מבני נתונים
תרגיל 5

גלעד אשרוב נטע ברקאי צבי קופולוביץ'

23 באפריל 2012

תאריך הגשה: בתרגול 6, בקבוצת התרגול (תאריך אחרון: 02.05.12)

ההגשה ביחידים. מותר להתייעץ ולפתור את התרגילים בקבוצה אך יש לכתוב את הפתרונות באופן **עצמאי**. חל איסור מוחלט להחזיק פתרון כתוב של סטודנט אחר.

שאלה 1.

א. תאר אלגוריתם המקבל כקלט רשימה מקושרת ממויינת ומחזיר רשימת דילוגים אידיאלית *ideal skip list*. מהו זמן הריצה של האלגוריתם שהצגת?

ב. מה ההסתברות שבהכנסת n איברים לרשימה, נגיע למצב שבו מספר הרמות מגיע ל - $\omega(\log n)$?

ג. כדי להימנע ממצב שבו מספר הרמות גדל וגדל ללא הגבלה, סטודנט הציע את השיפור הבא: כאשר מספר הרמות גדול מ־$5\log n$, נעצור, ונבנה מחדש את הרשימה כרשימת דילוגים אידיאלית. ברור כי זמן חיפוש הוא בתוחלת $O(\log n)$, ומספר הרמות במקרה הגרוע הוא לכל היותר $5\log n$. מהי תוחלת זמן ההכנסה?

ד. הצג אלגוריתם לאיחוד שתי רשימות דילוגים. מהו זמן הריצה? הסבר מדוע הפתרון שלך הוא האופטימלי.

שאלה 2. בד"כ, יותר קל למצוא איבר "קרוב" אחרי שכבר מצאנו איבר בסביבה. לשם כך, נתבונן בשאילתא הבאה: בהינתן צומת במבנה הנתונים המחזיק את המפתח x ובהינתן מפתח - y, נרצה למצוא את הצומת שמחזיק את y.

תאר אלגוריתם העונה לשאילתא הנ"ל ברשימת דילוגים, העובד בזמן $O(\log(2+|rank(x)-rank(y)|))$ בתוחלת. הוכח את תשובתך.

הערה: $rank(x)$ מוגדר להיות ה"מיקום" של x ברשימה. כלומר, בהינתן הרשימה $1,2,5,7,9$ נקבל כי: $rank(5)=3$, $rank(9)=5$ וכן על זה הדרך.

שאלה 3. (ממבחן) נתבונן בשיטה הבאה למיון קבוצה A_0 של n מספרים שונים: נחלק את A_0 ל־$\sqrt{n}$ תת־קבוצות זרות בגודל $\sqrt{n}$ כל אחת. אח"כ נבחר בכל פעם את האיבר הקטן ביותר מתוך אלה שעוד לא נבחרו ע"י שנמצא (בזמן לינארי) את האיבר הקטן ביותר בתוך כל אחת מ - $\sqrt{n}$ הקבוצות (נקרא לקבוצות $\sqrt{n}$ הקטנים האלה A_1), ואז נמצא את האיבר הקטן ביותר ב - A_1.

א. כמה זמן לוקח מציאת האיבר הקטן ביותר? והשני? ה - ℓ, כאשר $\ell=2,3,\ldots,n$? סכם - מה הסיבוכיות של שיטת מיון זו?

ב. האם יעזור אם, במקום לחזור ולפש שוב ושוב את האיבר הקטן ביותר בכל קבוצה, נמיין את $\sqrt{n}$ הקבוצות מיד בהתחלה?

הנח: משתמשים במיון מיזוג.

ג. כדי לשפר נחלק את A_0 ל - $n^{2/3}$ קבוצות בגודל $n^{1/3}$ כל אחת. נקבל קבוצה A_1 בגודל $n^{2/3}$ ולכן גם אותה נחלק ל - $n^{1/3}$ קבוצות בגודל $n^{1/3}$. נקרא לקבוצת האיברים המינימאליים בשם A_2. כמה זמן לוקח מציאת האיבר הקטן ביותר? והשני? ה - ℓ, n, $\ell = 2, \ldots, n$? סכם: מהי הסיבוכיות של שיטת מיון זו?

ד. תאר בקצרה הכללה של השיטה של סעיף א' (רמה נוספת אחת) ושל סעיף ג (2 רמות נוספות) ל - k רמות נוספות. מה תהיה הסיבוכיות?

ה. מהו הערך הגדול ביותר של k שהוא אפשרי בסעיף ג (רמז: חייבים להיות לפחות 2 איברים בכל קבוצה בחלוקה). מה תהיה הסיבוכיות במקרה זה?

שאלה 4. נניח שנשתמש באלגוריתם Boyer and Moore רק ב - Δ_1 בלי Δ_2 ונכתוב:

$$i \quad \leftarrow \quad i + \Delta_1(T(i))$$

במקום:

$$i \leftarrow i + \max(\Delta_1(T(i)), k+1)$$

בנה דוגמה של תבנית באורך 6 מעל א"ב $\{A, B, C\}$ וקטע טקסט מתאים שעבורם התוכנית תיכנס ללולאה אינסופית.

שאלה 5. באלגוריתם של Boyer and Moore הגדרנו את Δ_2 עם שני תיקונים:

- **תיקון 1:** לכל סיפא נחפש מופע מוקדם יותר, אך נדרוש שהאות הקודמת לסיפא תהיה שונה.
- **תיקון 2:** בחיפוש אחר מופע מוקדם יותר, נאפשר גלישה של המופע מעבר להתחלת המחרוזת.

סמן "כן" או "לא" במשבצות המתאימות בטבלה הבאה, והוסיפו לאחר מכן הסברים.

	אם נשתמש בתיקון 2 אך לא בתיקון 1	אם נשתמש בתיקון 1 אך לא בתיקון 2
האלגוריתם לא יעבוד האלגוריתם יעבוד, אך יהיה פחות מהיר האלגוריתם יעבוד, מהירותו לא תשתנה, אך הוא ידרוש יותר זיכרון האלגוריתם יעבוד, מהירותו לא תשתנה, אך הזכרון הדרוש יקטן		