234124 - מבוא לתכנות מערכות

תרגיל בית 5

סמסטר חורף 2022-23

תאריך פרסום: 19/01/2023
תאריך הגשה: 26/01/2023
מתרגלים אחראים: אסף במברגר

1 הערות כלליות

- הציון על תרגיל זה מהווה 4% מהציון הסופי.
- התרגיל להגשה בזוגות בלבד. חריגה מהנחייה זו רק באישור המתרגל האחראי של הקורס.
- מענה לשאלות בנוגע לתרגיל יינתן אך ורק בפורום התרגיל בפיאצה או בסדנאות. לפני פרסום שאלה בפורום אנא בדקו אם כבר נענתה – מומלץ להיעזר בכלי החיפוש שהוצגו במצגת האדמיניסטרציה בתרגול הראשון.
- שימו לב: לא תינתנה דחיות במועד הגשת התרגיל פרט למקרים חריגים. תכננו את הזמן בהתאם.
- קראו את התרגיל עד סופו לפני שאתם מתחילים לממש. חובה להתעדכן בעמוד ה-FAQ של התרגיל ובפורום הפיאצה, הכתוב שם מחייב.
- העתקות קוד בין סטודנטים ובפרט גם העתקות מסמסטרים קודמים תטופלנה. עם זאת – מומלץ ומבורך להתייעץ עם חברים על ארכיטקטורת המימוש
- המסמך נכתב בלשון זכר מטעמי נוחות בלבד ומיועד לשני המינים.
- מטרת תרגיל זה היא היכרות עם תכנות ב-Python.

2 מערכת הצפנה ופענוח

2.1 רקע

גנדלף האפור, סטודנט קורס מת"ם, רוצה לזמן ישיבה דחופה בריוונדל כדי לדון כיצד יש לנצח את כוחות האופל אשר רוצים שהוא יכשל בקורס. על מנת למנוע מכוחות האופל לקרוא את ההודעה לגבי מועד ומיקום הפגישה ולתקוף את ריוונדל, גנדלף החליט להצפין את ההודעה.
לצערו, גנדלף עסוק כרגע במסע להר הבודד ולכן הוא לא יכול לכתוב את התוכנה שתצפין ותפענח עבורו את ההודעה, ולכן הוא פנה אליכם בבקשה שתעזרו לו לכתוב תוכנת הצפנה משוכללת.

2.2 קצת על צפנים

צופן (Cipher) הוא זוג אלגוריתמים ("הצפנה" ו-"פענוח") אשר מקבלים כפרמטר מחרוזת ומידע נוסף (אשר נקרא "מפתח") ומחזירים מחרוזת.

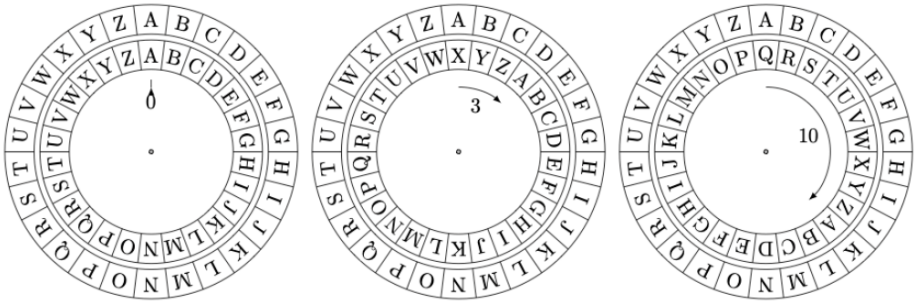
הצפנה היא אלגוריתם אשר, בעזרת שימוש במפתח, מחזיר מחרוזת חדשה אשר מקודדת בתוכה את המחרוזת המקורית באופן שקשה לשחזר ממנו את המחרוזת המקורית ללא ידיעת המפתח.

פענוח הוא אלגוריתם אשר, בעזרת שימוש באותו מפתח כשל ההצפנה, מחזיר את המחרוזת המקורית בהינתן המחרוזת המוצפנת.

כלומר- מתקיים כי הצפנה של מחרוזת ופענוח של המחרוזת המוצפנת לפי אותו מפתח מחזירים את המחרוזת המקורית: $\text{decrypt}(\text{encrypt}(M, K), K) = M$.

בתרגיל זה נתמקד בצפנים אשר מצפינים כל תו במחרוזת בנפרד והמחרוזת המוצפנת היא שרשור הצפנות התווים.[1]

2.3 צופן קיסר



גנדלף למד קצת היסטוריה, ובמיוחד מעניין אותו יוליוס קיסר מהאימפריה הרומית. לפי הסיפורים, יוליוס קיסר השתמש בצופן הנקרא צופן קיסר[2] (או הסטה קיסרית, צופן היסט...) או באנגלית Caesar Cipher.

נפתח בכך שנתאר כיצד פועל צופן קיסר באופן לא פורמלי:

המצפין מקבל מחרוזת כלשהי s וערך הזזה k, שהוא מספר שלם ומהווה המפתח של ההצפנה. עבור כל תו $c \in s$, המצפין יבצע k הזזות של c בכיוון המתאים (לפי הסימן של k , חיובי יוזז ימינה ושלילי שמאלה). למשל, אם $k = 2$, וקיבלנו את התו 'A', אז נזיז אותו פעמיים – פעם אחת ל-'B', ופעם שנייה ל-'C'. הערך 'C' אפוא הוא הערך המתקבל מהצפנת 'A' עם מפתח 2 לפי צופן קיסר.

<u>הערות:</u>

1. הזזות מתבצעות כך שהתווים מסודרים במעין מעגל. כלומר, אחרי Z יש A, לאחר מכן B וכן הלאה.
2. אם k שלילי, ההזזה תתבצע לכיוון השני. למשל, אם $k = -2$ וקיבלנו את התו 'D', נזיז אותו פעמיים – פעם אחת ל-'C', ופעם אחת ל-'B'. גם כאן, לפני A יש Z, ולפני זה יש Y וכן הלאה.

דוגמאות נוספות:

$$\text{encrypt}('\text{a}', 2) = 'c', \qquad \text{encrypt}('\text{A}', 2) = '\text{C}',$$

$$\text{encrypt}('\text{K}', -2) = '\text{I}', \qquad \text{encrypt}('\text{A}', -2) = '\text{Y}'$$

$$\text{encrypt}('\text{a}', 28) = 'c', \qquad \text{encrypt}('\text{A}', -28) = \text{Y}',$$

<u>שימו לב!</u> אנו מתייחסים לאותיות גדולות וקטנות **בנפרד**, ומצפינים רק אותיות אנגליות גדולות וקטנות, כל שאר התווים נשארים ללא שינוי (לכל תו c שאינו A עד Z או a עד z, תוצאת ההצפנה של c היא c בלי תלות במפתח).

נזכור שהגדרנו שתוצאת הצפנה (או פיענוח) של מחרוזת היא שרשור תוצאות ההצפנה (או הפענוח) של התווים במחרוזת. למשל, $\text{encrypt}('\text{aB}', 2) = '\text{cD}'$.

---

[1] אם הנושא של צפנים לא ברור לכם לחלוטין בשלב זה, לא נורא. זאת לא מטרת התרגיל. תוכלו לקחת בהמשך את הקורסים 236350 הגנה ברשתות ו-236506 קריפטולוגיה מודרנית.

[2] על שמו של יוליוס קיסר כמובן

המפענח מקבל מחרוזת מוצפנת כלשהי s' וערך הזחה k. עבור כל תו $c' \in s'$, המפענח יבצע k הזחות של c' בכיוון המנוגד לכיוון ההצפנה.

למשל:

$$\text{decrypt}('c', 2) = 'a', \qquad \text{decrypt}('C', 2) = 'A',$$
$$\text{decrypt}('I', -2) = 'K', \qquad \text{decrypt}('Y', -2) = 'A'$$

הבחנה חשובה: $\text{decrypt}(M, k) = \text{encrypt}(M, -k)$. חשבו כיצד ניתן להשתמש בעובדה זו כדי להימנע משכפול קוד.

מימוש צופן קיסר

ממשו את המחלקה CaesarCipher אשר מכילה את המתודות הבאות:

1. בנאי המקבל מספר ומאתחל את העצם לייצג הצפנה עם המספר בתור מפתח. ניתן להניח שהתקבל מספר ואין צורך לבדוק זאת.
2. **encrypt(self, plaintext: str) -> str**

   המתודה encrypt מקבלת מחרוזת ומחזירה את ההסטה הקיסרית של המחרוזת לפי המפתח של העצם עליו מופעלת המתודה.
   דוגמאות:

```
>>> caesar_cipher = CaesarCipher(3)
>>> caesar_cipher.encrypt('a')
'd'
>>> caesar_cipher.encrypt('Mtm is the BEST!')
'Pwp lv wkh EHVW!'
```

3. **decrypt(self, ciphertext: str) -> str**

   המתודה decrypt מקבלת מחרוזת ומחזירה את הפענוח של המחרוזת שהוצפנה באמצעות צופן קיסר לפי המפתח של העצם עליו מופעלת המתודה.
   דוגמאות:

```
>>> caesar_cipher.decrypt('d')
'a'
>>> caesar_cipher.decrypt('Pwp lv wkh EHVW!')
'Mtm is the BEST!'
```

2.4 צופן ויז'נר (Vigenère)

בניגוד לגנדלף האפור, גנדלף הסגול למד על צופן קיסר, אבל הוא חושב שהצופן חלש מידי וכוחות האופל יצליחו לפענח אותו בקלות.
לכן, גנדלף המשיך הלאה בלימודי ההיסטוריה שלו עד שלמד על צופן ויז'נר (שפותח באיטליה במאה ה-16). בצופן החדש, המפתח יכיל **רשימה של ערכים** (ולא ערך בודד) כאשר כל אות בטקסט המוצפן תוצפן באמצעות מפתח מתוך רשימת המפתחות בצורה מחזורית.

למשל, עבור המחרוזת "come to Rivendell!" והמפתח [7,8,11,13, −2,4] נקבל כי המחרוזת המוצפנת תהיה:

c	o	m	e		t	o		R	i	v	e	n	d	e	l	l	!
+7	+8	+11	+13		-2	+4		+7	+8	+11	+13	-2	+4	+7	+8	+11	
j	w	x	r		r	s		Y	q	g	r	l	h	l	t	w	!

כלומר: "jwxr rs Yqgrlhltw!"

<u>הערה:</u> בדומה לצופן קיסר, בצופן ויז'נר מצפינים רק אותיות אנגליות גדולות וקטנות, כל שאר התווים נשארים ללא שינוי. בצורה דומה, הפענוח של מחרוזת גם הוא יבוצע אות–אות באמצעות מפתח מרשימת המפתחות.

למשל, עבור הדוגמא לעיל, עבור המחרוזת המוצפנת "jwxr rs Yqgrlhltw!" והמפתח [7,8,11,13, −2,4] נקבל:

j	w	x	r		r	s		Y	q	g	r	l	h	l	t	w	!
+7	+8	+11	+13		-2	+4		+7	+8	+11	+13	-2	+4	+7	+8	+11	
c	o	m	e		t	o		R	i	v	e	n	d	e	l	l	!

מימוש צופן ויז'נר

<u>חלק ראשון</u>

ממשו את המחלקה VigenereCipher אשר מכילה את המתודות הבאות:

1. בנאי המקבל <u>רשימה של מספרים</u> כמפתח ומאתחל את העצם לייצג הצפנה עם הרשימה בתור מפתח.
   ניתן להניח שהתקבלה רשימה של מספרים ואין צורך לבדוק זאת.
2. **encrypt(self, plaintext: str) -> str**
   המתודה encrypt מקבלת מחרוזת ומחזירה את הצפנת ויז'נר של המחרוזת לפי המפתח של העצם עליו מופעלת המתודה.
3. **decrypt(self, ciphertext: str) -> str**
   המתודה decrypt מקבלת מחרוזת ומחזירה את הפענוח של המחרוזת שהוצפנה באמצעות צופן קיסר לפי המפתח של העצם עליו מופעלת המתודה.

דוגמאות:

```
>>> vigenere_cipher = VigenereCipher([3])
>>> print(vigenere_cipher.encrypt('l'))
o
```

```
>>> vigenere_cipher = VigenereCipher([2, -4, -14, -16, -17, -17])
>>> print(vigenere_cipher.encrypt('we wish you best of luck in all of your exams'))
ya isbq akg lnbv kr vdlm ez kuu qb kyda gtmwb
>>> print(vigenere_cipher.decrypt('ya isbq akg lnbv kr vdlm ez kuu qb kyda gtmwb'))
we wish you best of luck in all of your exams
```

```
>>> vigenere_cipher = VigenereCipher([1,2,3,4,-5])
>>> print(vigenere_cipher.encrypt('Hello World!'))
Igopj Xqupy!
>>> print(vigenere_cipher.decrypt('Igopj Xqupy!'))
Hello World!
```

חלק שני

על מנת להקל על זכירת המפתח נוהגים שהמפתח הוא משפט באנגלית, למשל:
"python rules, C drools", כאשר כמובן מתעלמים מרווחים וכל תו שאינו אות.
בשיטה זו כל אות מיתרגמת למספר לפי האינדקס שלה $(a \to 0, b \to 1, \ldots)$.
למשל, המחרוזת "abAbc" מתורגמת למפתח [0, 1, 26, 1, 2].
כתבו את הפונקציה החיצונית **getVigenereFromStr(keyString: str) -> VigenereCipher** אשר מקבלת משפט מפתח (**מחרוזת**) ומחזירה **אובייקט מטיפוס ויז'נר** אשר מתאים למפתח המתקבל.
דוגמה:

```
>>> vigenere_from_str = getVigenereFromStr('python rules, C drools')
>>> vigenere_from_str.encrypt('JK, C is awesome')
'YI, V pg nnydseg'
>>> vigenere_from_str.decrypt('YI, V pg nnydseg')
'JK, C is awesome'
```

2.5 מימוש המערכת

בחלק זה נממש את המערכת בה גנדלף ישתמש להצפנת ופענוח ההודעות שלו.

ממשו את הפונקציה **processDirectory(dir_path: str) -> None** אשר מקבלת נתיב לתיקייה[3] שבה יש קובץ בשם config.json אשר בו שמורים הערכים הבאים:

- **type**: מחרוזת שהיא אחת מבין: Caesar, Vigenere אשר מגדירה באיזה סוג הצפנה עלינו להשתמש.
- **mode**: מחרוזת שהיא אחת מבין: encrypt, decrypt אשר מגדירה האם נבצע הצפנה או פענוח.
- **key**: מפתח שבאמצעותו נצפין/נפענח. אם שיטת ההצפנה היא Caesar הוא יהיה מספר, אם השיטה היא Vigenere הוא יכול להיות מחרוזת או רשימה של מספרים.

הפונקציה תטען את הקובץ ותיצור מערכת מתאימה, ואז תבצע את הפעולה המתוארת ב-config על כל הקבצים בתיקייה (באופן לא רקורסיבי) באופן המתואר כאן. הפונקציה לא תמחק את הקבצים המקוריים אלא רק תיצור קבצים חדשים. אם קבצי המטרה כבר קיימים, הם יידרסו ע"י הפעולה.

אם הפעולה הדרושה היא הצפנה, הפונקציה תצפין כל קובץ בתיקייה בעל סיומת txt. לקובץ בעל שם זהה אך עם סיומת enc.
txt → enc

אם הפעולה הדרושה היא פענוח, הפונקציה תפענח כל קובץ בתיקייה בעל סיומת enc. לקובץ בעל שם זהה אך עם סיומת txt.
enc → txt

הערה: במידה ואנו מצפינים/ מפענחים בשיטת צופן ויז'נר בהצפנה של כל קובץ עלינו להתחיל מתחילת המפתח.

רשות (לא ייבדק): נסו לממש את המחלקות Caesar ו-Vigenere באמצעות ירושה בפייתון כפי שראינו בתרגול.

דגשים נוספים:

- בסעיף זה אתם רשאים להשתמש בספריה os ובכל ספריה אחרת שנלמדה בקורס.

[3] directory, folder, ספריה

3 הערות

- בכל התרגיל אין להשתמש במספרי קסם למעט 0/1.
- ניתן להניח כי הקלט תקין מבחינת טיפוסים בכל התרגיל.
- ודאו כי אתם מריצים פייתון גרסה 3.6. שימו לב כי גרסה זו אינה גרסה ברירת מחדל על השרת. כדי להריץ פייתון 3.6 השתמשו בפקודה python3.6.
- פתרון התרגיל צריך לעבוד בכל מערכת הפעלה (platform independent).
- שימו לב שאתם לא נדרשים לעמוד בקונבנציות קוד, אבל אתם כן נדרשים לעמוד בכללי התכנות הנכון שלמדנו בקורס.

4 הגשה

את ההגשה יש לבצע דרך אתר הקורס, תחת Assignments -> HW5 -> Electronic Submission

הקפידו על הדברים הבאים:

- יש להגיש את קבצי הקוד מכווצים לקובץ zip (לא פורמט אחר).
- אין להגיש קבצים נוספים מלבד קובץ אחד בשם ex5.py בתוך ה-ZIP.
- לתרגיל זה לא יפורסמו טסטים פומביים או קובץ finalCheck, זאת במטרה להרגיל אתכם לכתוב קוד על בסיס מפרט כתוב ולא רק על בסיס טסטים.
- ניתן להגיש את התרגיל מספר פעמים, רק ההגשה האחרונה נחשבת .
- על מנת לבטח את עצמכם נגד תקלות בהגשה האוטומטית, שימרו את קוד האישור עבור ההגשה. עדיף לשלוח גם לשותף.
  - כל אמצעי אחר לא יחשב הוכחה לקיום הקוד לפני ההגשה .

בהצלחה!