



מבני נתונים ב07

תרגול 7

13/6/2007

מיון ובחירה

ליאור שפירא

תרגיל 1

□ בהינתן מערך בעל n אברים, מצא את k האברים הקטנים ביותר במערך בזמן $O(n)$

□ פתרון

■ בהנחה שאנו יודעים מיהו האבר ה- k הקטן ביותר (ז"א k statistic order), כמה זמן דרוש למצוא את שאר האברים הדרושים?

■ סה"כ

□ $O(n)$ למצוא את האבר הא קטן ביותר

□ $O(n)$ לעבור על המערך ולסמן את כל הקטנים ממנו

תרגיל 2

□ אבר מסוים מופיע במערך $n/5$ פעמים, באיזו מהירות ניתן למצוא אותו?

■ האבר חוזר במערך הרבה פעמים



■ מה אם המערך היה ממורכז?



■ איך ננצל את העובדה הזו?



$n/6$
intervals



תרגיל 2

פתרון

■ נמצא את ה-order statistics הבאים

■ $n/6$

■ $2n/6$

■ $3n/6$

■ ...

■ אחד מהם חייב להיות בוודאות האיבר המופיע $n/5$ פעמים

■ סה"כ סיבוכיות - לינארית

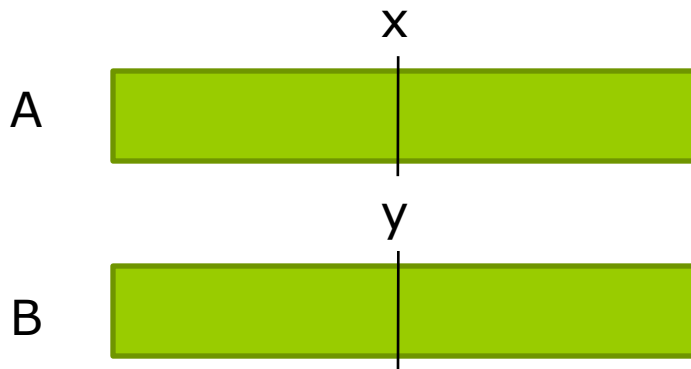
תרגיל 3

□ מציאת חציון משותף לשני מערכים ממוינים שאין להם אברים משותפים

□ פתרון

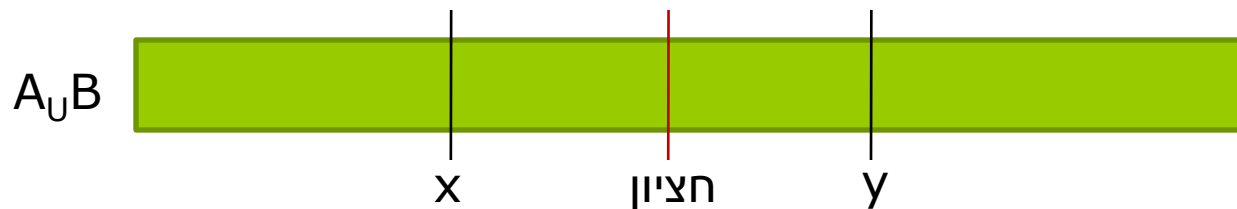
■ זמן לינארי?

■ זמן לוגריתמי?



□ נמצא חציון של כל אחד מהם, נניח בה"כ כי $x < y$

□ איפה יהיה החציון של האיחוד?



□ ניפטר מכל מה שמתחת לא וכל מה שמעל ל y

□ נחזור על התהליך רקורסיבית

מציאת חציון בזמן לינארי

- Step 1: If n is small, for example $n < 6$, just sort and return the k^{th} smallest number in constant time i.e; $O(1)$ time.
- Step 2: Group the given number in subsets of 5 in $O(n)$ time.

מציאת חציון בזמן לינארי

- Step3: Sort each of the group in $O(n)$ time. Find median of each group.
- Given a set
(.....2,5,9,19,24,54,5,87,9,10,44,32,21,1
3,24,18,26,16,19,25,39,47,56,71,91,61,4
4,28.....) having n elements.

מציאת חציון בזמן לינארי

- Arrange the numbers in groups of five

.....	2	54	44	4	25	
.....	5	5	32	18	39	
.....	9	87	21	26	47	
.....	19	9	13	16	56	
.....	24	10	2	19	71	

Find median of $N/5$ groups

..... (2) (5) (2) (4) (25)

..... (5) (9) (13) (16) (39)

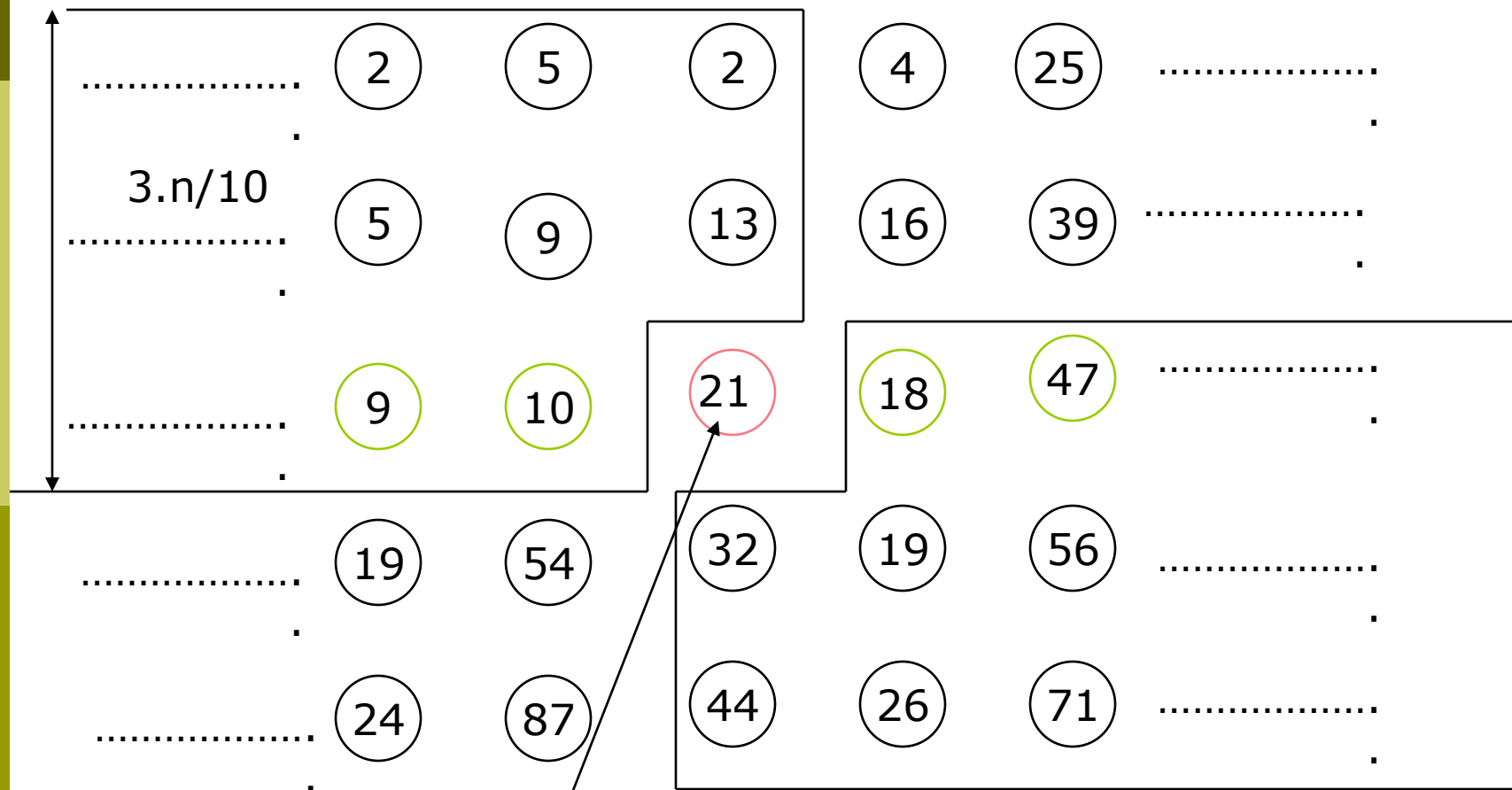
..... (9) (10) (**21**) (18) (47)

..... (19) (54) (32) (19) (56)

..... (24) (87) (44) (26) (71)

Median of each group

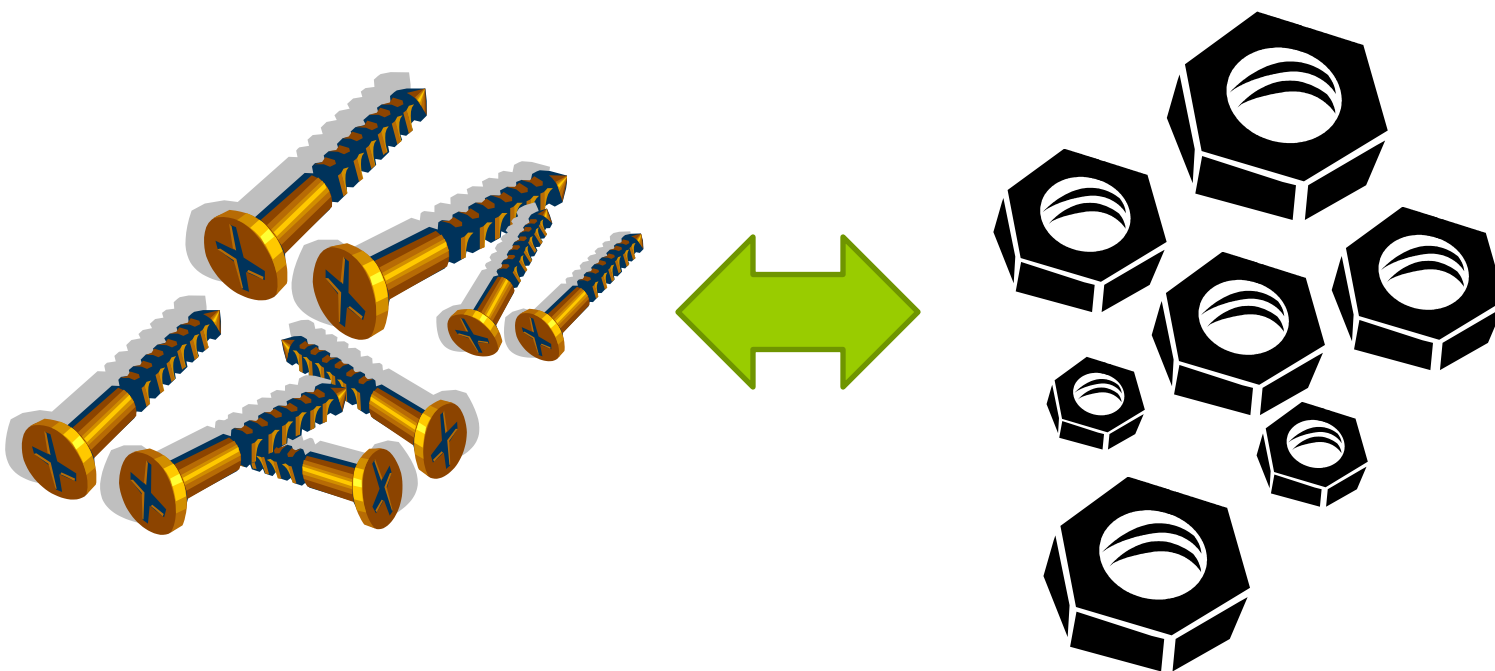
Find the Median of each group



Find m , the median of medians

תרגיל 4: ברגים ואומים

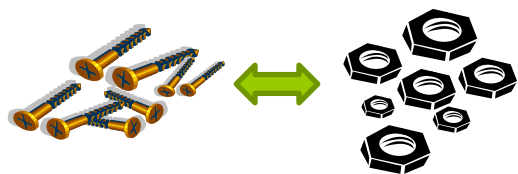
□ נתונים החלקים הבאים:



המטרה: להתאים כל אחד מ-ח הברגים ל-ח האומים (יש התאמה יחידה)



תרגיל 4: ברגים ואומים



□ נתונים החלקים הבאים:

□ צריך להתאים כל אחד מ- n הברגים ל- n האומים (יש התאמה יחידה)

□ הפעולה היחידה שמותרת: השוואת בורג לאום

□ התשובה: קטן, גדול או מתאים

□ כיצד ניתן לפתור את הבעיה?

1. הפתרון הבנאלי

2. אלגוריתם רנדומי

3. אלגוריתם דטרמיניסטי



תרגיל 4: ברגים ואומים

פתרון 1: אלגוריתם רנדומי

- נבחר בורג באקראי ונמצא לו אום מתאים
- נסדר את האומים ל-2 קבוצות
 - קטנים מדי עבור הבורג
 - גדולים מדי עבור הבורג
- נסדר את הברגים ל-2 קבוצות לפי האום המתאים לבורג הנבחר
 - האום המתאים קטן מדי
 - האום המתאים גדול מדי
- נמשיך רקורסיבית
- מה לפי דעתכם הסיבוכיות בתוחלת?



תרגיל 4: ברגים ואומים

□ פתרון 2: אלגוריתם דטרמיניסטי

■ יותר מסובך ממה שזה נשמע

□ בשנת 94 נמצא פתרון בזמן

□ בשנת 96 נמצא פתרון בזמן $\log n$ אך הוא מאד מסובך, דה-רנדומיזציה של האלגוריתם... לא נעים.

תרגיל 5

□ במערך אבר מסוים מופיע $n/2$ פעמים, אחר $n/4$ פעמים, הבא $n/8$ פעמים וכו'

■ כיצד ניתן למיין את המערך ובאיזו מהירות?

□ פתרון

■ ניתן לעשות זאת בזמן לינארי

■ מוצאים את זה שמופיע הכי הרבה, רצים ומוציאים אותו מהמערך, ממשיכים רקורסיבית ובסוף ממיינים את $\log n$ הערכים

■ אפשר להראות למה החסם התחתון לינארי

$$\frac{n!}{(n/2)!(n/4)!(n/8)!...}$$

Hashing

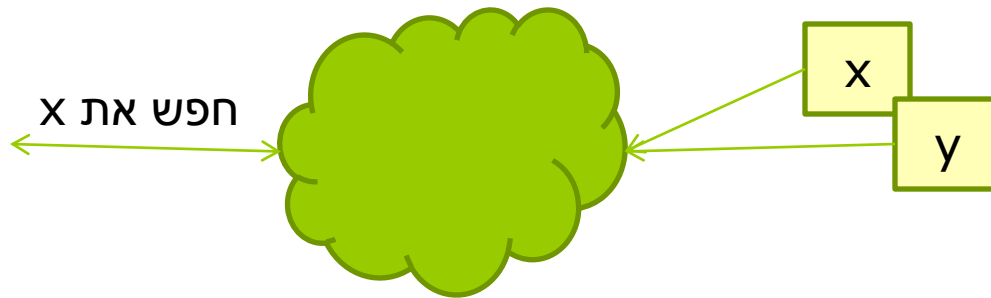
□ נרצה ליצור מבנה נתונים כך ש-

■ יכול להכיל את המספרים בין 1 ל- m

■ תומך בפעולות הבאות

Insert □

Find □



□ אילו מבני נתונים ראינו שתומכים בפעולות אלו?

■ עצי חיפוש (הכנסה ומציאה בזמן לוגריתמי)

■ Bit vectors (הכנסה ומציאה בזמן קבוע)

Hashing

□ מבנה הנתונים יכול ח אברים

□ נניח שקיימת פ' רנדומית הממפה $[1, m] \rightarrow [1, 2n]$



□ התנגשות: כאשר $x \neq y, f(x) = f(y)$

□ הסיכוי להתנגשות מאד נמוך, תוחלת ההתנגשויות מאוד נמוכה

Hashing

□ סוגי Hash שונים

■ Universal hash

□ פ' המבטיחה כי לא ניתן "להכריח" ביצועים רעים על הפונקציה שלנו (מספר התנגשויות גבוה)

□ מגדיר משפחה של פונקציות שנבחר ביניהן באקראי

□ תוחלת פעולה בהכנסת n אברים **שונים** היא $O(1)$

■ Closed vs. Open hashing

□ ב-Closed hash נאחסן את הערכים ישירות בתוך המערך

□ ב-Open hash כל תא במערך יצביע לרשימה מקושרת (או מבנה דומה)

להכניס אברים זהים ל-Hash הורס את חישובי היעילות!

תרגיל 5

□ מיינו מערך A ובו כל אבר חוזר k פעמים

■ תרגיל לבית

□ הראו חסם תחתון $\Omega(n \log(n/k))$ למיון.

□ הראו חסם עליון דומה

■ כיצד ניתן לפתור ע"י אלגוריתם רנדומי?

□ נשתמש בhash (לא במודל ההשוואות)

□ נספור כמה יש מכל אבר ונכניס רק עותק אחד

□ נמין n/k אברים ולכן חסם הרבה יותר טוב – $(n/k) \log(n/k)$