בהנחה שאברי המערך הם מס' שלמים בין 0 ל-k-1

Counting Sort (A[0...n-1])

```
Θ(1) < create an array C[0...k]
       create an array B[0...n-1]

Θ(k) { for i ← 0 to k
           C[i] ← 0

Θ(n) { for i ← 0 to n-1                         (2*)
           C[A[i]] ← C[A[i]] + 1

Θ(k) { for i ← 1 to k                           (3*)
           C[i] ← C[i] + C[i-1]

Θ(n) { for i ← 0 to n-1
           B[C[A[i]]] ← A[i]
           C[A[i]] ← C[A[i]] - 1
```

סיבוכיות כוללת: $\Theta(n+k)$.

(או למעשה $\Theta(\max\{n,k\})$)

לדוגמה: $(k=3)$

A: | 1 | 0 | 3 | 1 | 1 | 3 |

C (0 1 2 3): | 1 | 3 | 0 | 2 | → | 1 | 4 | 4 | 6 | (3*)

B: | 0 | 1 | 1 | 1 | 3 | 3 |

(*) הקצאת מערך, לא עבור מערכים גדולים, עולה מעט במונחי סיבוכיות רגילה, $\Theta(1)$.

(2*) כאן אנו בעצם מבצעים ספירה של כמות האיברים מכל סוג במערך A. אנו בעצם משתמשים בערך של הנתון ממערך A כאינדקס במערך C. ניתן לעשות זאת בהסתמך על ההנחה שאברי A מכילים מספרים שלמים בין 0 ל-k.

(3*) לאחר הרצת לולאה זו, נקבל במערך C את המיקום האחרון בו מופיע הערך של אינדקס i במערך A.

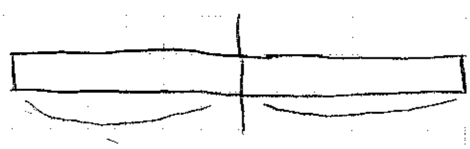
הטענה שהוכחנו לגבי מספר ההשוואות הנחוץ באלגוריתם מיון היתה תחת ההנחה כי הפעולות היחידות* שמתבצעות במהלך האלגוריתם הן העתקה של ערכים מתא לתא (הזכרון) והשוואה של ערכים הנמצאים בתאים. ואסור לבצע פעולה חשבון כגון חיבור וכפל. מודל מצומצם זה (ביחס למודל RAM) קרוי מודל ההשוואה.

* של הערכים שבתאי הזכרון.

האם זו הנחה סבירה (כאשר ממיינים קלט כללי)?
במשך הרבה שנים האמינו שבאמת שימוש בפעולות כגון חיבור וכפל לא יעזרו למיין קלט כללי, ואכן כל האלגוריתמים שהיו ידועים למיון קלט כללי השתמשו רק בהשוואות.
ב-1990 נמצא אלגוריתם שמשתמש בחיבור וכפל וממיין קלט של n מספרים בזמן $O(\frac{n\log n}{\log\log n})$.

מיון מיזוג Merge Sort

(במסגרת מודל ההשוואות)



מחלקים את המערך ל-2 תתי מערכים באורך "$\frac{n}{2}$". ממיינים כל אחד מהם רקורסיבית ב-Merge Sort ואז ממזגים את 2 המערכים הממוינים למערך אחד.

כיצד ממזגים?

$$\text{Merge}(\overbrace{A[0\ldots K-1],\ B[0\ldots l-1]}^{\text{מערכים ממוינים}})$$

$$\theta(1)\begin{cases}\text{create an array } C[0\ldots(k+l)-1]\\ a\leftarrow 0\\ b\leftarrow 0\end{cases}$$

יש $k+l$ איטרציות. וכל אחת בזמן של $\theta(1)$ ולכן המיזוג לוקח זמן $\theta(k+l)$.

```
for i <- 0 to k+l-1
    do if a=K then C[i] <- B[b]
                   b <- b+1
       else
          if b=l then C[i] <- A[a]
                      a <- a+1
          else if A[a]<B[b] then C[i] <- A[a]
                                 a <- a+1
               else C[i] <- B[b]
                    b <- b+1
```

מהי הסיבוכיות של המיון כולו?

(*) כיוון שמדובר בתהליך רקורסיבי, עלינו לסמן את מה שאנו מחפשים.
נסמן ב- $T(n)$ את זמן הריצה של האלגוריתם על קלט באורך n. במקרה הגרוע ביותר.

$$T(n)=\underbrace{2T(\lceil\tfrac{n}{2}\rceil)}_{\text{קריאות רקורסיביות}}+\underbrace{\theta(n)}_{\text{מיזוג}}\le 2T(\lceil\tfrac{n}{2}\rceil)+c_2n\le 2\left[2T(\lceil\tfrac{\lceil\frac{n}{2}\rceil}{2}\rceil)+c_2\lceil\tfrac{n}{2}\rceil\right]+c_2n \le$$

$$c_1n\le\ \ \le c_2n$$

$$2^{r-1}<n\le 2^r$$

$$\frac{n}{2}\le 2^{r-1}$$

$$\lceil\tfrac{n}{2}\rceil\le 2^{r-1}$$

$$\left\lceil\frac{\lceil\frac{n}{2}\rceil}{2}\right\rceil\le 2^{r-2}$$

$$\vdots$$

$$2^3\left\lceil\frac{\lceil\frac{\lceil\frac{n}{2}\rceil}{2}\rceil}{2}\right\rceil\le 2^r=2\cdot 2^{r-1}<2n$$

לאחר r צעדים

$$\le 2^r\cdot\overbrace{T(1)}^{c_3}+c_2\Big(\underbrace{\ldots+2^2\left\lceil\tfrac{\lceil\frac{n}{2}\rceil}{2}\right\rceil+2\lceil\tfrac{n}{2}\rceil+n}_{\text{סכום } r \text{ מחוברים}}\Big)\le c_3 2^r+c_2 r\cdot(2n)\le$$

$$\le c_3 2n+c_2\lceil\log_2 n\rceil\cdot 2n=O(n\log_2 n)$$

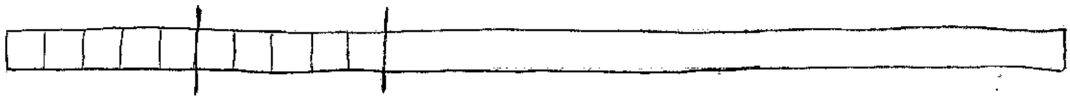
אלגוריתם בזמן לינארי ($O(n)$) למציאת חציון במערך (לא ממוין) ובאופן כללי מציאת האיבר ה-i בגודלו (i מבוקש בין 1 ל-n).

<u>הקלט:</u> מערך $A[0...n-1]$ ואינדקס $0 \le i \le n-1$

<u>הפלט:</u> האיבר $B[i]$ מהמערך B שיתקבל מ-A לאחר מיון.

<u>האלגוריתם:</u>

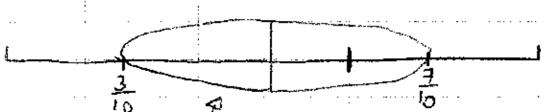
1. נחלק את המערך המקורי ל"חמישיות"



2. בכל חמישיה מוצאים את החציון (המקומי)
3. ע"י קריאה רקורסיבית נמצא את החציון <u>המדויק</u> של $\lfloor \frac{n}{5} \rfloor$ חציונים שמצאנו. זהו <u>לא בהכרח</u> החציון המדויק של המערך אבל זהו חציון מקורב במובן <u>הבא:</u>

(נניח לשם הפשטות $n = 5(2r+1)$ וכל איברי המערך המקורי שונים.
כמה "חציונים" $\le$ מ"החציון המקורב"? בדיוק $r+1$ (כולל שלו)
בכל חמישיה יש בדיוק 3 איברים $\le$ החציון של החמישיה.

לפיכך, יש לפחות $3(r+1)$ איברים במערך המקורי שהם גדולים או שווים מ"החציון המקורב"



בתחום זה נמצא ה"חציון המקורב".