

מבוא למדעי המחשב, סמסטר קיץ 2018

תרגיל בית 4

מועד אחרון להגשה: יום ב' – 23:59 10.9.2017

המתרגל האחראי על תרגיל זה: עידו רפאל

E-mail: ungulient@gmail.com (נא להוסיף את מספר הקורס – "234114" לכתובת המייל, כדי שהמייל יגיע לתיקייה הנכונה ולא אפספס אתכם)
שעת קבלה רגילה: אחרי התרגולים בימים ב' ו-ד' (התרגול נגמר ב-15:20. נא לא לאחר, כדי שלא אפספס אתכם).

הנחיות:

- הגשה **בבודדים**. עליכם לכתוב את הפתרונות לבד ולהגיש ביחידים.
- קראו את השאלות בעיון לפני שתתחילו בפתרון.
- הקפידו לתעד את הקוד שלכם בהערות באנגלית.
- מלבד מילואים, לא יתקבלו תרגילים אחרי מועד הגשה. הגשה באיחור לאחר מועד הגשה נחשבת כאי-הגשה.
- כל יום מילואים = יום דחייה. על מנת לקבל את הדחייה, עליכם לשלוח באי-מייל למתרגל האחראי על תרגיל זה עותק של האישור המראה שהייתם במילואים (טופס 3010). אם האישור יגיע אליכם בתאריך מאוחר, יש להודיע על כך למתרגל האחראי על התרגיל.
- לא ניתן לערער על תוצאות הבדיקה האוטומטית.
- שימו לב! הבדיקה הינה אוטומטית, ולכן הקפידו להדפיס בדיוק בפורמט שהתבקשתם ובידקו עם אתר הבדיקה ועם DiffMerge את הפלט שלכם מול הפלט של הדוגמאות שקיבלתם.
 - השתמשו ב-redirection כדי להפנות את הפלט לקובץ טקסט.
 - וודאו את האותיות הגדולות והקטנות לפי הדוגמאות וההסברים בתרגיל.
 - אין להדפיס רווחים שלא התבקשתם להדפיס (בתחילת שורה או בסופה).
- בתרגיל זה מותר להשתמש בפונקציות מהספרייה `stdbool.h`, `stdlib.h`, `stdio.h` למעט במקרים בהם נאמר אחרת. החומר הנדרש לתרגיל זה שייך לתרגולים 0-9 (כלומר עד התרגול של Complexity, כולל). אין להשתמש בחומר שאינו מופיע במצגות אלה.
- ההגשה הינה אלקטרונית ו**בבודדים** דרך אתר הקורס. קובץ ההגשה יהיה מסוג **zip** (ולא אף פורמט אחר) ויכיל בתוכו את הקבצים הבאים בלבד, ללא כל תיקיות:
 - קובץ `students.txt` עם שמך **באנגלית**, מספר תעודת הזהות וכתובת האי-מייל שלך.
 - קובץ פתרון `hw4q1.c` עבור שאלה 1.
 - קובץ פתרון `hw4q2.c` עבור שאלה 2.
 - קובץ פתרון `hw4q3.c` עבור שאלה 3.
- **חובה לשמור את קוד אישור ההגשה שמקבלים מהמערכת לאחר שמגישים, עד לסיום הקורס.**
- יש להקפיד להגיש את כל הקבצים בדיוק עם השמות שמופיעים לעיל. הגשה שלא תעמוד בתנאי זה **לא תתקבל ע"י המערכת!** אם המערכת לא מקבלת את התרגיל שלכם, חפשו את הפתרון לבעיה באתר הקורס תחת הכפתור FAQ.

בדיקה ידנית:

בנוסף לבדיקה האוטומטית, התרגיל ייבדק גם בבדיקה ידנית. הבדיקה תתמקד בנושאים הבאים:

- אורך מימוש (לא כולל ההכרזה, או שם הפונקציה) כל פונקציה לא יעלה על 16 שורות קוד. הגבלה זו תקפה לכל הפונקציות, כולל main.
רק שורות ריקות, שורות עם סוגר מסולסל **בלבד**, ושורות עם הערות **בלבד** לא נספרות. רווחים לבנים לא נספרים.
אסור לכתוב כמה פקודות שונות משמעותית באותה השורה, למשל כותרת תנאי־לולאה צריכה להיות בשורה נפרדת מגוף התנאי־לולאה.
• רוחב שורה - רוחב כל שורה (כולל הערות והזחות) לא יעלה על 80 תווים.
ניתן לסמן אורך של שורה בקודבלוקס:

Code::Blocks -> Settings -> Editor -> Margins and caret ->

Right margin hint: Visible line, **Hint column:** 80

- קבועים בקוד:
 - יש להגדיר בעזרת #define כל קבוע משמעותי. שמות קבועים צריכים להיות באותיות גדולות בלבד, אם השם מכיל יותר ממילה אחת מילים יופרדו בעזרת מקף תחתון (למשל STUDENTS_NUM).
 - אין להשתמש בערכי ASCII ישירות. יש להשתמש בייצוג של התווים (למשל 'a').
 - הזחות:
 - שיטת הזחה מקובלת – הזחת קוד בכל בלוק, למשל:

```
int main()
{
    // your code here
    while (...)
    {
        // your code here
    }
}
```
 - אין להשתמש במשתנים גלובאליים או סטטיים.
 - שמות משתנים/קבועים/פונקציות צריכים להיות אינפורמטיביים, להעיד על מטרתם.
 - בהירות הקוד ותיעוד:
 - יש לתעד את הקוד באמצעות **הערות באנגלית בלבד**. במידה ויש כמה שורות קוד שניתן להסביר בקצרה מה המטרה שלהן, יש לשים הערה בהתחלה ואין צורך לתעד כל שורה.
 - יש לתעד פונקציות – לפני הפונקציה להוסיף הערה שמסבירה בקצרה מה הפונקציה עושה ומה המשמעות של הפרמטרים שלה.
 - התיעוד צריך להיות אינפורמטיבי, כלומר יש להסביר מה המטרה של שורות הקוד ולא לכתוב את הקוד במילים.
 - שכפול קוד שלא לצורך (למשל ריבוי קוד זהה במספר מקרי if-else שונים).
 - אי-עמידה באחת מדרישות התרגיל (שימוש בחומר שהיה אסור בתרגיל וכו').
- באופן כללי – הקפידו על כתיבת קוד מסודר ומוכן ככל שניתן תוך יישום העקרונות שנלמדו בכיתה. מותר לכם לממש פונקציות עזר משלכם ולהשתמש בהן.

חלק א'

ממשו פונקציה עם החתימה:

```
void compute_integral_image(int image[][M], int n, int m, int integral_image[][M])
```

הפונקציה מקבלת תמונת grayscale בגודל $n \times m$ (כמו קודם, n - מספר השורות, m - מספר העמודות) ומחשבת את טבלת השטח המסומם עבורה לתוך `integral_image`.

הניחו שאורך השורה המירבי הינו 50 פיקסלים, והגדירו את הקבוע M בהתאם.

הניחו שהקלט תקין, ז"א ש-`image` ו-`integral_image` מצביעים חוקיים לזיכרון מוקצה בגודל המתאים ו- n ו- m גדלים חוקיים שאינם חורגים מגודל הזיכרון המוקצה.

סיבוכיות: על הפונקציה לעבוד בסיבוכיות זמן $O(n \cdot m)$ (לינארית בגודל התמונה) וסיבוכיות זיכרון $O(1)$.

רמז: אם נוכל לחשב כל איבר במספר קבוע של פעולות ($O(1)$), נוכל לעמוד בדרישות הסיבוכיות (כי יש $n \cdot m$ איברים).
לשם כך, חישבו מה הקשר בין האיבר ה- (i,j) ב-`integral_image` (אותו אנו רוצים לחשב בזמן קבוע) לבין האיברים ה- (i,j) , $(i,j-1)$ ו- $(i-1,j)$ ב-`integral_image`.
כמו כן, שימו לב מה קורה כאשר חורגים מגבולות התמונה, ואז תשכחו לטפל במקרה זה.

חלק ב'

ממשו פונקציה עם החתימה:

```
int sum_rect(int integral_image[][M], int rect[4])
```

הפונקציה תחשב את Σ_{Rect} לפי הנוסחא מעלה. גם כאן הניחו שהקלט תקין, בפרט `rect` מציין חלון מלבני בתמונה בדומה לדוגמא לעיל - $Rect = (i_{top}, j_{left}, i_{bottom}, j_{right})$

סיבוכיות: על הפונקציה לעבוד בסיבוכיות זמן $O(1)$ וסיבוכיות זיכרון נוסף $O(1)$.

חלק ג'

ממשו פונקציה עם החתימה:

```
void sliding_average(  
    int integral_image[][M], int n, int m, int h, int w, int average[][M]  
)
```

הפונקציה מקבלת את טבלת השטח המסומם - `integral_image`, בגודל $n \times m$, ומחשבת לתוך `average` ממוצע על פני חלון נע בגודל $h \times w$ של הערכים בתמונה המקורית, כלומר הערך באיבר ה- (i,j) ב-`average` (כלומר `average[i][j]`) יהיה הערך הממוצע במלבן בתמונה המקורית שמרכזו בפיקסל (i,j) וגודלו $h \times w$.

הגודל של המטריצה `average` כגודל תמונת הקלט.
כמו כן, הניחו שמידות החלון אי-זוגיות (גם h וגם w), וגם שהינן קטנות מגודל התמונה.

עבור הערכים בקצוות (כאשר המלבן "לא נכנס" כולו בתמונה. כלומר "גולש" מחוץ לתמונה), הניחו שמחוץ לתמונה הערכים הם 0 (מה זה אומר על הערכים בטבלת השטח המסומם שמחוץ לתחום?).
למשל עבור מלבן 3×3 והדוגמא לעיל, מתקיים:

$$average[0][1] = (0 + 0 + 0 + 1 + 6 + 8 + 0 + 0 + 3) / 9 = 2$$

הערה: כידוע, בפועל בשפת סי הערכים מחוץ למטריצה שיצרנו אינם בהכרח 0, עליכם לדאוג למקרים אלה במיוחד.

average =

1	2	3	2
2	4	6	4
2	4	6	5
2	4	5	4

טבלת הערכים המלאה:

פורמלית, המטריצה average מחושבת לפי (כאשר image היא התמונה המקורית):

$$average[i][j] = \frac{1}{h \cdot w} \sum_{k=-\lfloor \frac{h}{2} \rfloor}^{\lfloor \frac{h}{2} \rfloor} \sum_{l=-\lfloor \frac{w}{2} \rfloor}^{\lfloor \frac{w}{2} \rfloor} image[i+k][j+l]$$

(כאשר ההנחה שהערכים מחוץ ל-image הם 0).

שימו לב, מכיוון ש-average מכילה מספרים שלמים, את הממוצע יש לחשב בעיגול **למספר השלם הקרוב ביותר** (לשם כך יש להשתמש בפונקציה round בספריית math.h). שימו לב שכאשר אתם מחשבים את הממוצע אתם לא מבצעים חילוק שלמים בטעות.

סיבוכיות: על הפונקציה לעבוד בסיבוכיות זמן $O(n \cdot m)$ וסיבוכיות זיכרון נוסף $O(1)$. כאשר n ו- m הם גודל התמונה שקיבלנו. שימו לב שלא צריכה להיות תלות בגודל החלון הנע! כדי שנוכל לממש פונקציה זו בסיבוכיות הנדרשת, עלינו להיות מסוגלים לחשב כל איבר בזמן קבוע $O(1)$. לשם כך, עשו שימוש בפונקציה sum_rect, ב-integral_image ובנוסחא מההקדמה. שימו לב שיש לטפל במקרים של חריגה מגבולות התמונה.

רמז: אנו רוצים לעשות שימוש ב-integral_image ובנוסחא מההקדמה. חישבו מה קורה כאשר גולשים\חורגים מגבולות התמונה לשמאל ו\או למעלה. מה קורה כאשר גולשים לימין ו\או למטה. שני המקרים שונים.

חלק ד'

כתבו תכנית שקולטת מהמשתמש תמונה ומידות חלון נע, ומשתמשת בפונקציות מהסעיפים הקודמים על מנת לפלוט את טבלת השטח המסומם וממוצע על פני חלון נע (תמונת ממוצעים). הניחו כי גודל התמונה לא יעלה על 50×50 פיקסלים.

כדי להבין מה בדיוק פורמט הפלט המצופה מכם (למשל מתי לשים רווח ומתי לרדת שורה), בדקו את קבצי הקלט\פלט המסופקים לכם.

דוגמא להרצת התכנית:

```
Enter image dimensions:
4 4
Enter image:
1 6 8 3
0 0 3 7
4 7 8 8
5 0 9 9
Enter sliding window dimensions:
3 3
Integral image is:
1 7 15 18
1 7 18 28
5 18 37 55
10 23 51 78
Smoothed image is:
1 2 3 2
2 4 6 4
2 4 6 5
2 4 5 4
```

דוגמאות נוספות בקבצי הקלט\פלט המסופקים לכם.

הנחיות לפתרון התרגיל:

- לאורך התרגיל הניחו שהקלט מהמשתמש חוקי, ז"א כאשר נדרש – המשתמש מזין מימדי תמונה תקינים ($n, m \geq 1$), תמונה בגודל המתאים עם ערכים תקינים עבור הפיקסלים וכן מידות תקינות לחלון נע – ז"א מצבים לא תקינים לא יינתנו כמקרי קלט, ולכן הפלט המדויק במקרים אלו אינו משנה. יחד עם זאת, עליכם להקפיד על כללי תכנות נכון כפי שנלמדו בכיתה, לוודא תקינות קלט (למרות שלא ייבדק בבדיקה האוטומטית), להקפיד על חלוקה הגיונית לפונקציות, להקפיד על קריאות הקוד, שמות משמעותיים למשתנים ופונקציות, הגדרת קבועים וכו'.
- מומלץ לפתור באמצעות עקרונות top-down design. התחילו את הפתרון מסעיף ד'. הניחו שגודל התמונה לא יעלה על 50×50 פיקסלים, והקצו את התמונה וטבלת השטח המסוכם כמטריצות דו-מימדיות בגודל המתאים. הגדירו פונקציות ריקות (או עם מימוש נאיבי) עם החתימות הדרושות, וכן פונקציות נוספות לפי הצורך. חלקו את העבודה לתתי-משימות, ובכל פעם ממשו תת-משימה אחרת, בדיוק כמו שלמדנו.

שאלה 2: מחרוזות

בשאלה זו נרצה לממש מנוע חיפוש מבוסס ASCII – המשתמש יקליד את מספר המחרוזות אותן הוא מעוניין לחפש, לאחר מכן יקליד את המחרוזות והתכנית תפלוט מחרוזות הנמצאות במילון, השקולות בערך ה-ASCII למחרוזת שהוכנסה.

אנחנו נתייחס לערך ה-ASCII הפשוט של מחרוזת. כלומר ערך ה-ASCII של מחרוזת שווה לסכום ערכי ה-ASCII של התווים המרכיבים אותה.

לדוגמא, ערך ה-ASCII של המחרוזת "BRAZIL" הינו:
$$\text{ASCII_Value}(\text{"BRAZIL"}) = 'B' + 'R' + 'A' + 'Z' + 'I' + 'L' = 66 + 82 + 65 + 90 + 73 + 76 = 452$$

נממש את מנוע החיפוש:

1. יצירת מילון מחרוזות המכיל את המחרוזות הבאות:

```
"AVOCADO", "BANNANA", "WATER", "CAT", "DOG", "HELLO",  
"ISRAEL", "ITALY", "GARMISH", "USA", "BRAZIL", "KING", "ABCDy"  
"INT", "CHAR", "DOUBLE", "hardcheese", "realize", "farther"
```

2. נבקש מהמשתמש להקיש מספר שלם ואי שלילי, אשר יסמן כמה מחרוזות הוא רוצה לחפש, ע"י הדפסת ההודעה:

```
"Please enter a non-negative number:\n"
```

- ניתן להניח כי המספר שלם.
- יש לבדוק כי המספר אי שלילי. כל עוד אינו אי שלילי, יש לבקש מהמשתמש להכניס מספר חדש ע"י הדפסת ההודעה:

```
"Please enter a NON-NEGATIVE number:\n"
```

- חישובו על מימוש פשוט ע"י לולאת do-while, והדפסה בעזרת printf והדגל %s.

3. נבקש מהמשתמש להכניס את מספר המחרוזות אותו הוא ביקש לחפש ע"י הדפסת ההודעה:

```
"Please enter %d strings:\n"
```

- אין לטפל במיוחד במקרה של מחרוזת אחת (כלומר עדיין יש להדפיס "1 strings").
- יש לטפל במיוחד במקרה של 0 מחרוזות, ולהדפיס את ההודעה:

```
"no strings\n"
```

4. המשתמש יכניס את המחרוזות (ניתן להניח שמופרדות בירידת שורה וכן כי לא מכילות רווחים\whitespaces. כלומר ניתן לקלוט ע"י scanf ו-s% ולא לדאוג שתתעלם מרווחים). לכל מחרוזת (ניתן להניח שבגודל 10 לכל היותר):

- במידה ונמצאו התאמות (כל מחרוזת עם ערך ASCII שווה לערך ה-ASCII של המחרוזת שלנו):

נדפיס:

```
"the string %s has %d matches in the dictionary:\n"
```

כאשר %s היא המחרוזת, ו-d% הוא מספר ההתאמות שמצאנו במילון. ומתחת להודעה זו, נדפיס את כל המחרוזות מהמילון שערך ה-ASCII שלהן שווה לערך ה-ASCII של המחרוזת שלנו. אם יש יותר מאחת, סדר ההדפסה הוא לפי סדר הופעת המחרוזות במילון.

כמו כן, יש להפריד בין המחרוזות שנמצאו מתאימות ע"י ירידת שורה אחת בין מחרוזת למחרוזת (וכמובן גם ירידת שורה אחרי המחרוזת האחרונה, ולא רק ביניהן).

• במידה ולא נמצאו התאמות:

נדפיס:

```
"the string %s has no matches in the dictionary\n"
```

כאשר %s היא המחרוזת שלנו.

• לאחר כל איטרציה של הלולאה הראשית שמוגדרת בסעיף 4, יש להדפיס בנוסף 2 ירידות שורה.

דוגמאות הרצה:

```
Please enter a non-negative number:
1
Please enter 1 strings:
example
the string example has 2 matches in the dictionary:
realize
farther
```

```
Please enter a non-negative number:
0
no strings
```

```
Please enter a non-negative number:
-6
Please enter a NON-NEGATIVE number:
3
Please enter 3 strings:
WATER
the string WATER has 3 matches in the dictionary:
WATER
ITALY
ABCDy

ITALY
the string ITALY has 3 matches in the dictionary:
WATER
ITALY
ABCDy

hsercdhee
the string hsercdhee has no matches in the dictionary
```

לדוגמאות נוספות, בדקו את קבצי הקלט\פלט המסופקים לכם.

שאלה 3: ניתוח סיבוכיות

כל הסעיפים דורשים סיבוכיות זמן במקרה הגרוע.

בשאלות 0-5 (בעמוד הבא) בחרו את האות המייצגת את הסיבוכיות המתאימה מתוך טבלת הפונקציות להלן:

- | | |
|--------------------------------|---|
| a. $\Theta(1)$ | n. $\Theta(n^2 \cdot \sqrt{n})$ |
| b. $\Theta(\log n)$ | o. $\Theta(n^3)$ |
| c. $\Theta(\log^2 n)$ | p. $\Theta(n^3 \log n)$ |
| d. $\Theta(\sqrt{n})$ | q. $\Theta(n^3 \log^2 n)$ |
| e. $\Theta(\sqrt{n} \log n)$ | r. $\Theta(2^n)$ |
| f. $\Theta(\sqrt{n} \log^2 n)$ | s. $\Theta(n \cdot 2^n)$ |
| g. $\Theta(n)$ | t. $\Theta(n^2 \cdot 2^n)$ |
| h. $\Theta(n \log n)$ | u. $\Theta(n^3 \cdot 2^n)$ |
| i. $\Theta(n \log^2 n)$ | v. $\Theta(3^n)$ |
| j. $\Theta(n \cdot \sqrt{n})$ | w. $\Theta(n \cdot 3^n)$ |
| k. $\Theta(n^2)$ | x. $\Theta(n^2 \cdot 3^n)$ |
| l. $\Theta(n^2 \log n)$ | y. $\Theta(n^3 \cdot 3^n)$ |
| m. $\Theta(n^2 \log^2 n)$ | z. <i>the correct answer doesn't appear</i> |

השאלות:

```
void f1(int n) {
    int s = 1;
    for (int i = 0; i < n; i++) {
        s *= 2;
        for (int j = 0; j < s * n; j++) {
            printf("hi!\n");
        }
    }
}
```

0. מה היא סיבוכיות הזמן של הפונקציה f1 כתלות ב-n?
1. מה היא סיבוכיות המקום של הפונקציה f1 כתלות ב-n?

```
#define K 5

int f2(int n) {
    int s = 0;
    int d, m;

    while (n > 1) {
        scanf("%d", &m);
        d = (m > 2 * K) ? 2 * K : (m < K ? K : m);
        n = n / d;
        s += n;
    }
    return s;
}
```

2. מה היא סיבוכיות הזמן של הפונקציה f2 כתלות ב-n?
3. מה היא סיבוכיות המקום של הפונקציה f2 כתלות ב-n?

```
int f3(int k) {
    int s = k;
    for (int m = 2; k > 0; k /= m) {
        s = s + k;
    }
    return s + 1;
}

void f4(int n) {
    for (int j = 0; j < n * f3(n); ++j) {
        for (int i = 1; i * i < n; i *= 2) {
            printf("Hello j!\n");
        }
    }
}
```

4. מה היא סיבוכיות הזמן של הפונקציה f4 כתלות ב-n?
5. מה היא סיבוכיות המקום של הפונקציה f4 כתלות ב-n?

אופן הגשת התשובות בשאלה 3:

מסופק לכם קובץ שלד hw4q3.c אשר נראה כך:

```
int main()
{
    int question;
    char answer[] = {
        'a', 'b', 'c', 'd', 'e', 'f'
    };

    scanf("%d",&question);

    printf(
        "The answer to question %d is %c\n",
        question, answer[question]
    );

    return 0;
}
```

עליכם לערוך את התשובות במערך answer בהתאם.

למשל, בקובץ השלד, התשובות לשאלות 0-5 הן a-f בהתאמה, כלומר לשאלה 0 התשובה היא a, לשאלה 1 התשובה היא b, לשאלה 2 התשובה היא c וכו'. עליכם לשנות את התשובות במערך לתשובות הנכונות.

קלט/פלט צפויים:

כפי שניתן לראות מקובץ השלד, הקלט לתרגיל זה יהיה מספר השאלה (0 עד 5).

סיפקנו לכם קבצי **קלט** של שאלות 0-5 (שבסה"כ מכילים את מספר השאלה וירידת שורה), אך כמובן שלא סיפקנו קבצי **פלט** לדוגמא (ולכן גם לא באתר), מכיוון שאתם בסה"כ מתבקשים להחליף את הערכים שיש במערך, ואחרת הייתם יכולים לדעת את התשובות בלי לפתור את השאלות 😊.