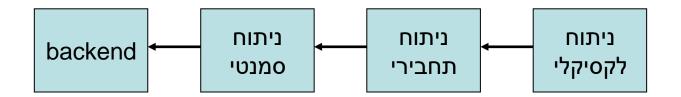
ניתוח תחבירי Top-Down

עודכן: סמסטר חורף 2022

תזכורת מתרגול אחרון

• ראינו מהו המבנה הסכמתי של קומפיילר



- לקסמות, אסימונים, וביטויים רגולריים.
 - דנו בתכונות המנתח הלקסיקלי.
 - .Flex ראינו את הכלי

דרישת מהמנתח התחבירי

- שלמות: המנתח צריך לזהות <u>כל</u> מילה בשפה:
 - אם התכנית היא חוקית,המנתח יצליח לזהות אותה ויבנה עץ גזירה.

- **נאותות**: אסור למנתח לקבל מילים שאינן בשפה:
 - אם התכנית אינה חוקית,המנתח לא יצליח לזהותה ויודיע על שגיאה.

ניתוח Top-Down

- הניתוח מתחיל מהמשתנה התחילי (S).
 - מפעילים רצף של כללי גזירה.
 - :עוצרים כאשר
- מגיעים למילת הקלט (התוכנית כולה) ⇒ המילה בשפה.
- נתקעים (לא ניתן להפעיל אף כלל גזירה) ⇒ המילהאינה בשפה.

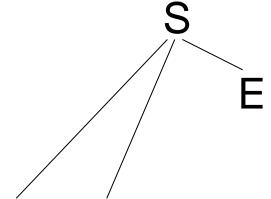
$$x = y + 1$$

• בקוד המקור:

id assign id op num

• אחרי ניתוח לקסיקלי:

:Top-Down ניתוח תחבירי



שימוש בכלל S
$$\rightarrow$$
 id assign E

id assign id op num

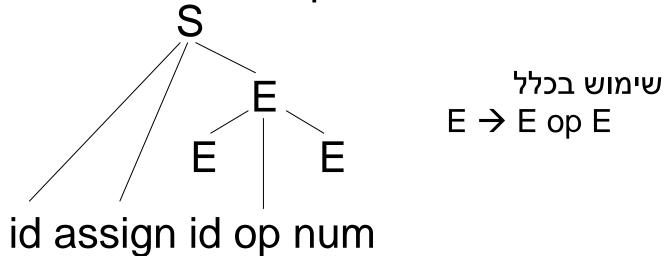
$$x = y + 1$$

• בקוד המקור:

id assign id op num

• אחרי ניתוח לקסיקלי:

:Top-Down ניתוח תחבירי



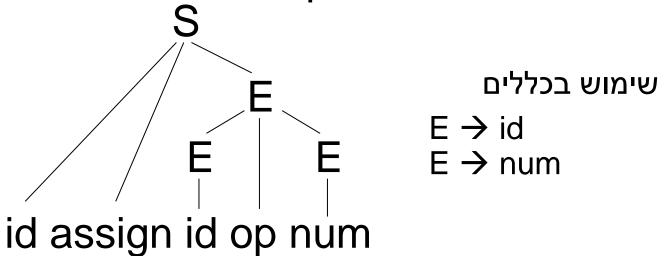
$$x = y + 1$$

• בקוד המקור:

id assign id op num

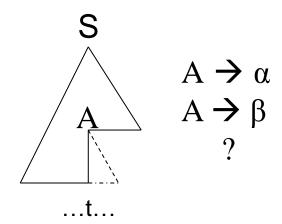
* אחרי ניתוח לקסיקלי:

:Top-Down ניתוח תחבירי



ניתוח Top-Down

? הבעיה העיקרית: איזה כלל להפעיל בכל צעד



- הסיטואציה: •
- .t וכרגע בקלט האסימון A, וכרגע בקלט האסימון
 - השאלה: באיזה כלל של A להשתמש?

Recursive Descent שיטת

- הרעיון: נכתוב תכנית ספציפית לדקדוק.
- נכתוב לכל משתנה פונקציה שיודעת "לגזור" את כל המילים שניתן לקבל ממנו.
 - נבחר את כלל הגזירה לפי הסדר הבא:
- :t והטרמינל הבא בקלט A כשמנסים לגזור את המשתנה
 - יש כלל שמתחיל ב-t, בחר בו A יש כלל
- אחרת, אם ל-A כלל יחיד שמתחיל במשתנה, בחר בו
 - אחרת, החזר הודעת שגיאה

RD בעיות בשיטת

- המנתח לא יודע לקבל כל מני דקדוקים יחסית טריוויאליים:
 - ?עבור $A \rightarrow C\beta$, $A \rightarrow B\alpha$ עבור •
 - ?עבור $A \rightarrow a\beta$, $A \rightarrow a\alpha$ עבור •
 - המנתח לא שלם הוא לא מזהה את כל המילים בשפה
 - ,B→a ,S→a | Bb = בדקדוק: ומילת קלט "ab":
 - נשמח אם המנתח יבחר בכלל S→Bb, אבל הוא יבחר בכלל S→a אבל הוא יבחר
 - רקורסיה שמאלית תגרום לריצה אינסופית י
 - .b כאשר התו הבא בקלט אינו , S ightarrow Sa | b -

שיפור מנגנון ההחלטה לבחירת כלל

קבוצת הטרמינלים בשפה קבוצת כללי הגזירה

נגדיר פונקציה: $P \rightarrow 2^T$ נגדיר פונקציה: $P \rightarrow 2^T$ שמגדירה לכל כלל ב P, איזו קבוצת טרמינלים שמגדירה לכל כלל ב P, ליטרלים) בקלט יגרמו לבחירת הכלל הזה.

אלגוריתם RD משופר:

זהה לאלגוריתם RD שראינו, פרט לשינוי הבא: כל פעם ששוקלים אם להפעיל חוק, בודקים אם התו הבא בקלט שייך לselect של החוק

מנתח (1) LL

- לבחירת הכלל הבא שייגזר select •
- קורא קלט משמאל לימין, מייצר ניתוח שמאלי ביותר ונעזר lookahead-ב-lookahead

$?A \rightarrow \alpha$ עבור טרמינל t בקלט, מתי ייבחר הכלל

:פתרון נאיבי

- אם [t∈T, β ∈(V∪T)*] \rightarrow t β אם נבחר בכלל מהצורה
- [B \in V, $\beta\in$ (V \cup T)*] \Rightarrow **B\beta** נבחר בכלל מהצורה B β יכולה לגזור מילה המתחילה ב-t:
 - : נשתמש בפונקצית עזר first, עבור כל תבנית פסוקית lpha בדקדוק

$$\mathsf{T}\supseteq\mathsf{first}(lpha)=\left\{egin{array}{l} \mathsf{first}(lpha)\ \mathsf{array}\end{array}
ight.$$
 שיכולים להופיע בתחילת מילה שהתבנית $lpha$ גוזרת

 $select(A \rightarrow \alpha) = first(\alpha)$:ננסה להגדיר

Oh Noez!

$$S o AB$$
 ?פה יקרה בדקדוק הבא $A o C$ $select(S o AB) = first(AB) = \{t\}$ $C o \epsilon$: $A o C$ $Select(AB) = first(C) = \emptyset$

- **•** כלומר:
- לעולם לא ייבחר A o C –
- אין אף כלל אחר ל-A, הדקדוק לא יגזור כלום! מכיוון שאין אף כלל

פתרון: להתייחס לכללים אפיסים

אם הכלל כולו אפיס $(A \to^* \epsilon)$ נרצה לבחור אותו • לפי טרמינלים שיוכלו להופיע אחריו בכללי גזירה

$$S \to AB$$

$$A \to C$$

$$C \to \epsilon$$

$$B \to t$$

 $select(A \rightarrow C) = first(B) = \{t\}$ נרצה •

:עבור כל משתנה בדקדוק , follow גדיר פונקצית עזר נוספת -

הגדרות פורמאליות

- First: $(V \cup T)^*$ → 2^T •

כל הטרמינלים שיכולים להופיע בתחילת מילה שתבנית α גוזרת:

First
$$(\alpha) = \{ t \in T \mid \alpha \rightarrow^* t\beta \}$$

- Follow: $V \rightarrow 2^T$ •

כל הטרמינלים שיכולים להופיע מיד לאחר המשתנה A בגזירה <u>כלשהי</u>:

Follow (A) = {
$$t \in T \cup \{\$\} \mid S\$ \rightarrow^* \alpha At\beta$$
 }

. משתנה תחילי של הדקדוק, \$ - סימן סוף קלט (אינו חלק מאוסף הטרמינלים). $\alpha,\beta\in (V\cup T)^*$

- Select: $P \rightarrow 2^T$ •
- לכל כלל בדקדוק, זו קבוצת הטרמינלים (אסימונים) האפשריים שיגרמו לכלל להיבחר, אם האסימון נמצא בראש הקלט כרגע:

Select
$$(A \rightarrow \alpha) = \begin{cases} first(\alpha) \cup follow(A) & \alpha \rightarrow^* \epsilon \\ first(\alpha) & otherwise \end{cases}$$

first אלגוריתם לחישוב פונקצית

- . שלב 0 בודקים אילו משתנים הם אפיסים.
- עבור כל סימן X∈(V∪T) עבור כל סימן first $X \in (V \cup T)$ ע"פ האלגוריתם הבא:

:אתחול

```
first(t) := \{t\} בצע: t \in T לכל טרמינל –
```

$$first(A) := \emptyset$$
 בצע: $A \in V$ – לכל משתנה $-$

- $A \rightarrow Y_1 Y_2 ... Y_k$ צעד: כל עוד יש שינויים, לכל כלל $Y_i \in (V \cup T)$ צעד: $Y_i \in (V \cup T)$
 - $first(A) := first(A) \cup first(Y_1) -$
 - :עבורו $Y_1 Y_2 ... Y_{i-1} \rightarrow *\varepsilon$ עבורו i > 1, בצע

$$first(A) := first(A) \cup first(Y_i)$$

<u>שימו לב</u>: יש כאן תלות מעגלית! האלגוריתם הוא איטרטיבי ופועל עד שאין יותר שינויים.

first אלגוריתם לחישוב פונקצית

• שלב 2 – חישוב first עבור תבנית פסוקית $\alpha = x_1 x_2 \dots x_k$

$$first(\alpha) := \bigcup_{\substack{x_1 x_2 \dots x_{j-1} \to {}^* \varepsilon}} first(x_j)$$

- כלומר:
- איחוד עם, first(x_1) –
- במקרה ש- \mathbf{x}_1 אפיס, איחוד עם first(\mathbf{x}_2) –
- ... במקרה ש- \mathbf{x}_1 אפיסים, איחוד עם first(\mathbf{x}_3) –

• נתון דקדוק:

 $B \rightarrow b \mid \epsilon$

 $C \rightarrow c \mid \epsilon$

.S,C,B :שלב 0 – המשתנים האפיסים

ישלב 1 - חישוב first עבור המשתנים: •

	S	В	С
1	Ø	Ø	Ø

```
Y_i \in (V \cup T) אבעד: כל עוד יש שינויים, לכל כלל Y_1 = Y_1 = Y_1 = Y_1 = Y_2 + Y_2 = Y_1 = Y_1 = Y_1 = Y_2 + Y_2 = Y_1 = Y_
```

 $first(A) := first(A) \cup first(Y_i)$ בצע: $Y_1Y_2...Y_{i-1} \rightarrow *\epsilon$ עבורו - לכל 1 - לכל 1 -

• נתון דקדוק:

 $B \rightarrow b \mid \epsilon$

 $C \rightarrow c \mid \epsilon$

S,C,B – המשתנים האפיסים:• שלב 0 – המשתנים האפיסים:

ישלב 1 - חישוב first עבור המשתנים: •

	S	В	С
1	Ø	Ø	Ø
2	{a,d}	{b}	{C}

צעד: כל עוד יש שינויים, לכל כלל $Y_1 = Y_1 =$

 $first(A) := first(A) \cup first(Y_1)$ -

 $first(A) := first(A) \cup first(Y_i)$ בצע: $(Y_1Y_2...Y_{i-1} \rightarrow *\epsilon$ עבורו i > 1 -

• נתון דקדוק:

 $B \rightarrow b \mid \epsilon$

 $C \rightarrow c \mid \epsilon$

.S,C,B :שלב -0 – המשתנים האפיסים

ישלב 1 - חישוב first עבור המשתנים: •

	S	В	С
1	Ø	Ø	Ø
2	{a,d}	{b}	{c}
3	{a,d,b,c}	{b}	{c}

• נתון דקדוק:

 $B \rightarrow b \mid \epsilon$

 $C \rightarrow c \mid \epsilon$

.S,C,B :שלב 0 – המשתנים האפיסים

ישלב 1 - חישוב first עבור המשתנים: •

	S	В	С
1	Ø	Ø	Ø
2	{a,d}	{b}	{c}
3	{a,d,b,c}	{b}	{c}
4	{a,d,b,c}	{b}	{c}

```
S \rightarrow aB \mid BC \mid CBd
B \rightarrow b \mid \epsilon
C \rightarrow c \mid \epsilon
                        :שלב 2 - חישוב first לאגפי ימין
first(aB) = \{a\}
first(BC) = first(B) \cup first(C) = \{b,c\}
first(CBd) = first(C) \cup first(B) \cup first(d) = \{c,b,d\}
```

$$first(\alpha) := \bigcup first(x_j)$$

$$x_1x_2...x_{i-1} \rightarrow^* \varepsilon$$

אלגוריתם לחישוב פונקצית follow

follow(A) = { $\mathbf{t} \in \mathsf{T} \cup \{\$\} \mid \mathsf{S}\$ \rightarrow^* \alpha \mathsf{A} \mathbf{t} \beta$ } רוצים לחשב

:אתחול

- follow(S):= $\{\$\}$ = :S למשתנה התחילי
- follow(A):= \emptyset : A \neq S לכל משתנה –
- :צעד: כל עוד יש שינויים, לכל משתנה $A \in V$ בצע
- :בצע: $\mathbf{Y} \rightarrow \alpha \mathbf{A} \boldsymbol{\beta}$ בדקדוק (כל כלל בו \mathbf{A} מופיע איפשהו באגף ימין) בצע $\mathbf{Y} \rightarrow \alpha \mathbf{A} \boldsymbol{\beta}$
 - $follow(A) := follow(A) \cup \underline{first(B)} \bullet$
 - $follow(A) := follow(A) \cup \underline{follow(Y)}$ בצע גם: $\beta \rightarrow *\epsilon$ ואם \bullet
 - .αAβ יכול להופיע אחרי Y, אזי הוא יכול להופיע גם אחרי t אבל אם t אפיס, אזי t בעצם יכול להופיע גם אחרי t ...]

follow המשך דוגמה - חישוב

ישוב follow (לכל משתנה דקדוק): •

	S	В	С
1	{\$ }	Ø	Ø

follow המשך דוגמה - חישוב

 $S \rightarrow aB \mid BC \mid CBd$ $B \rightarrow b \mid \epsilon$ $C \rightarrow c \mid \epsilon$

ישוב follow (לכל משתנה דקדוק): •

	S	В	С	
1	{\$ }	Ø	Ø	
2	, {\$}	{ç,\$,q}	{b,d,\$}	
מתוך S→BC מתוך S→CBd				

משתנה S לא מופיע בצד ימין של אף כלל גזירה

לכל כלל $\alpha A \beta$ בדקדוק בצע:

• נתון דקדוק:

 $follow(A) := follow(A) \cup \underline{first(\beta)}$

 $follow(A) := follow(A) \cup \underline{follow(Y)}$ בצע גם: $\beta \rightarrow *\epsilon$ ואם $\beta \rightarrow *\epsilon$

follow המשך דוגמה - חישוב

$$S \rightarrow aB \mid BC \mid CBd$$
 : נתון דקדוק: • $B \rightarrow b \mid \epsilon$
$$C \rightarrow c \mid \epsilon$$

ישוב follow (לכל משתנה דקדוק): •

	S	В	С
1	{\$ }	Ø	Ø
2	{\$ }	{c,\$,d}	{b,d,\$}
3	{\$ }	{c,\$,d}	{b,d,\$}

select המשך דוגמה - חישוב

$$S \rightarrow aB \mid BC \mid CBd$$
 : נתון דקדוק: • $B \rightarrow b \mid \epsilon$
$$C \rightarrow c \mid \epsilon$$

י חישוב קבוצת הטרמינלים select לכל כלל בדקדוק:

```
select(A \rightarrow α) = \begin{cases} first(α) \cup follow(A) & α \rightarrow * ε \\ first(α) & otherwise \end{cases}
```

select המשך דוגמה - חישוב

```
S \rightarrow aB \mid BC \mid CBd
B \rightarrow b \mid \epsilon
C \rightarrow c \mid \epsilon
select(A \rightarrow \alpha) = \begin{cases} first(\alpha) \cup follow(A) & \alpha \rightarrow^* \epsilon \\ first(\alpha) & otherwise \end{cases}
select(A \rightarrow \alpha) = \begin{cases} first(\alpha) \cup follow(A) & \alpha \rightarrow^* \epsilon \\ first(\alpha) & otherwise \end{cases}
```

• חישוב select לכל כלל בדקדוק:

```
select(S→aB) = first(aB) = {a}

select(S→BC) = first(BC)∪follow(S) = {b,c,$}

select(S→CBd) =

select(B→b) =

select(B→\epsilon) =

select(C→\epsilon) =

select(C→\epsilon) =
```

select המשך דוגמה - חישוב

$$S \rightarrow aB \mid BC \mid CBd$$

$$B \rightarrow b \mid \epsilon$$

$$C \rightarrow c \mid \epsilon$$

$$select(A \rightarrow \alpha) = \begin{cases} first(\alpha) \cup follow(A) & \alpha \rightarrow^* \epsilon \\ first(\alpha) & otherwise \end{cases}$$

$$select(A \rightarrow \alpha) = \begin{cases} first(\alpha) \cup follow(A) & \alpha \rightarrow^* \epsilon \\ first(\alpha) & otherwise \end{cases}$$

• חישוב select לכל כלל בדקדוק:

```
= first(aB)
select(S \rightarrow aB)
                                                                            = \{a\}
select(S \rightarrow BC)
                                 = first(BC)∪follow(S)
                                                                            = \{b,c,\$\}
select(S \rightarrow CBd)
                                 = first(CBd)
                                                                            = \{b,c,d\}
select(B \rightarrow b)
                                 = first(b)
                                                                            = \{b\}
select(B \rightarrow \epsilon)
                                 = first(\epsilon)\cupfollow(B)
                                                                            = \{c,d,\$\}
select(C \rightarrow c)
                                 = first(c)
                                                                            = \{C\}
select(C \rightarrow \varepsilon)
                                 = first(\epsilon)\cupfollow(C)
                                                                            = \{b,d,\$\}
30
```

מתי קיים מנתח (1)_L.

 לדקדוק G קיים מנתח (LL(1) אם ורק אם לא קיים קונפליקט בדקדוק, כלומר לכל שני כללים בדקדוק עבור אותו המשתנה- A→β , A→α:

 $select(A \rightarrow \alpha) \cap select(A \rightarrow \beta) = \emptyset$

בניית מנתח (LL(1

מבני נתונים של המנתח

- מחסנית הניתוח שמחזיקה את מה <u>שרוצים עדיין לראות: Q</u>
 - בראש המחסנית מה שרוצים לראות מיד (טרמינל/משתנה)
 - S אתחול: בראש המחסנית יהיה –
 - pop, push, top תומכת בפעולות –
 - שאומרת איזה כלל גזירה יש להפעיל בהינתן המשתנה הבא לגזירה והטרמינל הבא בקלט

$$M(X,t) = \begin{cases} X \to \alpha & t \in select(X \to \alpha) \\ X \in V & error & otherwise \\ t \in T \end{cases}$$

כאשר אין קונפליקטים, כל תא של M מכיל איבר אחד בלבד

• נתון דקדוק הבא:

 $S \rightarrow Ab \mid bC$

 $A \rightarrow a$

 $C \rightarrow cA$

- נבנה טבלת M של מנתח (LL(1).
- select לצורך כך, נחשב תחילה את פונקצייתעבור כל כלל גזירה.

דוגמה (המשך)

$$select(S \rightarrow Ab) = \{a\}$$
 $select(S \rightarrow bC) = \{b\}$
 $select(A \rightarrow a) = \{a\}$
 $select(C \rightarrow cA) = \{c\}$

M:

M:

a
b
c
\$

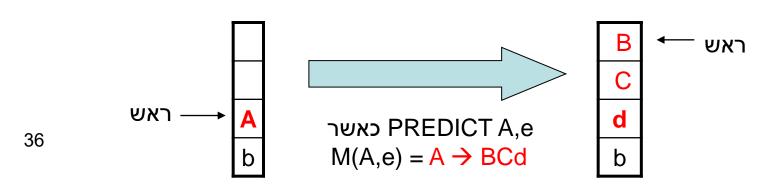
S
S \rightarrow Ab
S \rightarrow bC

A
A \rightarrow a
C
C \rightarrow cA

error המקומות הריקים מכילים

פעולות המנתח

- . הוצאת טרמינלים תואמים מהקלט ומהמחסנית. MATCH t,t אם הם לא תואמים עוצרים עם הודעת שגיאה.
 - בצד ימין PREDICT X,t פרע משתנה X בצד ימין: ימין פרל גזירה מתאים
 - איזה כלל גזירה להפעיל M(X,t)– חפש ב-
 - עצור עם הודעת שגיאה error אם מצאת •
 - אחרת, נניח שזהו כלל $X \to Y_1 Y_2 ... Y_n$. הוצא את X מהמחסנית יניח שזהו כלל ידחוף $Y_1 Y_2 ... Y_n$ למחסנית בסדר הפוך כאשר $Y_1 Y_2 ... Y_n$



אלגוריתם הניתוח

Init: push(Q,S), t is the next terminal in input

- 1. If (isEmpty(Q))
 - if (t == \$), report success // \$ end of input
 - else, report error
- 2. else
 - -X = top(Q)
 - if (X is terminal)MATCH (X,t)
 - if (X is variable)PREDICT (X,t)
- 3. goto 1

	מחסנית		קלט	פעולה
S		bca\$		

```
X = top(Q)
if (X is terminal)

MATCH (X,t)
if (X is variable)

PREDICT (X,t)
```

	а	b	С
S	S → Ab	S → bC	
Α	A→a		
С			C→
			cA

:bca ריצת המנתח על המילה •

	מחסנית		קלט	פעולה
S		bca\$		predict S,b
Cb		bca\$		

ראש המחסנית

```
X = top(Q)
if (X is terminal)

MATCH (X,t)
if (X is variable)

PREDICT (X,t)
```

a b c S S → S → bC ... A A → a ... C C → cA

	מחסנית		קלט	פעולה
S		bca\$		predict S,b
Cb		bca\$		match b,b
С		ca\$		predict C,c

```
X = top(Q)
if (X is terminal)

MATCH (X,t)
if (X is variable)

PREDICT (X,t)
```

a b c S S → S → bC ... A A → a ... C C → cA

	מחסנית		קלט	פעולה
S		bca\$		predict S,b
Cb		bca\$		match b,b
С		ca\$		predict C,c
Ac		ca\$		match c,c

```
X = top(Q)
if (X is terminal)

MATCH (X,t)
if (X is variable)

PREDICT (X,t)
```


חסנית	מ	קלט	פעולה
S	bca\$		predict S,b
Cb	bca\$		match b,b
С	ca\$		predict C,c
Ac	ca\$		match c,c
Α	a\$		predict A,a
а	a\$		match a,a
<empty></empty>	\$		success

שאלה ממבחן - מחסנית

נתון הדקדוק הבא:

 $S \rightarrow AAB$

 $A \rightarrow a$

 $B \rightarrow b$

עבור כל אחת מהאפשרויות הבאות, קבעו האם תוכן המחסנית הנתון בה, יכול להתקבל בריצה של המנתח (LL(1) עבור הדקדוק הנ"ל על מילת קלט בלשהי (שאינה בהכרח בשפה). הסימן הימני ביותר הוא בראש המחסנית:

BA .א ✓

Ba .ב ✓

k. Ad

Аа .т

פתרון שאלה

 $S \rightarrow AAB$

 $A \rightarrow a$

 $B \rightarrow b$

BA .א

Ba .:

bA .a

Аа .т

שלב 1: נבנה את הטבלה:

	a	b
S	S →AAB	
A	$A \rightarrow a$	
В		$B \rightarrow b$

	а	b	DΛ
S	S →AAB		BA Ba
Α	$A \rightarrow a$		bA
В		$B \rightarrow b$	Aa

שלב 2: נבנה לפי הטבלה את המצבים האפשריים במחסנית:

→ BA and Ba are ok ...

סיכום

- שמחקה את פעולת LL(1) הגדרנו מנתח LL(1) באמצעות מחסנית. Recursive Decent
- המנתח הוא Top Down, בכל שלב בניתוח נעזרים בפונקציית select לקבל החלטות.
- A קיים קונפליקט, אם עבור משתנה בדקדוקמתקיים:

 $select(A \rightarrow \alpha) \cap select(A \rightarrow \beta) \neq \emptyset$

בשיעור הבא

• מנתח (1) מתקשה להתמודד עם רקורסיה שמאלית. למשל

$$A \to A\alpha | b$$

אינו ב(1) c: (4) – (4)

 $select(A \rightarrow A\alpha) \cap select(A \rightarrow b) = \{b\}$

Bottom Up בשיעור הבא נדבר על מנתחי
 שמצליחים להתמודד עם בעיה זו.