

תרגיל בית מס' 1 – Bash

תאריך הגשה: 23.11.2022 בשעה 23:59

הקדמה

בתרגיל זה נבנה script מסוג bash שמטרתו לעזור לנו לאתר ולבחון את המילים הכתובות במסמכים שונים.

נעשה זאת כל ידי יצירת רשימה של כל המילים המופיעות במסמך או קבצים מסויימים המתחילות באות מסויימת, מכילות כמות מסויימת של תווים וספירת ההופעות של כל מילה.

קריאה לסקריפט

התרגיל מורכב מקובץ אחד לפחות, בשם wordFinder.sh, המקבל רשימת פרמטרים. כל פרמטר ברשימה הוא path לקובץ שאנו רוצים להוסיף לרשימת המילים שלנו. הpath יכול להיות או רלטיבי או אבסולוטי. יש לשים לב לשני המקרים! . שני הפרמטרים האחרונים יספקו לנו את המידע שיש לחפש לפיו. בפרמטר האחד לפני אחרון יש לתת תו שהינו התו הראשון של המילים אותם נרצה למצוא. בפרמטר האחרון יש לתת מספר חיובי המסמל את מספר התווים המינימלי של המילים שיש לחפש בטקסט. דוגמא לקריאה לסקריפט:

```
> ./wordFinder.sh ~/myFirstFile.txt ~/mySecondFile.txt a 2
```

בדיקת קלט

שימו לב – יש לקבל בסקריפט לפחות שלושה פרמטרים על מנת שהסקריפט יעבוד והם – (1) נתיב לקובץ אחד תקין, (2) תו אחד תקין, (3) מספר אחד תקין. במידה והתקבלו פחות מ-3 פרמטרים (לא משנה מה הם), הסקריפט ידפיס stderr את הודעת השגיאה הבאה :

Number of parameters received : [number]

רק במידה ויש 3 פרמטרים לפחות יבדק האם הפרמטר האחד לפני אחרון הוא תו אחד בלבד תקין. תו אחד בלבד תקין זה אחד מהתווים שהסקריפט יודע לחשב, ולכן כל תו אחד מתוך 0-9, a-z, A-Z נחשב לתו תקין וחוקי. כל דבר אחר נחשב ללא תקין ויתקבל עבורו הודעת שגיאה.

לאחר מכן יבדק האם הפרמטר האחרון הוא מספר חיובי שלם תקין הגדול מאפס.

- במידה והפרמטר אחד לפני אחרון אינו תו אחד בלבד תקין, הסקריפט ידפיס stderr את הודעת השגיאה הבאה:

Only one char needed : [parameter received]

- במידה והפרמטר האחרון הוא אינו מספר חיובי שלם תקין הגדול מאפס, הסקריפט ידפיס stderr את הודעת השגיאה הבאה :

Not a positive number : [parameter received]

- אם גם התו אינו תקין וגם המספר אינו תקין יודפסו שתי השגיאות כך שהשגיאה של התו ראשונה והשגיאה של המספר שניה.

רק במידה ויש 3 פרמטרים והתו והמספר הינם תקינים, יבדק כי כל רשימת הקבצים תקינה. אם אחד מהפרמטרים ברשימה אינו קובץ תקין (אלא תיקייה או לא קיים כלל), הסקריפט ידפיס stderr את הודעת השגיאה הבאה על כל קובץ לא תקין:

File does not exist : [file_name received]

יש להדפיס שורה כזו עבור כל הקבצים שלא תקינים במידה ויש יותר מאחד (עם ירידת שורה ביניהם), על פי הסדר בהם מסודרים כפרמטרים. שימו לב כי לכל ההדפסות **מעלה** יש להכניס את הפרמטר שהתקבל להדפסת השגיאה. כל מה שבתוך ה [] כולל הסוגריים עצמם זה place holder עבור המידע שמתקבל בזמן אמיתי בתוכנית. ראו דוגמאות הרצה בהמשך העבודה להבהרת ההדפסות.

במידה וכתבתם הודעת שגיאה כלשהי למשתמש, יודפס למסך (**stdout**) גם הודעת העזרה הבאה עבור המשתמש (כאן אין ממלאי מקום, יש להדפיס בדיוק כך) :
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
מטרת ההודעה היא להודיע למשתמש מהי הצורה התקינה שיש להשתמש בסקריפט זה, בדומה ל-synopsis של פקודות ב-man pages.
שימו לב שהיא תודפס **לstdout** ורק פעם אחת, תמיד בסוף כל הודעות השגיאה שהודפסו **stderr**.

במידה ויש אפילו שגיאה אחת (גם אם רק קובץ אחד לא תקין מתוך רשימה תקינה) הסקריפט יסיים את עבודתו ולא ימשיך לפעולת החיפוש.

אופי פעולת הסקריפט

הסקריפט יודע לקרוא את כל המילים הבנויות מהאותיות Z-A, a-z, 0-9. כל תו אחר מהווה רווח בין מילים.
כדי שלא לספור את המילה Cat ו-cat כמילים שונות, על הסקריפט להתייחס לכל המילים כעל lowercase, מבלי לשנות את הקובץ המקורי.
על הסקריפט יש לספור רק את המילים שמתחילות בתו שהתקבל והן בעלות מספר תווים השווה או גדול מהמספר שהתקבל בסקריפט.
לאחר סיום הספירה, על התוצאות להיות מודפסות (**stdout**) לפי **כמות הופעות המילים** ולאחר מכן לפי **סדר לקסיקוגרפי**. על כל המילים המודפסות להיות באותיות קטנות בלי קשר לכיצד מופיעות במסמך וללא שינוי המסמך.

לדוגמא, המשפט "Isn't that that Wow!" עם האות t והמספר 1 תוצג הרשימה הבאה:
1 t
2 that

פקודות מומלצות:

אתם רשאים לממש את העבודה בעזרת כל אוסף של פקודות bash שתמצאו לנכון. אין להשתמש בפקודות הדורשות התקנה מוקדמת.
אסור לסקריפט לפנות לאינטרנט או להתקין תוכנות חדשות בסביבה בה מריצים אותו.

שימו לב:

- הכתובת "~" מתפרשת בקבלת הפרמטר לנתיב המלא ולכן יש להדפיס במידת הצורך את הערך שאתם מקבלים בסקריפט כפי שהוא. אין צורך לבצע מניפולציות או שינויים על הסטרינג (כלומר החל מHOME).
 - זכרו – מבחינת bash, אותיות קטנות אינן זהות לאותיות גדולות (a != A). שימרו על אחידות בין העבודה שלכם למסמך זה, ועל אחידות בתוך העבודה עצמה.
 - מותר לכם ליצור קבצים נוספים מתוך הסקריפט (כל עוד אתם מנקים אחריכם).
 - יש תווים שמתקבלים כאופרטורים. אנחנו לא(!) ננסה להכשיל לכם את העבודה עם תווים שהם מתקבלים כאופרטורים בפרמטרים של הסקריפט.
- בבדיקות יתקבלו רק תווים אמיתיים ולא אופרטורים (לדוגמא, התו # מתקבל כאופרטור אבל כדי להכניס ממש את התו # יש צורך לפני בסלש אחורי, כלומר #\ יכניס את התו עצמו ולא את האופרטור)

תרחיש לדוגמא:

בתרחיש הדוגמא הנ"ל נרצה לאתר מילים משתי הפואמות המפורסמות של המשוררת הידועה פ. בופה : smelly cat, sticky shoes .
(הקבצים מצורפים לעבודה זו על מנת שתוכלו להתנסות בהם בעצמכם)

הקריאות הראשונות מייצגות דוגמאות לקריאות עם קלט לא חוקי, והקריאות האחרונות מייצגות דוגמאות לקריאות חוקיות עבור שני הקבצים smelly_cat.txt ו sticky_shoes.txt המכילות את השירים.

שימו לב להדפסות, לשגיאות ולהתעלמות מתווים כגון " , ! .

תרחישי שגיאה אפשריים:

```
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh
Number of parameters received : 0
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist aa 1
Only one char needed : aa
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist aa bb
Only one char needed : aa
Not a positive number : bb
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist a b
Not a positive number : b
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist a -1
Not a positive number : -1
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist a 1
File does not exist : notExist
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh notExist a
Number of parameters received : 2
Usage : wordFinder.sh <valid file name> [More Files] ... <char> <length>
```

תרחישי הצלחה אפשריים:

```
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt s 2
1 smell
14 smelly
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt s 6
14 smelly
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt n 3
1 noses
7 not
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt n 1
1 no
1 noses
7 not
```

```

inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt
sticky_shoes.txt s 1
1 say
1 smell
1 smile
1 so
1 something
1 stepped
4 sticky
5 shoes
6 s
14 smelly
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt
sticky_shoes.txt s 3
1 say
1 smell
1 smile
1 something
1 stepped
4 sticky
5 shoes
14 smelly
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt
sticky_shoes.txt s 8
1 something
inbal@latitude:~/Desktop/assignment1$ ./wordFinder.sh smelly_cat.txt
sticky_shoes.txt s 30

```

* הקריאה האחרונה היא חוקית. אך כיוון שאין אף מילה שעומדת בתנאים לא יודפס **כלום**
למסך באף אחד מהזרמים.
* שימו לב לאיזה זרם יש להדפיס את התוכן (err,out) בדוגמא בקובץ לא היה אפשרות
להראות את השוני.
* פלט ההדפסה מצורף גם בתור קובץ טקסט על מנת שיהיה קל לבצע השוואות.

להעשרה והעלאת המורל בזמן העבודה:

<https://www.youtube.com/watch?v=XXNIZulBJKs>

<https://www.youtube.com/watch?v=rjDWOQHJtBM>

הנחיות הגשה

ההגשה ביחידים בלבד. עבודות שיוגשו בצורה אחרת – לא ייבדקו.
נא לעבור על קובץ הצ'ק ליסט לפני ההגשה ולוודא כי מילאתם את כל הנקודות!!
קראו את מסמך הוראות ההגשה למידע נוסף לגבי הגשת העבודה.

הגישו בדיוק קובץ אחד בשם wordFinder.sh אותו יש להריץ.

התרגיל יעבור בדיקה אוטומטית. כדי שהתוכנית תעבור בשלום את הבדיקה אין לשנות את הפורמט של ההדפסות שניתנו – הן מבחינת כמות enters, upper/lower case רווחים ובאיזה זרם ההודעה.

את העבודות יש להגיש לאתר הקורס, לתוך הפעילות שתיקרא Assignment1.

שאלות לגבי העבודה ניתן יהיה לשאול בפורום מיוחד שייפתח באתר הקורס – כל הנכתב בפורום זה הוא חלק בלתי נפרד מהוראות העבודה ויש להתעדכן שם.
בפורום יש לכתוב שאלות עם **כותרות אינפורמטיביות** ולוודא כי השאלה שאתם שואלים לא נשאלה כבר בעבר.

בהצלחה !!!

