

# תכנות מתקדם

## תרגיל 2

### חיפוש מהיר בקובץ

מועד אחרון להגשה: 2.12.2020

משקל התרגיל: 5 נקודות

מטרות התרגיל: STL, shared\_ptr, Inheritance, RegEx

בתרגיל נממש חיפוש מהיר בקובץ על ידי מפה מהמילים לשורות בקובץ.

#### הנחיות הגשה

- הגשת התרגיל באתר:

<http://submit.org.il/ariel>

שמות המשתמשים בקובץ login.pdf הנמצא בתיקיית תרגיל 1.

- יש להוריד מאתר מודל את הקבצים הנמצאים בתיקיית תרגיל 2.

#### תיאור התוכנית

תכנית זו מבוססת על תכנית שמתוארת בספר C++ Primer.

(מצורפים דפים מתוך הספר, התכנית דומה אך לא זהה.)

התכנית קוראת את הקובץ לתוך ווקטור של מחרוזות:

```
vector<string>>* file; // simplified
```

התכנית יוצרת מפה מהמילים בקובץ, למצביעים לסט של השורות בהן הן נמצאות:

```
map<string, set<int>*> wm; // simplified, wm - word map
```

כשמחפשים מילה, אפשר למצוא במפה את סט השורות ולפי הסט לדעת את מספר השורות בהן היא נמצאת ואת רשימת השורות המכילות אותה.

אם המילה מופיע כמה פעמים באותה שורה, השורה צריכה להופיע בתוצאה רק פעם אחת.

רשימת השורות בתוצאה צריכה להופיע בסדר עולה.

סט מתאים לכך כי הוא מכיל כל פריט רק פעם אחת ונותן את הפריטים בסדר עולה.

### תיאור התרגיל

התכנית כפי שהורדתם, כבר יוצרת את האינדקס, ומממשת:

חיפוש של שורות המכילות מלה.

חיפוש של שורות המכילות שתי מלים (AND).

חיפוש של שורות המכילות אחת משתי מלים (OR).

התוספות הדרושות:

בקובץ main.cpp ישנה לולאה שמקבלת קלט מהמשתמש, המשתמש יוכל להקליד:

1.

```
word
```

כדי למצוא שורות המכילות מילה אחת.

2.

```
AND word1 word2
```

כדי למצוא שורות המכילות שתי מלים.

3.

```
OR word1 word2
```

כדי למצוא שורות המכילות אחת משתי מלים.

AD word1 word2

כדי למצוא שורות המכילות את שתי המילים בשתי שורות עוקבות (Adjacent).  
מילה ראשונה בשורה אחת ומילה שניה בשורה שאחריה, או להיפך.  
תוצאת החיפוש תדפיס את מספר הפעמים בהם יש שורות עוקבות ואת השורות העוקבות.  
(הגישו את המטלה וראו את הפלט הנכון)

בקובץ Query.cpp ישנה מתודה factory שיוצרת את אובייקט החיפוש לפי מחרוזת החיפוש.  
מחרוזת חיפוש תקין מכילה מילה אחת או שלוש מילים שהראשונה בהם היא AND , OR , או AD.  
המילים במחרוזת החיפוש מופרדים בתוו רווח אחד או יותר (יתכן גם לפני המילה הראשונה).  
כל חיפוש אחר אינו תקין.  
אם החיפוש לא מתאים לאף אחד מהחיפושים התכנית תדפיס :

Unrecognized search

### המטלה

חלק ניכר מהמטלה הוא להבין את התכנית הקיימת.  
מתוך הקבצים שהורדתם עליכם לערוך ולהגיש רק את הקובץ Query.cpp.  
בקובץ Query.cpp צריך להשלים (בין השורות המסומנות //) את QueryBase::factory  
את Adjacent::eval ואת פונקצית ההדפסה כדי שתדפיס נכון עבור חיפוש AD.  
לאחר שהשלמתם תוכלו לבנות את התכנית ולבדוק אותה.  
ישנה אפשרות בקובץ main להדפיס את המפה לצורך בדיקה (להוריד את ההערה).

### בהצלחה