תורת הקומפילציה

תרגיל בית 4

מגישים:

יונדב קדז 207068973

לירן כהן 209043470

שאלה 1 – Parsing

1. הדקדוק אינו LL(1), יש קונפליקט:

$select(S \to A\ B) = first(A\ B) = \{a, b\}$

$select(A \to a\ A) = first(a\ A) = \{a\}$

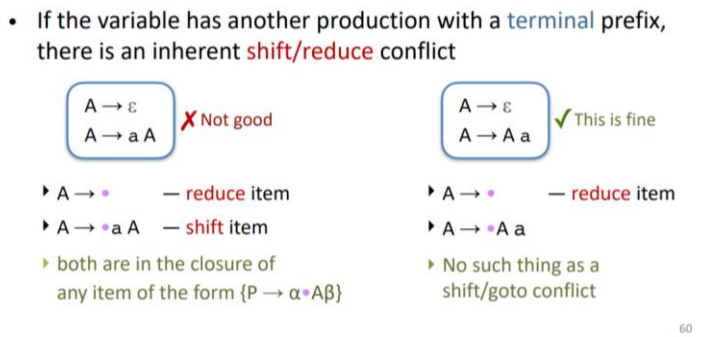
$select(A \to \varepsilon) = first(\varepsilon) \cup follow(A) = \emptyset \cup \{b\} = \{b\}$

$select(B \to b\ a) = first(b\ a) = \{b\}$

$select(B \to b\ B\ a) = first\ (b\ B\ a) = \{b\}$

	a	b	$
S	$S \to A\ B$	$S \to A\ B$	
A	$A \to a\ A$	$A \to \varepsilon$	
B		$B \to b\ a$ $B \to b\ B\ a$	

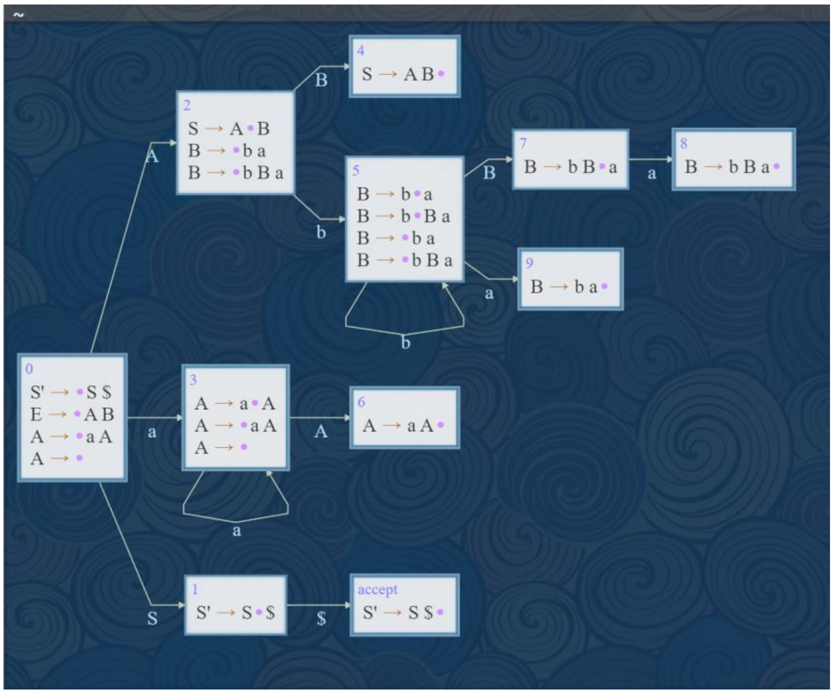
2. הדקדוק אינו LR(0), יש קונפליקט shift-reduce מכלל האפסילון, כפי שנלמד בהרצאה. בכל היתקלות עם המשתנה A, יהיה קונפליקט האם לעשות reduce עבור כלל האפסילון, או shift עבור a:



3. הדקדוק SLR. נראה שאין קונפליקטים ע"י בניית האוטומט והטבלה.
נמספר את הכללים:

1. $S \to A\ B$
2. $A \to\ a\ A$
3. $A \to \varepsilon$
4. $B \to b\ a$
5. $B \to b\ B\ a$

אוטומט:



טבלת מצבים:

$$Follow(S) = \{\$\}\ ;\ Follow(A) = \{b\}\ ;\ Follow(B) = \{a, \$\}$$

	SLR actions			goto		
	a	b	$	S	A	B
0	s3	r3		1	2	
1			acc			
2		s5				4
3	s3	r3			6	
4			r1			
5	s9	s5				7
6		r2				
7	s8					
8	r5		r5			
9	r4		r4			

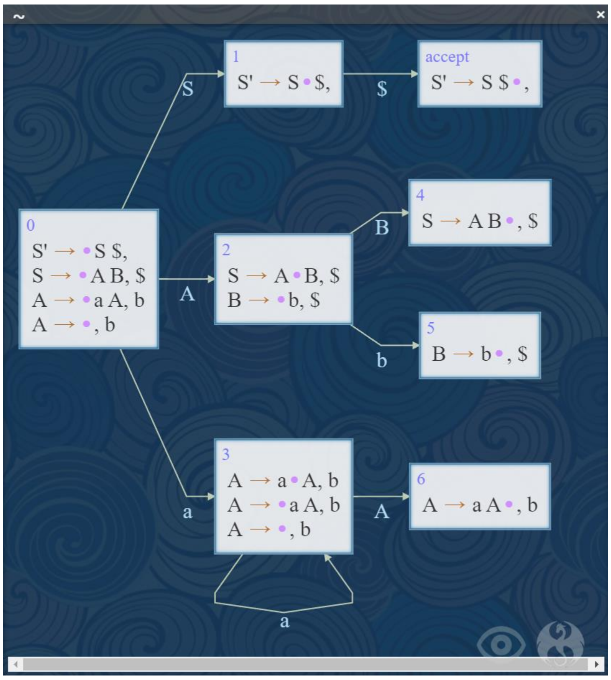
אין קונפליקטים לכן הדקדוק SLR.

ד. כפי שלמדנו בכיתה, כל דקדוקי SLR מוכלים בכל דקדוקי LR(1). בשילוב עם הסעיף הקודם נקבל כי הדקדוק הוא LR(1).

ה.   נמספר את הכללים:

1. $S \rightarrow A\ B$
2. $A \rightarrow\ a\ A$
3. $A \rightarrow \varepsilon$
4. $B \rightarrow b$

אוטומט:



טבלת מצבים:

	LR(1) actions			goto		
	a	b	$	S	A	B
0	s3	r3		1	2	
1			acc			
2		s5				4
3	s3	r3			6	
4			r1			
5			r4			
6		r2				

ו.

קלט	פעולה	מחסנית
ab$	s3	(0,)
b$	r3	(0,)(3,a)
		(0,)(3,a)(goto(3,A),A)
b$	r2	(0,)(3,a)(6,A)
		(0,)(goto(0,A),A)
b$	s5	(0,)(2,A)
$	r4	(0,)(2,A)(5,b)
		(0,)(2,A)(goto(2,B))
$	r1	(0,)(2,A)(4,B)
		(0,)(goto(0,S))
$	acc	(0,)(1,S))

ז.

קלט	פעולה	מחסנית
bb$	r3	(0,)
		(0,)(goto(0,A))
bb$	s5	(0,)(2,A)
b$	error	(0,)(2,A)(5,b)

שאלה 2 – DFA

1. נגדיר:

$$A = set\ of\ all\ ascii\ characters$$
$$N = Vars \times A$$
$$Domain = \mathcal{P}(N)$$
$$\sqcup = \cup$$
$$\sqsubseteq = \subseteq$$
$$\bot = \emptyset$$
$$\top = N$$

2. נגדיר מעברים בעזרת טבלת Kill/Gen. נגדיר בנוסף את הפעולות והקבוצות הבאות:

$$Lower = [a - z]$$
$$Upper = [A - Z]$$
$$upper(x)\ takes\ a\ char\ in\ Lower\ and\ turns\ it\ into\ uppercase$$
$$lower(x)\ takes\ a\ char\ in\ Upper\ and\ turns\ it\ into\ lowercase$$
$$\underline{\boldsymbol{upper\ and\ lower\ return\ the\ same\ char\ if\ it\ isn't\ in\ Upper/Lower}}$$

Statement	Kill(B)	Gen(B)
$x := const\_string$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid a \in const\_string\}$
$x := y$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B)\}$
$x := y + z$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B) \lor (z,a) \in in(B)\}$
$x := y.upper()$	$\{(x,a) \mid a \in A\}$	$\{(x, upper(a)) \mid (y,a) \in in(B)\}$
$x := y.lower()$	$\{(x,a) \mid a \in A\}$	$\{(x, lower(a)) \mid (y,a) \in in(B)\}$
$x.find(y)$	$\emptyset$	$\emptyset$
$x := y * n \mid n \le 0$	$\{(x,a) \mid a \in A\}$	$\emptyset$
$x := y * n \mid n > 0$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B)\}$
$x := y * n \mid n\ is\ unknown$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B)\}$
$x := z.replace(c,y) \mid c\ is\ an\ ascii\ char\ and\ (z,c) \in in(B)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B) \lor (z,a) \in in(B)\}$
$x := z.replace(c,y) \mid c\ is\ an\ ascii\ char\ and\ (y,c) \notin in(B)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B) \lor (z,a) \in in(B)\} \backslash \{(x,c)\}$
$x := z.replace(s,y)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B) \lor (z,a) \in in(B)\}$

נעדיף להשתמש באפשרות השנייה של ה-replace כי כך נדע ש-c לא יכול להיות ב-x.

עבור תתי-חישובים בתוך בלוק נניח לשם הפשטות שהתוצאה שלהם נמצאת ב-in של הבלוק הזה.

3. נמספר את הבלוקים לפי מספרי השורות בהם. (לדוגמה, שורות 2-3 הן בלוק 1). מתקיים:

$in(1) = \{(x,a)|a \in A\}$ *(since x is unknown, it can have every character)*

$out(1) = in(1) - \{(y,a)|a \in A\} \cup \emptyset$

$in(2) = out(1)$

$out(2) = in(2) - \{(y,a)|a \in A\} \cup [\{(y,'s')\} \cup \{(y,a)|(x,a) \in in(2)\}]$

$in(3) = out(1)$

$out(3) = in(3) - \{(y,a)|a \in A\} \cup [\{(y,'s')\} \cup \{(x,a)|(x,a) \in in(3)\}]$

$in(4) = out(2) \cup out(3)$

$out(4) = in(4)$

נריץ את האנליזה:

	$\perp$	$F(\perp)$	$F^2(\perp)$	$F^3(\perp)$	$F^4(\perp)$	$F^5(\perp)$
$in(1)$	$\emptyset$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(1)$	$\emptyset$	$\emptyset$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$in(2)$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(2)$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$
$in(3)$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(3)$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$
$in(4)$	$\emptyset$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(x,a),(y,a)\|a \in A\}$
$out(4)$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$

	$F^6(\perp)$	$F^7(\perp)$
$in(1)$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(1)$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$in(2)$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(2)$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$
$in(3)$	$\{(x,a)\|a \in A\}$	$\{(x,a)\|a \in A\}$
$out(3)$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$
$in(4)$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$
$out(4)$	$\{(x,a),(y,a)\|a \in A\}$	$\{(x,a),(y,a)\|a \in A\}$

4. הרעיון של סת' לא מוכיח את ה-assert בשורה 7, כיוון שלפי הגדרתו הוא רק מראה ש-s **יכול** להיות בתוך y, אבל לא מראה ש-s **חייב** להיות בתוך y תמיד.

5. כעת נגדיר:

$$\sqcup = \cap$$

$$\sqsubseteq = \supseteq$$

$$\bot = N$$

$$\top = \emptyset$$

6. הסמנטיקה תשתנה עבור פעולות ה-replace והכפל, כי עכשיו צריכים לעקוב אחר התווים שבוודאות יהיו לאחר ההחלפה, ולא תווים שאולי יתווספו.
עבור הפעולות האחרות הסמנטיקה תהיה זהה. השינויים יהיו:

Statement	Kill(B)	Gen(B)
$x \coloneqq y * n \mid n \text{ is unknown}$	$\{(x,a) \mid a \in A\}$	$\emptyset$
$x \coloneqq z.replace(c,y) \mid c \text{ is an ascii char and } (z,c) \in in(B)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (y,a) \in in(B)\} \cup \{(x,a) \mid a \neq c, (z,a) \in in(B)\}$
$x \coloneqq z.replace(c,y) \mid c \text{ is an ascii char and } (y,c) \notin in(B)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (z,a) \in in(B)\} \setminus \{(x,c)\}$
$x \coloneqq z.replace(s,y)$	$\{(x,a) \mid a \in A\}$	$\{(x,a) \mid (z,a) \in in(B)\} \setminus \{(x,a) \mid (s,a) \in in(B)\}$

נעדיף להשתמש באפשרות הראשונה של replace כי כך נדע שכל התווים של y חייבים להיות ב-x.

7. כעת פונקציית המעברים תהיה:

$in(1) = \emptyset$ *(since x is unknown, there's no character it must have)*

$out(1) = in(1) - \{(y,a) | a \in A\} \cup \emptyset$

$in(2) = out(1)$

$out(2) = in(2) - \{(y,a) | a \in A\} \cup [\{(y,'s')\} \cup \{(y,a) | (x,a) \in in(2)\}]$

$in(3) = out(1)$

$out(3) = in(3) - \{(y,a) | a \in A\} \cup [\{(y,'s')\} \cup \{(x,a) | (x,a) \in in(3)\}]$

$in(4) = out(2) \cap out(3)$

$out(4) = in(4)$

נריץ את האנליזה:

	$\top$	$F(\top)$	$F^2(\top)$	$F^3(\top)$	$F^4(\top)$
$in(1)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$out(1)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$in(2)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$out(2)$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$
$in(3)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$out(3)$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$
$in(4)$	$\emptyset$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$	$\{(y,'s')\}$
$out(4)$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(y,'s')\}$	$\{(y,'s')\}$

כעת ניתן להוכיח את ה-assert בשורה 7 – אנחנו יודעים שבכניסה לבלוק 4, התו 's' חייב להיות בתוך המחרוזת y ולכן ה-find לא יחזיר 1-.

8. נרצה להשתמש באנליזה של סת' – כיוון שהיא מראה את כל התווים שעלולים להיות בכל מחרוזת (אנליזת may), כל תו שלא יימצא באנליזה עבור מחרוזת כלשהי מובטח לא להיות במחרוזת זו. כלומר, כשנריץ את האנליזה של סת' על bar נוכל לבדוק האם ב-in של הבלוק האחרון (שורה 17) מתקיים ש-(y,'s') שייך לו. אם לא, נדע שאכן אין את האות s במחרוזת y בשלב זה ונוכיח שה- assert מתקיים.