

# ניתוח לקסיקלי – מבט מדויק יותר

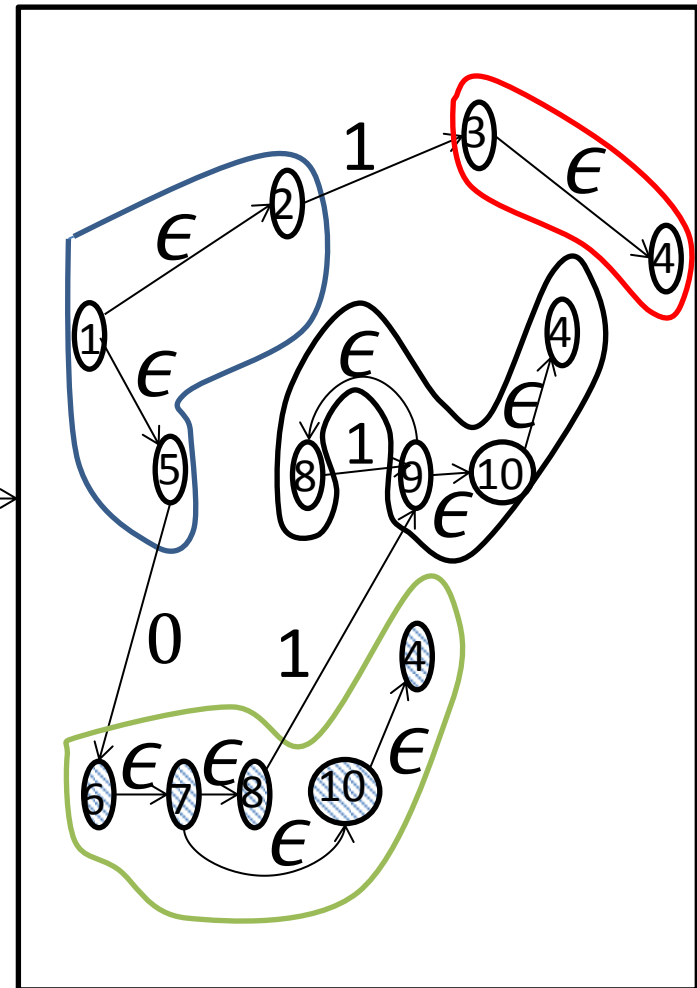
או: איך סידרתי את הסטודנטים בתרגול  
הראשון עם תיאור פשטני מדי של מנתח  
לקסיקלי!

# תזכורת מהתרגול שעבר

```
"+"      { return symbol(TokenNames.PLUS);}  
"- "     { return symbol(TokenNames.MINUS);}  
"do"     { return symbol(TokenNames.DO);}  
"double" { return symbol(TokenNames.DOUBLE);}  
"/"      { return symbol(TokenNames.DIVIDE);}  
"("      { return symbol(TokenNames.LPAREN);}  
")"      { return symbol(TokenNames.RPAREN);}  
[a-z]+   { return symbol(  
    TokenNames.ID,  
    new String( yytext()));  
}
```

מנתח

לקסיקלי



# תרגיל כיתה: נתחו את הקוד לקסיקלית

```
= class Person {  
    public:  
        int ID;  
        int age;  
        float lastYearSalaries[12];  
};  
  
= int main(void)  
{  
    Person MoishFrenkel;  
    Person *OrenIshShlom;  
    OrenIshShlom->lastYearSalaries[7] = 7800.50;  
    MoishFrenkel.ID = 38895247;  
    MoishFrenkel.age++;  
    OrenIshShlom->age >> 1;  
  
    return 0;  
}
```

# תיאור מדויק של פעולת המנתח הלקסיקלי

- ממקום הימצאו של הראש הקורא, החזר את הטוקן החוקי הארוך ביותר.
- אם יש יותר מאפשרות אחת, החזר את זאת שנמצאת ראשונה בקובץ הספסיפיקציה של הביטויים הרגולריים (קובץ הלקס)

# שאלה ממבחן (אונ' סטנפורד)

## 1. Lexical Analysis (10 points)

Consider the following flex-like lexical specification:

```
a*b    { print "1" }  
ca     { print "2" }  
a*ca*  { print "3" }
```

Given the following input string

abcaacacaaabbbaabcaaca

what does the lexer print?

# שאלה ממבחן (אונ' ת"א)

שגיאות?	קוד
<div data-bbox="357 588 1079 682">Build: 1 succeeded</div>	<pre data-bbox="1421 415 1846 853">void f() {     int i=8;     int j=3;      i--+--j; }</pre>
	<pre data-bbox="1421 911 1846 1349">void f() {     int i=8;     int j=3;      i-----j; }</pre>

# שאלה ממבחן (אונ' ת"א)

שגיאות?	קוד
Build: 1 succeeded	<pre>void f() {     int i=8;     int j=3;      i--++j; }</pre>
error C2105: '---' needs l-value	<pre>void f() {     int i=8;     int j=3;      i-----j; }</pre>

# מספרים שליליים במנתח הלקסיקלי

- מה המנתח הלקסיקלי אמור להחזיר כשהוא מזהה מספר שלילי? נפריד בין דרישות ההדפסה בתרגיל הראשון, לבין מנתח לקסיקלי שאמור לשמש את השלבים הבאים בקומפילציה.
- כדי להבין את הפעולה הנכונה יש להסתכל קדימה על מנתח הסינטקס ולהבין כיצד הוא רואה את הקלט.
- נסתכל למשל על פעולת החיסור הבאה: 8-num



# מספרים שליליים במנתח הלקסיקלי

- המנתח הלקסיקלי עובד על טוקנים שהחזיר לו הלקסר, ולכן, מה שהוא בעצם רואה זה:

- `INT(8) BINOP(-) ID(num)`

- כעת נחשוב, מה קורה בקלט הבא: 8-9?

- האם אפשר לנקוט באותה המדיניות שבה נוקט הלקסר של תרגיל 1?

- מה האלטרנטיבה? מה לקסר אמיתי עושה:

- <http://www.lysator.liu.se/c/ANSI-C-grammar-l.html>

# מילים שמורות במנתח הלקסיקלי

- אצלנו בתרגיל הראשון בחרנו להחזיר את המילה השמורה void בתור identifier. האם בחירה כזאת אפשרית גם עבור מנתח לקסיקלי שהפלט שלו דרוש להמשך שלבי הקומפילציה?
- מה האלטרנטיבה? מה לקסר אמיתי עושה:
- <http://www.lysator.liu.se/c/ANSI-C-grammar-l.html>
- זאת כנראה הזדמנות טובה להציץ בדקדוק של שפת C כדי לראות את ההתייחסות ל void שם:
- <http://www.lysator.liu.se/c/ANSI-C-grammar-y.html>