

מבני נתונים – תרגיל 3

תאריך פרסום: 10.5.2022

תאריך הגשה: 24.5.2022 בשעה 23:59

מרצה ומתרגלים אחראים: פרופ' דקל צור, אורי לוביש

הנחיות:

- הגשת העבודה הינה ביחידים או בזוגות, לבחירתכם. אנו ממליצים לעבוד בזוגות, כדי לעודד דיון והפריה הדדית. למען הסר ספק- גם אם העבודה הוגשה בזוג, כל סעיפי העבודה המוגשים צריכים להיות תוצאה של עבודה משותפת.
- חובה להגדיר קבוצה באתר המודל ולהשתייך אליה על מנת להגיש את העבודה, גם אם בחרתם להגיש לבד (ואז אתם בקבוצה המכילה רק אתכם). למידע נוסף בקרו בעמוד ההרשמה לקבוצה.
- העבודה חייבת להיות מוקלדת ומוגשת כקובץ בפורמט Pdf. שם הקובץ שלכם צריך להיות זהה לשם הקבוצה אליה נרשמתם באתר המודל, לדוגמא: Group300_Assignment3.pdf.
- שאלות לגבי העבודה יש לשאול בפורום הייעודי באתר המודל או בשעות קבלה של המרצה/המתרגלת האחראיות על העבודה. הפורום נועד לדיון בין הסטודנטים בנוגע לעבודה והצוות האחראי על העבודה לא עונה על שאלות בפורום. אחת ל-24 שעות הצוות האחראי על העבודה יקרא את השאלות שעלו בפורום ויפרסם הבהרות ב-FAQ-במידת הצורך.
- אין צורך לפרט דברים שנלמדו כיתה. עם זאת, יש להוכיח כל טענה שלא נלמדה בהרצאה או בתרגול.
- יש לנתח את זמן הריצה בצורה ההדוקה ביותר תוך התחשבות במקרה הגרוע ביותר.
- $\log()$ מתייחס ללוגריתם בבסיס 2.

שאלה 1 – עצי AVL (27 נקודות)

סעיף א' (9 נק'): הראו שאם x הוא עלה בעץ AVL אז $d(x) \geq \left\lfloor \frac{h}{2} \right\rfloor$ כאשר $d(x)$ הוא העומק של x ו- h הוא גובה העץ. באילו עצים החסם הדוק?

סעיף ב' (9 נק'): הראו שפעולות מחיקה בעץ AVL דורשת $\Omega(\log n)$ סיבובים. כלומר, עליכם להראות משפחה אינסופית של עצי AVL $\{T_n\}_{n \in \mathbb{N}}$ כך שפעולת מחיקה מסוימת בעץ T_n תדרוש לפחות $c \cdot \log(|T_n|)$ סיבובים, כאשר c הוא קבוע חיובי לא תלוי ב- n ו- $|T_n|$ הוא מספר הצמתים בעץ T_n .

סעיף ג' (9 נק'): הראו כיצד ניתן, ע"י תוספת זכרון, לתמוך בפעולת $\text{Find}(x, k)$ שתחזיר את k האברים הראשונים שגדולים מהמפתח x בעץ בזמן $O(\log n + k)$. אם מספר האברים עם מפתח גדול מ- x קטן מ- k יש להחזיר את כל האברים האלו.

שאלה 2 - קבוצות (16 נק')

הציעו מבנה נתונים למימוש ה-ADT הבא. פרטו איך תתבצע כל פעולה ונמקו את יעילותה.

תיאור ה-ADT: הערכים הנשמרים הם זוגות של $(key = (x,y), value=height)$. המפתחות הם נקודות ציון גאוגרפיות המיוצגות ע"י זוגות של קו אורך (x) וקו רוחב (y) כאשר x,y מספרים ממשיים, כלומר המפתחות הם מהצורה (x,y) . הערך $(height)$ הוא מספר ממשי המתאר את הגובה של נקודת הציון (x,y) .

פעולה	תיאור	זמן נדרש
$Init()$	אתחול מבנה נתונים ריק	$O(1)$
$insert(x,y, height)$	הכנסת נתון למבנה הנתונים. ניתן להניח שלא קיים עוד נתון בעל קואורדינטות אורך ורוחב זהות במבנה לפני פעולת ההכנסה.	$O(\log n)$
$delete(x,y)$	מסירה את הזוג $(x,y, height)$ מהמבנה. אם אין זוג כזה יש להדפיס הודעת שגיאה.	$O(\log n)$
$find(x,y)$	מחזירה את הגובה של נקודת הציון (x,y) אם קיימת כזו במבנה ובמידה ולא קיימת יוחזר NULL.	$O(\log n)$
$printAll(x)$	הדפסת הגבהים של כל נקודות הציון שקואורדינטת האורך שלהן שווה ל- x .	$O(\log n + k)$ כאשר k הוא מספר הנתונים בעלי קואורדינטת אורך x .

הבהרה: n הינו מספר הנתונים במבנה ברגע נתון. תארו בקצרה את מבנה הנתונים, ספקו אלגוריתם לכל אחת מהשיטות והוכיחו עמידה בזמני הריצה.

שאלה 3 - מיזוג/פיצול (48 נק')

סעיף א' (16 נק'): בהינתן שני עצי AVL, T_1 בגובה h_1 ו- T_2 בגובה h_2 , כך ש- $\max T_1 < \min T_2$ (כל מפתח ב- T_2 גדול מכל מפתח ב- T_1), ומפתח נוסף x המקיים $\max T_1 < x < \min T_2$, תארו אלגוריתם למיזוג T_1, x, T_2 לעץ AVL חדש בזמן $O(|h_1 - h_2|)$. כלומר, בנו עץ AVL חדש המורכב מכל מפתחות T_1 , כל מפתחות T_2 , ו- x .

הדרכה: אם $h_1 > h_2$, נסו להכניס את x לנקודה המתאימה במסלול הימני ביותר מהשורש לעלה בעץ (המסלול המתקבל מהליכות ימינה בלבד). הראו שהעץ שהתקבל ניתן לתיקון על ידי סיבובים. חשבו מה לעשות במקרה הסימטרי של $h_1 \leq h_2$.

סעיף ב' (16 נק'): פעולת $\text{Split}(T, k)$ על עץ AVL T ומפתח k מחזירה שני עצי AVL T_1, T_2 כאשר T_1 מכיל את כל המפתחות ב- T שקטנים או שווים ל- k , ו- T_2 מכיל את כל המפתחות ב- T שגדולים מ- k . תארו מימוש לפעולת $\text{Split}(T, k)$ בזמן $O(\log n)$ כאשר n כמות האברים בעץ T .

סעיף ג' (16 נק'):

תארו מבנה נתונים השומר אברים שלכל אחד מהם יש מפתח וצבע. הצבע הוא שחור או לבן. המבנה צריך לתמוך בפעולות הבאות:

זמן נדרש	תיאור	פעולה
$O(\log n)$	הכנס את האבר x (כאשר ל- x יש שדות $x.color$ ו- $x.key$, הניחו שהמפתחות הם מספרים טבעיים)	$Add(x)$
$O(\log n)$	החזר את הצבע של אבר עם מפתח k ($null$ אם לא קיים)	$Color(k)$
$O(\log n)$	החלף את הצבעים של כל האברים במבנה עם מפתח קטן או שווה ל- k . כלומר, אם הצבע של אבר כזה היה לבן הוא יהפוך לשחור ואם הצבע היה שחור הוא יהפוך ללבן.	$FlipColors(k)$

שאלה 4 – B-trees

בכיתה ראינו כיצד ע"י הוספת שדה $size$ השומר את גודל תת העץ המושרש בצומת ניתן, בהינתן עץ חיפוש מאוזן, למצוא את האיבר ה- k בסדר ההיררכי של המבנה (כלומר, האיבר שיש $k-1$ איברים קטנים ממנו) בזמן של $O(\log n)$. הראו כיצד ניתן לענות על אותה שאלתא בעץ B עם פרמטר t (כלומר מציאת האיבר ה- k בגודלו בעץ), שוב ע"י הוספת שדה יחיד לכל קודקוד. ניתן להניח שכל המפתחות בעץ שונים זה מזה. מה יהיה זמן הריצה של השאלתא?