

ב"ה

תרגיל מס' 3 – תהליכים וקבצים

הוראות הגשה

- מועד אחרון להגשה: 23:59 12/05/21.
- יש לשלוח את הקבצים באמצעות האתר:
<https://submit.cs.biu.ac.il/cgi-bin/welcome.cgi>
לפני חלוף התאריך הנקוב לעיל.
- שם ההגשה של התרגיל: ex3, יש להגיש רק את הקבצים ex31.c ו-ex32.c. את שני הקבצים יש להגיש יחד (ולא בהגשות נפרדות).
- להזכירכם, העבודה היא אישית. "עבודה משותפת" דינה כהעתקה.
- יש לוודא שהתרגיל מתקמפל ורץ על ה-planet ללא שגיאות/אזהרות.
- שימו לב להערות בסוף התרגיל.
- בראשי שני קבצי ה-c אותם אתם מגישים, יש לכתוב שם מלא ות.ז. בפורמט הבא:
// Israel Israeli 123456789
- מענה לשאלות בנוגע לתרגיל יינתן עד 12.5.2021 בלבד, בפורום הייעודי במודל.
- לפני פרסום שאלה יש לוודא שהשאלה לא נשאלה בעבר (ניתן להיעזר באופציית החיפוש במודל). שאלות חוזרות לא יענו.

בהצלחה!

השוואת קבצים – חלק א'

הנחיות עבור ex31

- שם קובץ שיש לשלוח: ex31.c

כתבו תכנית המקבלת נתיבים לשני קבצים כארגומנטים ל-main, ובודקת האם שני הקבצים זהים/דומים/שונים. למען הסר ספק, הנתיבים כוללים את שמות הקבצים. אם הקבצים זהים (מכילים בדיוק אותו תוכן), התוכנית **תחזיר 1**, אם הקבצים דומים (הסבר בהמשך) התוכנית **תחזיר 3**, אחרת, אם הקבצים שונים, **תחזיר 2** (שימו לב, לא מדובר כאן בתוכנית שמדפיסה למסך).

קבצים זהים הם קבצים שכל התווים בהם שווים (לדוגמא: Hello World ו-Hello World). קבצים דומים הם קבצים שאינם זהים, אך מכילים את אותו טקסט, וקיים הבדל בשימוש באותיות קטנות או גדולות, רווח ו/או ירידות שורה. קבצים שונים הם קבצים שאינם זהים וגם אינם דומים.

לדוגמא – קבצים אלו הינם דומים:

File A:

12ab23

File B:

12Ab23

File C:

12aB23

File D:

12AB23

File E:

12 aB 23

File F:

12

ab2

3

ראו דוגמאות נוספות בקבצים המצורפים לתרגיל.

שימו לב:

- הקבצים עשויים להכיל אותיות (גדולות או קטנות), מספרים, רווחים, או ירידות שורה.
- ניתן להשתמש ב-strcmp (למרות שבהערות שבסוף התרגיל צוין שאין להשתמש בפונקציות ספריה שניתן לממש באמצעות הכלים שלמדנו בתרגול).
- כאשר אתם מקמפלים את התוכנית תנו לה את השם comp.out במקום a.out.

דוגמא להרצת התוכנית:

```
[os2021@localhost ~]$ ./comp.out /home/os2021/code/1.txt /home/os2021/code/2.txt
```

לאחר הרצת התוכנית, אם הקבצים זהים ונרשום את הפקודה `$? echo` נקבל את הערך 1, אם הם רק דומים נקבל את הערך 3, אחרת נקבל את הערך 2.

כלומר, בהנחה שקבצים 1.txt ו 2.txt דומים:

```
[os2021@localhost ~]$ ./comp.out /home/os2021/code/1.txt /home/os2021/code/2.txt
```

```
[os2021@localhost ~]$ echo $?
```

```
3
```

```
[os2021@localhost ~]$
```

דוגמא נוספת:

```
[os2021@localhost ~]$ ./comp.out /home/os2021/code/1.txt /home/os2021/code/1.txt
```

```
[os2021@localhost ~]$ echo $?
```

```
1
```

```
[os2021@localhost ~]$
```

השוואת קבצים – חלק ב'

הנחיות עבור ex32

- שם קובץ שיש לשלוח: ex32.c

כתבו תכנית המקבלת נתיב לקובץ קונפיגורציה כארגומנט ל-main. ניתן להניח שהקובץ קיים (ואינו תיקיה).

קובץ הקונפיגורציה מכיל 3 שורות:

שורה 1: נתיב לתיקייה המכילה תתי תיקיות. ברמה הראשונה (אחת פנימה), כל תיקייה מייצגת משתמש, ואמורה להכיל קובץ c.

שורה 2: נתיב לקובץ המכיל קלט.

שורה 3: נתיב לקובץ המכיל את הפלט הנכון עבור קובץ הקלט משורה 2. קובץ הקונפיגורציה יסתיים בתו ירידת שורה.

על התוכנית שלכם להיכנס לכל תתי התיקיות (ולהתעלם מקבצים אחרים שאינם תיקיות, במידה וקיימים) שבתוך התיקייה משורה 1, לחפש בכל אחד מתתי התיקיות שלה (ולא ברמות עמוקות יותר) קובץ c, ולקמפל אותו. את קובץ הריצה שנוצר יש להריץ עם הקלט שמופיע בקובץ שבמיקום שבשורה 2 (התוכנית שתריצו קולטת מ-stdin ומדפיסה ל-stdout ולכן עליכם להשתמש ב- i/o redirection).

הניחו שיהיה לכל היותר קובץ c אחד בתיקייה (יכול להיות שלא יהיה קובץ c בכלל, וגם יכולים להיות קבצים ותיקיות מסוגים אחרים, למעט קבצים עם הסיומת out). ניתן להניח שלקובץ ה-c תהיה בהכרח סיומת מתאימה, לדוגמא file.c. אין משמעות לשם של קובץ ה-c.

את הפלט של התוכנית יש להשוות עם קובץ הפלט הרצוי, שמיקומו מגיע משורה 3, בעזרת התוכנית comp.out שמימשתם בחלק א' של התרגיל (הריצו את התוכנית comp.out ותתנו לה כארגומנטים ל-main את המיקום של הפלט הנכון, שנמצא בקובץ הקונפיגורציה בשורה 3, ואת הפלט של התוכנית של המשתמש אותה הרצתם).

התוכנית שלכם צריכה ליצור קובץ (בתיקייה שממנה הורצה התוכנית שלכם) בשם results.csv שמכיל עבור כל שם משתמש (שם תתי תיקייה), את ציונו (בין 0 ל-100) בהתאם לתשובה ש-comp.out החזירה, ואת הסיבה. יש לרשום את התו " ", בין שם המשתמש לבין הציון, לבין הסיבה שלו (ללא רווחים).

סיבות אפשריות:

1. NO_C_FILE - אין בתיקייה של המשתמש קובץ עם סיומת c. **הציון שינתן, 0.**
2. COMPILATION_ERROR - שגיאת קומפילציה (קובץ לא מתקמפל). **הציון שינתן, 10.**
3. TIMEOUT - קובץ ה-c המקומפל רץ יותר מ-5 שניות. **הציון שינתן, 20.**
4. WRONG - הפלט **שונה** מהפלט הרצוי. **הציון שינתן, 50.**
5. SIMILAR - הפלט **שונה** מהפלט הרצוי **אך דומה**. **הציון שינתן, 75.**
6. EXCELLENT - הפלט **זהה** לפלט הרצוי. **הציון שינתן, 100.**

דוגמא לתוכן קובץ results.csv:

```
Monica,100,EXCELLENT
Phoebe,0,NO_C_FILE
Rachel,20,TIMEOUT
Ross,10,COMPILATION_ERROR
Joey,50,WRONG
Chandler,75,SIMILAR
```

דוגמא לתוכן קובץ הקונפיגורציה:

```
/home/os2021/students
/home/os2021/io/input.txt
/home/os2021/io/output.txt
```

דוגמא לתוכן קובץ קלט (שורה 2):

```
1
5 4
4
```

דוגמא לתוכן קובץ פלט (שורה 3):

```
Please enter an operation
Please enter two numbers
The sum is 9
Please enter an operation
Bye
```

דוגמא להפעלת קובץ הריצה של התוכנית שלכם:

```
[os2021@localhost ~]$ ./a.out /home/os2021/conf.txt
```

שימו לב:

- לא ניתן להניח שתיקיית המשתמשים תכיל רק תיקיות. בנוסף, לא ניתן להניח שהשורה הראשונה בקובץ הקונפיגורציה בהכרח תייצג תיקייה, או שהיא תייצג נתיב שאכן קיים במערכת הקבצים. במקרים אלו יש להדפיס את ההודעה "Not a valid directory" ולאחריה תו ירידת שורה. לאחר מכן לצאת מהתוכנית ע"י החזרת 1-.
- בהקשר לשורות 2 ו-3 של קובץ הקונפיגורציה, לא ניתן להניח שהנתיבים אכן קיימים במערכת הקבצים. במידה והם לא קיימים, יש להדפיס "Input file not exist" או "Output file not exist" (בהתאם למקרה) ולאחר מכן תו ירידת שורה. לאחר מכן יש לצאת מהתוכנית ע"י החזרת 1-.
- הקובץ results.csv צריך להסתיים בירידת שורה.
- ניתן להניח שבזמן הרצת ex32 לא יהיה קיים בתיקייה קובץ results.csv (ישן).
- כל סיבה לציון עומדת בפני עצמה. כלומר, אין מצב שידרוש שרשור של סיבה אחת עם אחרת.

לנוחיותכם, מצ"ב לתרגיל זה התיקיה files המכילה דוגמאות לבדיקה. תיקייה זו מכילה:

1. תיקייה textComparison המכילה זוגות של קבצי טקסט לבדיקה של ex31.c
2. תיקייה students המכילה תיקיות של שמות משתמשים
3. קובץ קונפיגורציה - conf.txt
4. תיקייה בשם io ובה קובץ input וקובץ output
5. קובץ results.csv
6. קובץ Makefile בשביל לקמפל את קבצי ה-c שלכם (ex31.c, ex32.c)

הערות והנחיות נוספות:

1. אין להדפיס למסך מעבר למה שהוגדר בתרגיל.
בפרט, ב-ex32 אין להדפיס למסך שגיאות שמקורן בתוכניות אותן אתם מריצים. כלומר עליכם לכתב את file descriptor מספר 2 של הקוד שאתם מקמפלים, לקובץ errors.txt (ניתן להניח כי אינו קיים כאשר מריצים את ex32).
2. עבור כל system call או פונקציית ספרייה, חובה לבדוק האם היא הצליחה או לא.
אם היא לא הצליחה יש להדפיס למסך את ההודעה:

Error in: X

- כאשר X הוא שם הפונקציה שנכשלה.
- אין צורך לבדוק הצלחה עבור פונקציות string, לדוגמא strcpy, strcat, strcmp וכו'. כמו כן, נחריג את הפונקציה write, גם עבורה יש לבדוק כישלון, אך אם היא נכשלה אין להדפיס הודעת שגיאה או לנסות לקרוא לה שוב.
3. בנוסף להדפסת השגיאה:
 - ב-ex31 יש לצאת מהתוכנית לאחר מכן ולהחזיר 1-.
 - ב-ex32:
אם השגיאה קרתה בשלב כלשהו בעת, או לאחר כניסה לתיקיית משתמש, יש להמשיך בתוכנית לטיפול במשתמש הבא (אם קיים).
עבור אותו משתמש, אין צורך להדפיס דבר לקובץ results.csv.
במידה ונתקלתם בשגיאה חמורה שלא מאפשרת את המשך התוכנית – ניתן גם לצאת ולא להמשיך למשתמש הבא. אנחנו משאירים זאת לשיקול דעתכם בהתאם למקרה, אך חשוב להזכיר שאין צפי לשגיאות כאלו בשימוש נכון בפונקציה, לכן אם אתם נתקלים בשגיאות יש לבדוק היטב את מקורן.
 - אחרת (שגיאה שלא בהקשר של משתמש ספציפי), יש לצאת מהתוכנית ולהחזיר 1-.
ככלל, יציאה מהתוכנית צריכה להתבצע בצורה נקייה.

4. הניחו שה-gcc מוגדר במשתנה הסביבה PATH, ולכן אין בעיה להשתמש ב-execvp.
5. אתם רשאים להשתמש בכל קריאות המערכת שנלמדו בתרגולים עד היום. במידה וניתן לבצע פעולה באמצעות קריאות המערכת שלמדנו, אסור להשתמש בפונקציות אלטרנטיביות אחרות, כמו לדוגמא fread במקום read. בפרט, אין להשתמש ב-printf, אלא ב-write.
6. בנוסף, מותר להשתמש בפונקציות הבאות:
opendir, getcwd, chdir, memset, strtok, remove, unlink
פונקציות לעבודה עם strings: strcmp, strcpy, strcat, strlen.
7. אין להשתמש בפונקציה sleep או כל וריאציה אחרת שלה.
8. יש לדאוג לסנכרון בין תהליכים כפי שנלמד בתרגול (wait, waitpid).
9. הניחו שהתכנית comp.out תהיה בתקיה הנוכחית (./). שממנה יריץ הבודק את קובץ הריצה של חלק ב' של התרגיל.
10. כל שורה בקובץ הקונפיגורציה לא תעלה על 150 תווים. ניתן גם להניח שאורך הנתוב המקסימלי המוביל לקובץ C גם לא יעלה על 150 תווים.
11. אין חשיבות לסדר של הסטודנטים הרשומים בקובץ results.csv.
12. בסיום ההרצה, יש למחוק קבצים זמניים שיצרתם, למעט הקבצים הבאים, אותם יש להשאיר:
 - ex31.c
 - ex32.c
 - comp.out (שנוצר מקימפול ex31.c)
 - a.out (שנוצר מקימפול ex32.c)
 - results.txt
 - errors.txt
13. אם תכנית של סטודנט לא מסיימת את עבודתה תוך 5 שניות אתם לא צריכים לדאוג לסיים אותה, אך אתם צריכים להתייחס לזה בקובץ results.csv ע"י ציון 20 וסיבה תואמת.
14. בכל תיקיה בלינוקס יש " " " " " " (קישור לתיקיה הנוכחית ולתיקיית האב), שימו לב לכך כאשר אתם משתמשים ב-readdir.

15. יש לסגור file descriptors שנפתחו על ידכם.
16. שימוש במשתנים גלובליים הינו אסור.
17. אין צורך להשתמש בהקצאות דינאמיות בתרגיל.
18. בכל מקום בו מסופק נתיב – הנתיב יכול להיות יחסי או אבסולוטי.



בהצלחה!!!