של המינימלי או למקסימלי. ונחזיר את משתנה . כפי שהוכח בכיתה סיבוכיות פעולה זו הינה ב. נבחין כי השימוש בפונקציות חיצוניות הן בלתי תלויות ואינן גוברות על סיבוכיות זו, כלומר כולן לכל היותר בסיבוכיות .
* - פונקציה איטרטיבית אשר מטיילת משורש העץ ובכל פעם מתקדמת ימינה עד שהיא מגיעה לעלה שהוא הצומת בעל המפתח המקסימלי. סיבוכיותה כעומק העץ .
* - פונקציה איטרטיבית אשר מטיילת משורש העץ ובכל פעם מתקדמת שמאלה עד שהיא מגיעה לעלה שהוא הצומת בעל המפתח המינימלי. סיבוכיותה כעומק העץ .
* - פונקציה אשר מקבלת צומת. אם יש לו בן ימני היא תחזיר את הצומת הקטן ביותר של תת עץ ימני זה. אחרת, היא תעלה עד השורש עד שנעלה בפעם הראשונה מצד ימין ונחזיר צומת זה. אם מדובר באיבר הכי גדול נחזיר . סיבוכיותה כעומק העץ, שכן במקרה הגרוע אנו נטייל מעלה כלשהו ועד שורש העץ.
* - מחזירה את השדה של העץ . גישה לשדה בסיבוכיות .
* - מחזירה את השדה של העץ . גישה לשדה בסיבוכיות .
* - נאתחל מערך בגודל העץ. נמצא את המינימום של העץ באמצעות ונכניס את המפתח הראשון. נבצע פעולות ובכל פעם נכניס את המפתח המתאים למערך. סיבוכיות הפעולה, כפי שנלמד בתרגול, הינו . כל פעולות העזר שהשתמשנו הינן מסיבוכיות של אולם כפי שנלמד בתרגול הסיבוכיות הכוללת הינה .
* - נאתחל מערך בגודל העץ. נמצא את המינימום של העץ באמצעות ונכניס את המפתח הראשון. נבצע פעולות ובכל פעם נכניס את ערך הצומת המתאים למערך. סיבוכיות הפעולה, כפי שנלמד בתרגול, הינו . כל פעולות העזר שהשתמשנו הינן מסיבוכיות של אולם כפי שנלמד בתרגול הסיבוכיות הכוללת הינה .
* - מחזירה את השדה של העץ . גישה לשדה בסיבוכיות .
* - נבדוק תחילה אם שני העצים ריקים. אם כן, נחזיר עץ שבו רק הצומת המקשרת נמצאת. לאחר מכן נבדוק אם אחד העצים ריקים, אם כן, נוסיף את הצומת לעץ שאינו ריק, נבצע פעולות איזון באמצעות ונחזיר את גובה העץ שאינו ריק ועוד 1. אחרת, נבחין תחילה כי ישנם 4 מקרים זרים ומשלימים: העץ הנוכחי גבוה מהעץ הנוסף ועם ערכים גדולים יותר ממנו, העץ הנוכחי גבוה מהעץ הנוסף ועם ערכים קטנים יותר ממנו, העץ הנוכחי נמוך מהעץ הנוסף ועם ערכים גדולים ממנו, העץ הנוכחי נמוך מהעץ הנוסף ועם ערכים קטנים ממנו. כל המקרים סימטריים זה לזה ובה"כ נניח כי העץ הנוכחי גבוה יותר ובעל ערכים גדולים יותר מהעץ הנוסף. נטייל משורש העץ הנוכחי ועד שנגיע לצומת הראשונה שהיא בגובה זהה לגובה העץ הנוסף או אחד מעל. נחבר את הצומת המקשרת לשני העצים, נבדוק מקרה מיוחד שבו שני העצים באותו גובה ולכן אין לצומת המקשרת אבא. נעדכן את שדות (במקרים הזרים והמשלימים מעדכנים שדות אחרים בהתאם לאיזה עץ יש ערכים גדולים יותר). נבצע במידת הצורך פעולות איזון (כולל המקרה המיוחד ל- שהוסבר בהרצאה) על-ידי טיול מאביו של הצומת המקשרת (אם היא קיימת) ועד השורש של העץ הנוכחי. נעדכן את כל השדות המתאימים לכל השדות בזמן שמטיילים במעלה השורש, ונחזיר את הפרשי הגבהים בין העצים ועוד 1. סיבוכיות פעולה זו הינו כי במקרה הגרוע נעלה מעלה העץ ועד השורש וכן השימוש בפונקציות העזר הינו לכל היותר .
* - נמצא את הצומת המפריד באמצעות . נגדיר שני עצים חדשים שהשורש שלו הוא תת העץ השמאלי של הצומת המפריד, שהשורש שלו הוא תת העץ הימני של הצומת המפריד. נטייל מהצומת המפריד עד השורש. אם עלינו מצד ימין, נאחד באמצעות פעולת את הצומת שאליה עלינו, ואת תת העץ השמאלי של אותה צומת, לעץ . אחרת, עלינו משמאל ולכן נאחד באמצעות פעולת את הצומת שאליה עלינו, ואת תת העץ הימני של אותה צומת, לעץ . לאחר סיום הטיול, נעדכן את שדות לכל תת עץ באמצעות שימוש ב וכן נעדכן את שדה . נבצע השמה של שני העצים לרשימת עצים בגודל 2 כאשר עבורה האיבר במקום ה- הינו ובמקום ה- הינו ונחזיר את רשימה זו. במהלך הפעולה טיילנו במקרה הגרוע מהעלה ועד השורש ולמדנו בכיתה שלמרות פעולות ה- בדרך הסיבוכיות היא והשתמשנו לבסוף בפעולות החסומות ב אשר הינן בלתי תלויות ובסך הכל הסיבוכיות הינה .

שאלה 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי **i** | מספר חילופים **במערך ממוין-הפוך** | עלות החיפושים במיון AVL עבור **מערך ממוין-הפוך** | מספר חילופים **במערך מסודר אקראית** | עלות החיפושים במיון AVL עבור **מערך מסודר אקראית** |
| 1 | 1999000 | 38884 | 1001377 | 35990 |
| 2 | 7998000 | 85764 | 3963464 | 77981 |
| 3 | 31996000 | 187524 | 15984727 | 173861 |
| 4 | 127992000 | 407044 | 64626729 | 372355 |
| 5 | 511984000 | 878084 | 256675197 | 820189 |

1. נצטרך להשוות לכל מיקום בסדרה כל מיקום שגדול ממנה. כלומר, עבור האיבר ה- נשווה אותו ל האיברים האחרים. מכיוון שהמערך ממוין-הפוך, האיבר ה- יהיה גדול מכל שאר האיברים, האיבר ה- יהיה גדול מכל שאר האיברים שאחריו וכך הלאה כך שהתנאי מתקיים לכל איבר ונוצרת סדרה חשבונית נסכום אותה ונקבל: .

מכיוון שאנחנו מכניסים את האיברים בסדר ממוין-הפוך ואנו מבצעים לאיבר המקסימלי, הרי שבכל פעם אנחנו נצטרך להכניס מהשורש את האיבר שהתקבל בתור איבר מינימלי לעץ בשלב זה. כלומר בכל שלב נעלה מהמקסימלי אל השורש ומשם נרד אל המינימלי. כל זה ייקח לנו פעולות כפול קבוע מסוים. נבצע תהליך זה פעמים ובסך הכל נקבל

*- הוכחנו שוויון זה בתרגול 1*

1. *אכן נראה כי הערכים בטבלה מסעיף א' והניתוח מסעיף ב' מתאימים. בשני הגרפים ניתן לראות כי קיים מתאם כמעט מושלם בין כמות האיברים בעץ לבין מספר החילופים וגם בין כמות האיברים לבין עלות פעולות החיפוש. נבחין כי עבור מספר חילופים במערך ממוין-הפוך הערכנו כי הקשר הינו ולכן ביצענו טרנספורמציית שורש כדי לנסות לקרב את התוצאה לצורה שבה הקשר המצופה יהיה לינארי בהתאם להדרכה. בשל כך אנו יודעים שבהינתן מספר האיברים בעץ נוכל לנבא בדיוק רב את מספר החילופים וכן את עלות החיפושים במערך ממוין-הפוך באמצעות הנוסחאות שפיתחנו בסעיף הקודם.*
2. נבחין תחילה כי הגודל ש- מייצג הינו מספר האיברים שגדולים ממנו בעת הכנסתו, שכן זה מסביר את מספר החילופים שיהיו לאותו במערך. כלומר הוא האיבר ה- הכי גדול בעת הכנסתו. מכיוון שאנחנו עושים חיפוש ב אזי חיפוש האיבר ה- ייקח . ניקח בחשבון כי יש מקדם מסוים עבור ה- נכפיל במקדם זה את התוצאה. נסכום:

*כלומר*

שאלה 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי **i** | עלות join **ממוצע**  עבור split **אקראי** | עלות join **מקסימלי**  עבור split **אקראי** | עלות join **ממוצע** עבור  split של **האיבר המקסימלי** בתת העץ השמאלי | עלות join **מקסימלי** עבור  split של **האיבר המקסימלי** בתת העץ השמאלי |
| 1 | 2.5 | 10 | 2.4545 | 12 |
| 2 | 3 | 7 | 2.6363 | 13 |
| 3 | 2.916 | 5 | 2.8181 | 15 |
| 4 | 2.0714 | 6 | 3 | 16 |
| 5 | 2.6666 | 10 | 2.8461 | 17 |
| 6 | 2.4666 | 8 | 2.75 | 18 |
| 7 | 2.0555 | 7 | 3.285 | 19 |
| 8 | 3.0588 | 5 | 3.5 | 21 |
| 9 | 2.7 | 7 | 3.234 | 22 |
| 10 | 3.111 | 10 | 2.6 | 23 |

1. בפעולת אנו מטיילים מהצומת שקיבלנו עד לשורש ומבצעים פעולות איחוד עם תתי-העצים לקבלת עץ ו- כתלות באם עלינו מימין או משמאל במהלך הטיפוס. עלות פעולת ה הממוצעת תלויה ברצף פעולות ה שנעשו עם העץ ברצף או עם העץ ברצף. בה"כ, ומטעמי סימטריה, אם נבצע הרבה פעולות איחוד על ברצף, אז ממוצע הפעולות יהיה במקרה הגרוע שכן הפרש הגבהים בין כל תת-עץ חסום על-ידי 3. (ראו דוגמה על איחוד בין לבין Shape

   Description automatically generated). בכל פעם שנעשה חילוף כלומר מעבר מרצף פעולות איחוד לפעולת איחוד עם , פעולת איחוד זו תהיה עם עלות גבוהה, עם חסם עליון של רצף מספר פעולות איחוד עם ונכפיל הכל ב- ; וחסם תחתון רצף פעולות איחוד עם ונכפלי הכל ב- (הפרש הגבהים המינימלי והמקסימלי בעת מעבר מתת עץ שמאלי לימני). **במקרה של בחירת האיבר המקסימלי** בתת העץ השמאלי, הרי שביצענו רק פעולות איחוד עם עד השורש שם נבצע פעולת איחוד עם . פעולות האיחוד עם יהיו במקרה הגרוע ופעולת האיחוד עם תהיה כגובה תת העץ הימני של השורש, כלומר . בגלל שרוב הפעולות עלו במקרה הגרוע ופעולה אחת הייתה הרבה יותר יקרה, הממוצע יהיה סביב במקרה הגרוע עבור מאוד גדול.

**עבור תרחיש אקראי**, נבחין כי הסיכוי שהאיבר שייבחר יגרום לרצף בלעדי של איחוד עם או רק איחוד עם הינו קטן מאוד ((. כך שעלות פעולת ה תישאר סביב הממוצע, כלומר במקרה הגרוע. זאת מכיוון שכל שינוי מאיחוד עם לאיחוד עם מצמצם את המרחק בין גובה תת העץ השמאלי לתת העץ הימני והסתברותית אנו צופים שיהיו סדרות קצרות מאותו איחוד ויחסית הרבה חילופים (שכן סטטיסטית פחות סביר שייבחר באופן אקראי עץ שיגרום לרצף פעולות זהות גדול מאוד).

ניתוח זה מסביר מדוע בשני המקרים העלות הממוצעת נמצאת סביב הערך .

1. במקרה זה, נבצע רצף של פעולות איחוד עם ובסוף פעולת איחוד עם . כפי שהוסבר בסעיף ב', בתרחיש זה בעת האיחוד של תת העץ עם אנו נדרש לעלות גבוהה מאוד שכן הצומת שקיבלנו, הוא הצומת הגדול ביותר בתת העץ השמאלי. לכן, עץ ה יישאר ריק במשך כל פעולות האיחוד במעלה תת העץ השמאלי, עד שנגיע לשורש. בשלב זה, נאחד עץ ריק, , עם השורש וכל תת העץ הימני של השורש. לכן, עלות האיחוד תהיה כגובה תת העץ הימני ועוד אחד. זה החסם העליון וגם מסביר את העובדה שככל שהעץ עמוק יותר כך המספרים בטבלה גדלים.

1. מכיוון שמדובר בעץ עלות ה- המקסימלי חסומה מלמעלה על-ידי . מכיוון שהאקראיות היא אחידה על-פני בחירת מבנה העץ ובחירה האיבר בתוך העץ, נניח כי נבחר איבר כלשהו. נבחין כי הסתברותית הרצף הארוך ביותר אינו עולה על וזאת בשל החישוב הבא:

*יש לכל היותר מועמדים להיות בעלי רצף זה, הסיכוי הינו*

*כלומר מתקבל כי*

*ההסתברות שלא יהיה רצף חסום על-ידי 1 וכן נציב את הפיתוח שבצענו מעלה.*

*כעת נחסום מלמטה*

*נניח באופן שרירותי כי קיים רצף באורך . נחשב את ההסתברות לרצף זה*

*נבחין כי לא ייתכן שתהיה חלוקה גדולה יותר מ- רצפים כאלו. לכן, ההסתברות שלא תהיה סדרה כזו היא אפסית ולכן התוחלת תהיה*

*בסך הכל*

*כלומר היא חסומה על-ידי .*