

תרגיל 6, מבוא לתכנות מערכות, חורף 2020-2021

הגשה בזוגות או ביחידים דרך המודל
התרגיל הוא להגשה עד ליום ראשון, 24/1/2020 בשעה 23:59

תיאור התרגיל

בתרגיל זה תממשו שני סקריפטים ב-bash.

סקריפט 1: ספירת מספר המילים שמתחילות בכל אות בקובץ טקסט

כתבו סקריפט בשם count-words.sh אשר מקבל קובץ טקסט ומדפיס למסך את מספר המופעים של מילים שמתחילות בכל אחת מאותיות הא"ב האנגלי (ללא הבדל בין lowercase ל-uppercase). למשל: עבור הקובץ I_have_a_dream.txt המצורף במודל יוציא הסקריפט את הפלט הבא:

```
itaish@cs ~/Matam/Assignments/Ex6 $ ./count-words.sh I_have_a_dream.txt
a      160
b       88
c       72
d       56
e       23
f       76
g       37
h       72
i       86
j       17
k        2
l       58
m       56
n       68
o      152
p       34
q        2
r       46
s      100
t      268
u       20
v       12
w      122
x        0
y       17
z        0
```

הוציאו הודעת usage במקרה שלא התקבל ארגומנט אחד בדיוק. ודאו שהארגומנט שהתקבל הוא קובץ.

הצעת פתרון:

- ודאו שהתקבל ארגומנט אחד, הדפיסו הודעת שגיאה וצאו אם לא.
- ודאו שמה שהתקבל הוא שם חוקי של קובץ (העזרו ב-f-), הדפיסו הודעת שגיאה וצאו אם לא.
- השתמשו בפקודה tr בשביל להפריד כל מילה בקובץ הקלט לשורה.
- עברו על כל אותיות הא"ב האנגלי בלולאת for. ניתן למצוא פרטים על כך כאן:

<https://stackoverflow.com/questions/7300070/looping-through-alphabets-in-bash>

ה. השתמשו ב-egrep וב-regular expression המתאים על מנת למצוא את כל המילים שמתחילות באות אותה אתם בודקים. השתמשו בדגל i- על מנת ש-egrep יתעלם מ-lower/uppercase. ספרו את מספר השורות ש-egrep החזיר.

ו. הדפיסו את האות והמספר שמצאתם, הפרידו אותם ב-tab. השתמשו בדגל e- כדי לגרום ל-echo להתייחס ל-t\כאל tab.

סקריפט 2: ספירת מספר המילים שמתחילות בכל אות בקובץ טקסט

כתבו סקריפט בשם list-files.sh אשר מקבל שני ארגומנטים: הראשון הוא שם של תיקייה והשני הוא תיאור של תבנית (glob) שם של קבצים/תיקיות אותם הוא צריך לחפש. למשל: "c.*" יהיה תיאור של כל הקבצים/תיקיות שמסתיימים ב-c.. על הסקריפט לעבור על כל התיקיות שנמצאות תחת התיקייה שהועברה כארגומנט (כולל התיקייה עצמה) ולהדפיס את מספר הקבצים/תיקיות שמתאימים לתבנית שניתנה. למשל:

```
itaish@cs ~/Matam/Assignments/Ex6 $ ./list-files.sh /tmp/numbers "*.txt"
/tmp/numbers      13
/tmp/numbers/dir.10687: 1
/tmp/numbers/dir.10737: 1
/tmp/numbers/dir.21888: 1
/tmp/numbers/dir.22820: 1
/tmp/numbers/dir.28086: 1
/tmp/numbers/dir.29221: 1
/tmp/numbers/dir.30683: 1
/tmp/numbers/dir.5562: 1
```

הסקריפט צריך להדפיס הודעת usage במקרה ולא התקבלו שני הפרמטרים:

```
itaish@cs ~/Matam/Assignments/Ex6 $ ./list-files.sh
```

```
Usage: ./list-files.sh <root-directory> <pattern>
```

על מנת למצוא את כל הקבצים/תיקיות שמתאימים לתבנית מסויימת ניתן להשתמש בפקודה find. למשל:

```
itaish@cs ~/Matam/Assignments/Ex6 $ find /tmp/numbers -name "*.h"
/tmp/numbers/dir.10687/file.5827.h
/tmp/numbers/file.3050.h
/tmp/numbers/file.22806.h
```

על מנת לבדוק האם אלמנט מסויים הוא תיקייה או לא ניתן להשתמש ב-d-, בדומה ל-f- עבור קבצים. השתמשו בלולאה על מנת לעבור על כל מה שנמצא בתיקייה המבוקשת.

דגשים

- הקפידו על קוד קריא ושמות משתנים ברורים. הכניסו indentation. אל תשכחו להוסיף את השורה #!/bin/bash בתחילת הסקריפט.

הגשה

- הגשה אלקטרונית: יש להגיש במערכת Moodle את שני הסקריפטים מאוגדים בקובץ ex6.tar. כתבו את השמות+מספרי ת.ז. של המגישים בהערה בתחילת כל סקריפט.

בהצלחה!