מבוא למערכות מחשוב

**עבודה מספר 3.1 – Tokenizing**

**שימו לב‼** הסימסטר הולך ונגמר. עליכם לסיים את כל העבודות לפני סוף הסימסטר. לא יאושרו הארכות לתוך תקופת הבחינות!

**כללי** – בעבודה זאת תממשו Tokenizer המקבל קבצי קוד ומבצע עליהם tokenizing – הפיכת הקוד לרשימה של tokens.

ממשו את השיטה Tokenize במחלקה Compiler. על השיטה לתמוך בתחביר שנלמד בשיעור לשפת Jack. במחלקה Token כבר הוגדרו כל האטומים של השפה וכל סוגי האטומים. עליכם ליצור עבור כל token אובייקט מהטיפוס המתאים. המחלקות היורשות מ-Token כבר מוגדרות בפרוייקט.

בנוסף, המחלקה Compiler כבר מכילה שיטות לעבודה עם מחרוזות, כולל שיטה Next המחזירה את תת המחורזת הבאה בהינתן רשימה של delimiters (חוצצים), והשיטה Split המפצלת מחרוזת לתתי מחרוזות לפי delimiters.

שימו לב לזיהוי נכון של מיקום ה-Token. עליכם לזהות שגיאות ובמידת הצורך לזרוק SyntaxErrorException. בעבודה זאת עליכם לזהות מקרים בהם symbol או Identifier לא חוקי. מצורפת דוגמא בקבצי הקוד.

להזכירכם, העבודה אינה בזוגות. כל העברה של קטעי קוד בין סטודנטים אסורה בהחלט. כל צורה של העתקה תטופל בחומרה המירבית – אל תעתיקו!