

תרגיל 2 - פתרון

שאלה 1:

i. ההגדרות לאחר טוקניזציה:

1. [קובץ דפים מאוגד יחד, עטוף ב כריכה רכה או קשה, ה מיועד ל קריאה ו מכיל מידע או סיפור]
2. [סכס, מנה, קבע את ה מספר או ה כמות של אוסף פרטים ב אמצעות מנייתם]
3. [גלב, אדם אשר מקצועו הוא הסרת שערות ו עיצובן ב אמצעות גזירה, גילוח ו כדומה]
4. [רישום ה חשבונות של עסק]
5. [אזור או ישוב ה נמצא ליד ה גבול]

ii. המשפטים לאחר טוקניזציה:

1. [ה התיישבות ב ספר שינתה את אופיים של ה צפון ו ה דרום.]
2. [עוביו של ספר מוגבל ל סנטימטרים מעטים, ו לכן כאשר תכולתו עולה על 1,000 דפים, נהוג לפצלו.]
3. [ספר מקצועי כ יום מטפל ב דרך כלל רק ב שיער ה ראש.]
4. [ה ילד ספר את ה בתים ב שכונה.]
5. [אזור ה ספר הוא מקום מסוכן יחסית ל מרכז]
6. [ב ספר ה בישול יש עשרה מתכונים ל מנה ראשונה ו עשרים ל מנה עיקרית.]

משפט	משמעות נכונה	מילים רומזות	בחירת האלגוריתם	מילים חופפות
1	5	התיישבות, צפון, דרום	1	--
2	1	עוביו, דפים	1	דפים
3	3	מקצועי, שיער, ראש	1	--
4	2	-- (המילה משמשת כפועל)	1	--
5	5	אזור, מקום, מרכז	3	הוא
6	1	בישול	2	מנה

שאלה 2:

1. [The famous author]_{Arg0} **writes** [every single one of his books]_{Arg1} on a typing machine.
2. [His sister]_{Arg0} **wrote** [him]_{Arg2} [a note]_{Arg1} yesterday.
3. [The young children]_{Arg0} **write** slowly and carefully.
4. [This novel]_{Arg1} was **written** by [Jane Austen]_{Arg0}.
5. [Her story]_{Arg1} was never **written**.
6. Please **write** to [me]_{Arg2} when you get home.

שאלה 3:

overlap(L, M):

c ← 0

i ← 1

while i ≤ len(L):

תרגיל 2 - פתרון

```
j ← 1
while j ≤ len(M):
    if L[i] = M[j]:
        c ← c + 1
    j ← j + 1
i ← i + 1
return c
```