חלק ניסויי/תאורטי

בשאלה זו נדון ב insertion-sort באמצעות AVL Finger Tree. המיון מתבצע באופן הבא: מכניסים את האיברים לפי הסדר (הלא ממוין) אל העץ, כאשר החיפוש בהכנסת כל איבר חדש מתחיל מהמקסימום הנוכחי, ובסיום מבצעים סריקת in-order לקבלת הסדר הממוין. עבור עלות בניית העץ, ננתח בנפרד את עלות החיפושים ואת מספר פעולות האיזון.

* לצורך הניתוח, נמיין מערכים בגדלים שונים. גודל המערך שנמיין יהיה כאשר , ואיבריו יהיו הטבעיים עד . למשל, עבור המערך בגודל , ועבור המערך בגודל .
* לכל גודל של מערך, נבצע 4 ניסויים נפרדים:
  + בניסוי הראשון נמיין מערך **ממוין**, מקטן לגדול.
  + בניסוי השני נמיין מערך **ממוין הפוך**, מגדול לקטן.
  + בניסוי השלישי סדר האיברים במערך יהיה **אקראי**.
  + בניסוי הרביעי ניקח מערך ממוין ועבור כל אינדקס פרט לאחרון , נבצע החלפה עם האיבר הבא בסיכוי חצי (שימו לב שייתכן שאיבר יוחלף מספר פעמים).

הערה: בסעיפים הבאים, עבור ניסויים אקראיים, יש לקחת את הממוצע על פני 20 ניסויים.

1. יש למלא בטבלה הבאה את סך עלויות האיזון **ללא גלגולים** עבור כל אחד מהניסויים. הסבירו מהו החסם העליון התאורטי על סך עלויות האיזון כולל גלגולים, והאם הערכים בטבלה מתאימים. לסיום, נמקו מדוע תוספת הגלגולים לספירה אינה משנה אסימפטוטית.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי i | עלות איזון במערך ממוין | עלות איזון במערך ממוין-הפוך | עלות איזון במערך מסודר אקראית | עלות איזון במערך עם היפוכים סמוכים אקראיים |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| ... |  |  |  |  |

1. בהינתן מערך בגודל נגדיר היפוך בתור זוג אינדקסים כך שמתקיים , ונסמן את מספר ההיפוכים הכולל במערך ב-. ניתן לשים לב כי באופן כללי וככל שיש פחות היפוכים כך המערך קרוב יותר לממוין. יש למלא בטבלה הבאה את מספר ההיפוכים במערך הקלט עבור כל אחד מהניסויים (מספיק עד ).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי i | מספר היפוכים במערך ממוין | מספר היפוכים במערך ממוין-הפוך | מספר היפוכים במערך מסודר אקראית | מספר היפוכים במערך עם היפוכים סמוכים אקראיים |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| ... |  |  |  |  |

1. יש למלא בטבלה הבאה את סך עלויות החיפוש עבור כל אחד מהניסויים.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר סידורי i | עלות חיפוש במערך ממוין | עלות חיפוש במערך ממוין-הפוך | עלות חיפוש במערך מסודר אקראית | עלות חיפוש במערך עם היפוכים סמוכים אקראיים |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| ... |  |  |  |  |

1. כדי לחסום מלמעלה את סך עלויות החיפוש באופן תאורטי, נסמן ב- את מספר האיברים לפני האיבר באינדקס שגדולים ממנו.
   1. הסבירו מדוע .

יהי מערך A. נסמן .

מהגדרת היפוך נסיק כי ההגדרה נכונה גם בכיוון השני ולכן ניתן להגדיר אותה כך –  
לכל זוג אינדקסים המקיימים A[i] < A[j] נספור היפוך אחד.

מכך נסיק כי במעבר על כל האינדקסים. (מההגדרה נסיק כי

וספירת כל האיברים לפניהם הגדולים מהם נקבל את כמות ההיפוכים הנדרשת.  
**בהתאם על פי הגדרת נקבל כי**  .

* 1. חסמו את עלות החיפוש בעת הכנסת האיבר ה- כפונקציה של , והסיקו כי סך עלויות החיפוש לאורך סדרת ההכנסות הוא .  
     הערה: כאשר הסיבוכיות היא עבור , ניתן להמיר לביטוי כדי לעבור לביטוי אחד פשוט וחוקי שחוסם את שניהם מלמעלה.

יהי מערך A, נסמן .

נעבור על כל האינדקסים . (מההגדרה נסיק כי

מכיוון שאנו משתמשים ב-finger insert, בכל חיפוש נתחיל מהמקסימום, נעלה למעלה עד שנגיע לצומת מתאים לדרישה ואז נתחיל לרדת עד שנגיע לצומת מתאים להכנסה.

נחלק למקרים –

* . מכך נסיק כי הצומת החדש יוכנס כבן למקסימום. בהתאם, סך עלות החיפוש כלפי מעלה תהיה 1.
* . מכך נסיק כי נטפס למעלה לפחות צומת אחד ואז נתחיל לרדת.  
  כפי שנלמד על finger insert, הטיפוס למעלה חסום על ידי לוג 2 של כמות האיברים הגדולים מ-A[i] שכבר נמצאים בעץ. בהתאם, עלות החיפוש במקרה זה תהיה .

על-פי ההערה נסיק כי עלות החיפוש כלפי מעלה היא

בהתאם לאחר מעבר על כל האינדקסים נקבל כי עלות החיפוש כלפי מעלה תהיה . ועל-פי חוקי לוגריתמים נקבל .

ממבנה עץ AVL נסיק כי עלות החיפוש כלפי מטה חסומה על-ידי ביטוי זה +1.

בהתאם, **נקבל כי עלות החיפוש חסומה על-ידי .**

* 1. השתמשו באי-שוויון הממוצעים כדי לחסום מלמעלה את סך עלויות החיפוש כפונקציה של . הסבירו מדוע זוהי העלות הדומיננטית מבחינה אסימפטוטית.

תחילה נשים לב כי על-פי חוקי לוגריתמים מתקיים: (\*)

בנוסף, על-פי אי-שיוויון הממוצעים מתקיים   
מהגדרת I וסעיף א׳ נסיק כי.   
נציב זאת חזרה בביטוי ב-(\*) ונקבל

**סה״כ קיבלנו כי עלות החיפוש חסומה אסימפטוטית על-ידי .**

* 1. השוו בין החסם שהתקבל ותוצאות הניסויים.  
       
     ההשוואה מראה לנו כי אכן זהו חסם לעלות החיפוש.  
     ניתן לשים לב כי בניסוי הממוין, המקרה בעל עלות החיפוש הכי קטנה, עלות החיפוש חסומה ב- . (מכיוון שמתקיים I=0)  
     וכי בניסוי הממוין הפוך, המקרה בעל עלות החיפוש הכי גבוהה, עלות החיפוש חסומה ב-.  
     אלו הם מקרי הקצה, בשאר המקרים עלות החיפוש נמצאת בין שני המקרים ובהתאם גם כן חסומה על-ידי החסם של סעיף ג׳ – .