**ממ"ן 14 – פרוייקט 1 בקורס מבני נתונים ומבוא לאלגוריתמים**

(האלגוריתם מומש בשפת JAVA ומצורף בקובץ JAVA)

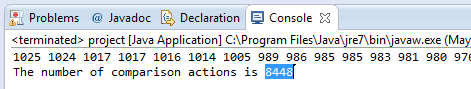
השפעת מספר החלוקה על יעילות אלגוריתמים

בפרויקט זה נוכיח כי מספר פעולות ההשוואה נובע אך ורק משיקולים מתמטיים אותם ניתן לחשב באופן ברור לאחר חקר יעילות הריצה, וכי המספר שיש לבחור לפי סוג המיון\אלגוריתם אינו יכול להיות שרירותי.

נעשה זאת בעזרת חקר אלגוריתם אשר הפלט שלו הינו מערך ממוין בסדר יורד.

המסקנה העיקרית כמובן ממחקר זה הוא כי ניתן לחשב באופן מדויק כיצד יש לחלק מערך על פי האלגוריתם וזמן הריצה שלו.

את מספר ההשוואות קיבלתי בעזרת הוספת counter לפני כל פעולת השוואה על המערך הראשי או התתי מערכים שאנו מקבלים, והצגתי אותו בסוף הרצת האלגוריתם באופן הבא:



את מספר ההשוואות בדקתי על:

1. מערך בסדר עולה.
2. מערך בסדר יורד.
3. מערך לא ממוין.

בכל אחד מסוגי המערכים האלה בדקתי את ערכי K הבאים:

1. K = 10
2. K = n/2 (128)
3. K = 2
4. K = 4
5. K = 8 (שורש N)
6. K = 12 (בקירוב שורש של (N\2))

זמן הריצה של האלגוריתם שכתבתי היה O(N\*K + N/(N^2/2K)

כאשר גוזרים את האלגוריתם לפי K ומשווים ל0, ערך הK שמתקבל שיהיה גם הערך בעל זמן הריצה הטוב ביותר הינו שורש (N/2), ולכן הוספתי את הערך השישי ברשימה שאכן התקבל כערך בעל מספר ההשוואות הקטן ביותר.

טבלת התוצאות בעמוד הבא:

|  |  |  |  |
| --- | --- | --- | --- |
| **ערך K/סוג מערך** | לא ממויין | בסדר עולה | בסדר יורד |
| **מספר פעולות השוואה** | | | |
| 10 | 5403 | 4290 | 4346 |
| 128 | 22306 | 16384 | 16399 |
| 2 | 16510 | 16384 | 16384 |
| 4 | 8824 | 8448 | 8448 |
| 8 | 5713 | 4864 | 4864 |
| 12 | 5267 | 3996 | 4041 |

ולאחר ניתוח נתונים פשוט – (מקרה 1 הינו מערך לא ממוין, מקרה 2 מערך בסדר עולה, ומערך בסדר יורד).

לפני הטבלאות ניתן להסיק את המסקנות הבאות:

1. לפי הנגזרת של זמן הריצה ניתן היה למצוא את המספר אותו הכי "משתלם" מבחינת מספר פעולות השוואה.
2. מערכים ממוינים דורשים מספר פעולות השוואה פחות באופן משמעותי, אך עקב חלוקת המערכים זמן הריצה עדיין נשאר גבוה.
3. בשל החלוקה, ההבדל בין מספר פעולות ההשוואה של המערך הממוין בסדר העולה מול הסדר היורד הוא קטן ולא משמעותי בכלל.
4. כל עוד K מתרחק מהערך הטוב ביותר באלגוריתם זה, מספר פעולות ההשוואה עולה משמעותית ומכפיל את עצמו (ניתן לראות זאת בטבלה השנייה גם כן)

**לסיכום:**

בעקבות הפרויקט למדתי כיצד בחירה של ערך חלוקה יכול להשפיע על יעילות התכנית. כמו כן, כי סוגי הקלטים אינם משנים באופן משמעותי את היעילות ולכן מספיק לנסות לבדוק תוצאות על הקלט הכי "גרוע" מבחינתנו.