

מבוא לתכנות מערכות תרגיל בית מספר 2

סמסטר חורף 2019-2020

תאריך פרסום: 11/12/2019

תאריך הגשה: 25/12/2019 שעה 23:55

מתרגל אחראי לתרגיל: אור אייזקס sisaacs@cs.technion.ac.il

1 הערות כלליות

- תרגיל זה מהווה 5% מהציון הסופי

2 הקדמה

מטרת תרגיל זה היא היכרות עם תכנות ב-Python והיכרות עם תפקידה של Python כ-gluing language באמצעות שימוש ב-swig. התרגיל מורכב משני חלקים:

1. שליפת נתונים מקובץ טקסט באמצעות Python.
 2. שימוש ב-swig כדי להשתמש ב-shared library שנכתבה ב-C בתוך תוכנית ה-Python.
- הערות:
- התרגיל להגשה רק בזוגות.

3 חלק א' - שליפת נתונים מקובץ טקסט באמצעות Python

חברת שוקולד עשתה סקר שבו הנשאלים התבקשו לדרג חמישה סוגי שוקולדים חדשים שהחברה הוציאה הממוספרים בין 0 ל-4. הנשאלים התבקשו לתת ציון (Score) מ-1 עד 10. הנתונים נמצאים בקובץ survey.txt, שיסופק על ידינו, ונתון בפורמט הבא (סוגריים הם מספר במקרה של int, אחרת הם מחרוזות), שימו לב כי עלולים להיות כמה רווחים בין כל מילה בשורה:

```
<id (int)> <eating habits (Vegan\Omnivore\Vegetarian)> <age (int)>
<gender(Woman\Man)> <Score of chocolate 1 (int)> <Score of chocolate 2
(int)> <Score of chocolate 3 (int)> <Score of chocolate 4 (int)> <Score of
chocolate 5 (int)>
```

לדוגמא, אם נעשה סקר עם שני נשאלים, הקובץ survey.txt ייראה כך:

```
39401830 Vegan 50 Woman 1 5 3 8 10
29441133 Omnivore 15 Man 1 10 1 3 1
```

היו אנשים שהצביעו יותר מפעם אחת בסקר כאשר הם גילו שהם כתבו פרט לא נכון או החליטו על דירוג שונה (או שכחו שכבר הצביעו). כמו כן, ישנם אנשים שהכניסו מידע לא נכון, כאשר הדרישות למידע בתרגיל הן:

- תעודת זהות בת 8 ספרות.
- גיל בין 10 ל-100 (כולל).

- ציון לשוקולדים חייב להיות בין 1 ל-10 (כולל).

אנחנו מעוניינים להתייחס רק להצבעה האחרונה הנכונה שלהם. עליכם לפלוט למסך פלט מתוקן בו רשימת תשובות הנסקרים צריכה להיות מסודרת בסדר עולה, לפי מספר תעודת הזהות שלהם עם אותם רווחים שהיו בשורה המקורית. כדי לבצע זאת, עליכם למלא בקובץ pythonFuncs.py את הפונקציה בשם correct_myfile שמקבלת את ה-path לקובץ הסקר המקורי. אין צורך להתייחס למקרים בהם הקובץ אינו קיים או בפורמט שונה ממה שהוגדר.

אז לדוגמא, אם הקלט שלנו נראה כך:

```
12345678 Vegan 50 Woman 1 5 3 8 10
02305678 Vegetarian 70 Woman 10 2 1 8 10
39401830 Omnivore 19 Man 7 10 3 5 1
39401830 Omnivore 19 Man 1 1 3 5 1
39401830 Omnivore 190 Man 1 10 3 2 6
02305678 Vegan 70 Woman 10 8 1 8 10
```

נרצה שהפלט שנייצר יראה כך:

```
02305678 Vegan 70 Woman 10 8 1 8 10
12345678 Vegan 50 Woman 1 5 3 8 10
39401830 Omnivore 19 Man 1 1 3 5 1
```

טיפ: השתמשו במתודה split של מחרוזות, ניתן לראות את התיעוד של פונקציה זו בכתובת הבאה:

<https://docs.python.org/3/library/stdtypes.html#str.split>

4 חלק ב' - עבודה עם swig

עליכם לעבור על הקובץ corrected_survey.txt שורה - שורה באמצעות Python, לשלוף את התשובות של הנסקרים ולהשתמש בספרייה C-הכתובה ב-survey והמצורפת לתרגיל שתבצע ניתוח של התשובות. הממשק של הספרייה נמצא בקובץ Survey.h והמימוש מקומפל ב-Survey.o, שני הקבצים מצורפים עם התרגיל.

עליכם להשלים את קובץ הממשק Survey.i שמתממשק עם swig.

בשביל לקמפל את הקבצים יש לבצע את הפקודות הבאות:

1. swig -python Survey.i
2. gcc -std=c99 -fPIC -c Survey_wrap.c -I/usr/local/include/python3.6m
3. ld -shared Survey.o Survey_wrap.o -L/usr/local/include/python3.6m/ -o _Survey.so

הערות:

- שימו לב שניתן לבצע את הקימפול רק על ה-cs13.

לאחר שהצלחתם לחבר את הספרייה עם swig, השתמשו בספריית survey על מנת לבדוק את ההתפלגויות של דירוגי שוקולדים שונים. עליכם להשלים בקובץ pythonFuncs.py:

- פונקציה בשם `scan_survey` שמקבלת כקלט את ה-`path` לקובץ סקר ומחזירה מבנה של `Survey` ב-C שניתן לבצע אינטרקציה איתו (אין צורך להפעיל את הפונקציה `correct_myfile` על הקובץ לפני כן).
- פונקציה בשם `print_info` שמקבלת כקלט מבנה של `Survey` ב-C, סוג שוקולד מ-0 עד 4 (לפי סדר ההכנסה של הדירוג), מין, גיל מינימלי ומקסימלי (כולל), הרגלי אכילה ומדפיסה למסך `list` של פייתון עם מספר ההצבעות הכולל לכל דירוג מ-0 על 9 (הרשימה תיפלט גם במקרים בהם ה-`Survey` לא מכיל מידע).
- פונקציה בשם `clear_survey` שמקבלת כקלט מבנה של `Survey` ב-C ומשחררת אותו.

בכל המקרים אין צורך להתייחס לקלט בלתי חוקי או למקרים של שגיאות הקצאת זיכרון.

דוגמא ל-`print_info`: אם היינו רוצים להפיק את התפלגות הציונים שניתנו לשוקולד 0 בקרב נשים בין גילאים 10 ל-30 בעלות הרגלי אכילה אומניבוריים בסקר שערכה החברה, ותוצאות הסקר שנשמרו ב- `survey.txt` הן

```
12345678 Omnivore 22 Woman 1 5 3 8 10
02305678 Vegan 30 Woman 2 2 1 8 10
39401890 Omnivore 13 Woman 7 10 3 5 1
94645678 Omnivore 21 Woman 1 5 3 8 10
12315678 Vegan 30 Woman 2 2 1 8 10
39401190 Omnivore 29 Woman 9 10 3 5 1
59400890 Omnivore 16 Woman 6 1 3 5 1
59400891 Omnivore 31 Woman 6 1 3 5 1
```

אז ההתפלגות המבוקשת תהיה:

1. שתי נשים נתנו לשוקולד ציון 1
2. אין אישה שנתנה לשוקולד ציון 2
3. אין אישה שנתנה לשוקולד ציון 3
4. אין אישה שנתנה לשוקולד ציון 4
5. אין אישה שנתנה לשוקולד ציון 5
6. אישה אחת נתנה לשוקולד ציון 6
7. אישה אחת נתנה לשוקולד ציון 7
8. אין אישה שנתנה לשוקולד ציון 8
9. אישה אחת נתנה לשוקולד ציון 9
10. אין אישה שנתנה לשוקולד ציון 10

והפלט שיתקבל הוא:

```
[2, 0, 0, 0, 0, 1, 1, 0, 1, 0]
```

5 קבצים מצורפים ופרטי הגשה

עם התרגיל אתם מקבלים את הקבצים: `Survey.h`, `Survey.i`, `Survey.o`, `pythonFuncs.py`. אתם נדרשים להשלים את `Survey.i` ואת `pythonFuncs.py`.

כמו כן אתם מקבלים קבצים `main1.py`, `main2.py`, `survey1`, `survey2`, `exp1`, `exp2`. לאחר כתיבת הפיתרון שלכם יש לוודא שהרצה של השורה הבאה ב-`cs13`:

```
python3.6 main1.py > out1
```

מייצרת קובץ זהה ל-`exp1`, וכנ"ל ל-`main2` עם `out2` ו-`exp2`.

יש להגיש רק את ארבעת הקבצים: `Survey.h`, `Survey.i`, `Survey.o`, `pythonFuncs.py` בתוך תיקיית `ZIP`, לא בתוך תיקיות פנימיות. לא יתקבלו ערעורים על מבנה הגשה לא תקין.

6 שינויים עדכונים והודעות בנוגע לתרגיל

כל ההודעות הנוגעות בתרגיל ימצאו באתר של הקורס <http://webcourse.cs.technion.ac.il/234124> בדף התרגילים. דף זה יכול שאלות ותשובות נפוצות. רק הודעות דחופות תשלחנה בדואר. עליכם לעקוב אחר האתר והעדכונים שיפורסמו בו.

בהצלחה !