Αρχές Γλωσσών Προγραμματισμού και Μεταφραστών ΣΕΤ ΑΣΚΗΣΕΩΝ #4 ΚΑΙ ΛΥΣΕΙΣ (V1.1) Συντακτική Ανάλυση Bottom-Up

Γκόγκος Χρήστος

17/12/2019

Δώστε ανοδικές συντακτικές αναλύσεις για:

- 1. την είσοδο 000111 και τη γραμματική $S \rightarrow 0S1|01.$
- 2. την είσοδο aaa*a++ και τη γραμματική $S\to SS+|SS*|a.$

Λύση

Υποερώτημα 1

stack	input	action
\$	000111\$	shift
\$0	00111\$	shift
\$00	0111\$	shift
\$000	111\$	shift
\$0001	11\$	reduce: $S \rightarrow 01$
\$00S	11\$	shift
\$00S1	1\$	reduce: $S \rightarrow 0S1$
\$0S	1\$	shift
\$0S1	\$	reduce: $S \rightarrow 0S1$
\$S	\$	accept

Πίνακας 1: Ανοδική συντακτική ανάλυση

Υποερώτημα 2

stack	input	action
\$	aaa*a++\$	shift
\$a	aa*a++\$	reduce: $S \rightarrow a$
\$S	aa*a++\$	shift
\$Sa	a*a++\$	reduce $S \to a$
\$SS	a*a++\$	shift
\$SSa	*a++\$	$\text{reduce } S \to a$
\$SSS	*a++\$	shift
\$SSS*	a++\$	$\operatorname{reduce} S \to SS*$
\$SS	a++\$	shift
\$SSa	++\$	reduce $S \to a$
\$SSS	++\$	shift
\$SSS+	+\$	reduce $S \to SS+$
\$SS	+\$	shift
\$SS+	\$	$\text{reduce } S \to SS +$
\$S	\$	accept

Πίνακας 2: Ανοδική συντακτική ανάλυση

Ερώτημα 2

Για την ακόλουθη επαυξημένη γραμματική.

$$E' \to E$$

$$E \rightarrow E + T|T$$

$$T \to T * F|F$$

$$F \rightarrow (E)|id$$

- 1. Να υπολογιστεί το κλείσιμο CLOSURE(I) θεωρώντας ως σύνολο στοιχείων I το $\{E' \to \bullet E\}$.
- 2. Να υπολογιστεί το GOTO(I, +) θεωρώντας ως σύνολο στοιχείων I το $\{E' \to E \bullet, E \to E \bullet + T\}$.

Λύση

Υποερώτημα 1

To CLOSURE(I) είναι:

$$E' \to \bullet E$$

$$E \to \bullet E + T$$

$$E \to \bullet T$$

$$T \to \bullet T * F$$

$$T \to \bullet F$$

$$F \to \bullet(E)$$

$$F \rightarrow \bullet id$$

Υποερώτημα 2

Για να υπολογιστεί το GOTO(I,+) αναζητείται στοιχείο του I που να έχει την τελεία ακριβώς πριν από το σύμβολο +. Τέτοιο στοιχείο είναι το $E \to E \bullet + T$. Η τελεία "περνά" στα δεξιά του συμβόλου + οπότε προκύπτει το στοιχείο $E \to E + \bullet T$.

Για το στοιχείο αυτό υπολογίζεται η κλειστότητα.

$$E \to E + \bullet T$$

$$T \to \bullet T * F$$

 $T \to \bullet F$

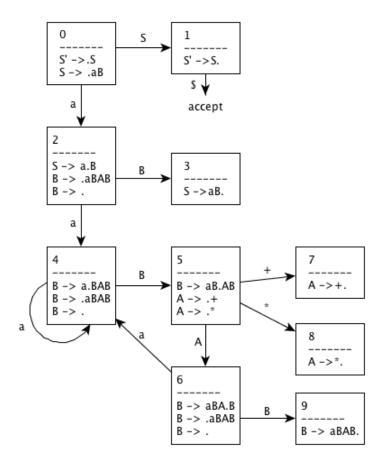
$$F \to \bullet(E)$$

 $F \to \bullet id$

Ερώτημα 3

Δίνεται η ακόλουθη γραμματική και το LR(0) αυτόματο της. Να κατασκευαστεί ο αντίστοιχος SLR πίνακας.

- $0. S' \rightarrow S$
- 1. $S \rightarrow aB$
- 2. $B \rightarrow aBAB$
- 3. $B \rightarrow \epsilon$
- 4. $A \rightarrow +$
- 5. $A \rightarrow *$



Σχήμα 1: LR(0) αυτόματο

 Δ ίνεται ότι FOLLOW(S) = [\$], FOLLOW(A) = [a, +, *, \$] και FOLLOW(B) = [+, *, \$].

Λύση

			ACT	ION	GO'	ТО	
	a	+	*	\$	S	A	В
0	s2				g1		
1				acc			
2	s4	r3	r3	r3			g3
3				r1			
4	s4	r3	r3	r3			g5
5		s7	s8			g6	
6	s4	r3	r3	r3			g9
7	r4	r4	r4	r4			
8	r5	r5	r5	r5			
9		r2	r2	r2			

Πίνακας 3: Πίνακας συντακτικής ανάλυσης SLR

Κατασκευάστε το LR(0) αυτόματο για την ακόλουθη γραμματική.

 $S \to aB$

 $B \to b$

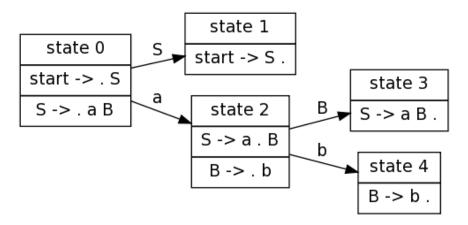
Στη συνέχεια κατασκευάστε τον SLR πίνακα συντακτικής ανάλυσης και πραγματοποιήστε συντακτική ανάλυση στη συμβολοσειρά ab.

Λύση

Η επαυξημένη γραμματική είναι η ακόλουθη:

- 0. start $\rightarrow S$
- 1. $S \rightarrow aB$
- 2. $B \rightarrow b$

Το LR(0) αυτόματο για τη γραμματική παρουσιάζεται στο σχήμα 2.



Σχήμα 2: LR(0) αυτόματο για τη γραμματική G

Tα follow sets είναι: $FOLLOW(start)=\{\$\}$, $FOLLOW(S)=\{\$\}$, $FOLLOW(B)=\{\$\}$. Ο πίνακας συντακτικής ανάλυσης είναι ο πίνακας 4.

	AC'	TION	GOTO		
	a	b	\$	S	В
0	s2			g1	
1			acc		
2		s4			g3
3			r1		
4			r2		

Πίνακας 4: Πίνακας συντακτικής ανάλυσης SLR

Η συντακτική ανάλυση της συμβολοσειράς ab παρουσιάζεται στον πίνακα 5.

stack	input	action
0	ab\$	s2
0a2	b\$	s4
0a2b4	\$	$r2:B \rightarrow b$
04204	Ф	g3
0a2B3	\$	$r1:S \rightarrow aB$
VaZDS	Ф	g1
0S1	\$	accept

Πίνακας 5: Συντακτική ανάλυση της συμβολοσειράς ab

Κατασκευάστε το LR(0) αυτόματο για την ακόλουθη γραμματική.

 $S \to AA$

 $A \rightarrow aA$

 $A \rightarrow b$

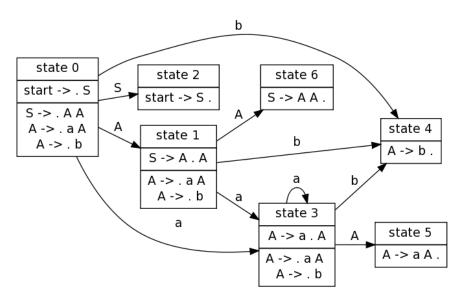
Στη συνέχεια κατασκευάστε τον SLR πίνακα συντακτικής ανάλυσης και πραγματοποιήστε συντακτική ανάλυση στη συμβολοσειρά aabb.

Λύση

Η επαυξημένη γραμματική είναι η ακόλουθη:

- 0. start $\rightarrow S$
- 1. $S \rightarrow AA$
- 2. $A \rightarrow aA$
- 3. $A \rightarrow b$

Το LR(0) αυτόματο για τη γραμματική παρουσιάζεται στο σχήμα 3.



Σχήμα 3: LR(0) αυτόματο για τη γραμματική G

Τα follow sets είναι: $FOLLOW(start) = \{\$\}$, $FOLLOW(S) = \{\$\}$, $FOLLOW(A) = \{a,b,\$\}$. Ο πίνακας συντακτικής ανάλυσης είναι ο πίνακας 6.

		acti	goto		
state	a	b	\$	A	S
0	s3	s4		g1	g2
1	s3	s4		g6	
2			accept		
3	s3	s4		g5	
4	r3	r3	r3		
5	r2	r2	r2		
6			r1		

Πίνακας 6: Πίνακας συντακτικής ανάλυσης SLR

Η συντακτική ανάλυση της συμβολοσειράς aabb παρουσιάζεται στον πίνακα 7.

stack	input	action
0	aabb\$	s3
0a3	abb\$	s3
0a3a3	bb\$	s4
0a3a3b4	b\$	$r3:A \rightarrow b$
0a3a30 4	U	g5
0a3a3A5	b\$	$r2:A \rightarrow aA$
VasasAs	Оф	g5
0a3A5	b\$	$r2:A \rightarrow aA$
043713	ОФ	g1
0A1	b\$	s4
0A1b4	\$	$r3:A \rightarrow b$
0/1104	Ψ	g6
0A1A6	\$	$r1:S \to AA$
0/11/10	Ψ	g2
0S2	\$	accept

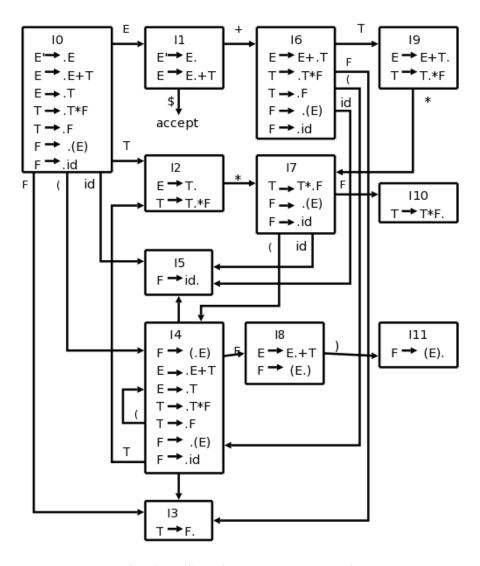
Πίνακας 7: Πίνακας συντακτικής ανάλυσης SLR

Ερώτημα 6

Για τη γραμματική εκφράσεων G, δίνεται το LR(0) αυτόματο του σχήματος 4.

Η γραμματική εκφράσεων G (σε επαυξημένη μορφή)

- 0. $E' \rightarrow E$
- 1. $E \rightarrow E + T$
- 2. $E \rightarrow T$
- 3. $T \rightarrow T * F$
- 4. $T \rightarrow F$
- 5. $F \rightarrow (E)$
- 6. $F \rightarrow id$



Σχήμα 4: LR(0) αυτόματο για τη γραμματική G

- 1. Πόσα είναι τα στοιχεία (items) που αντιστοιχούν στη γραμματική G;
- 2. Ποιές δράσεις προκαλούν οι ακόλουθες ενέργειες; Δίνεται ότι το FOLLOW(E)={\$, +, }}.
 - (a) ACTION[I0,(]
 - (b) ACTION[I0,id]
 - (c) ACTION[I1,\$]
 - (d) ACTION[I1,+]
 - (e) ACTION[I2,\$]
 - (f) ACTION[12,+]
 - (g) ACTION[I2,)]
 - (h) ACTION[I2,*]
- 3. Πραγματοποιήστε SLR συντακτική ανάλυση για τη συμβολοσειρά id * id, χρησιμοποιώντας το LR(0) αυτόματο και τη γραμματική G.
- 4. Τι θα σήμαινε το να βρεθούμε στην κατάσταση Ι1 και το \$ να είναι το επόμενο σύμβολο εισόδου;
- 5. Τι θα σήμαινε το να βρεθούμε στην κατάσταση Ι7 και το + να είναι το επόμενο σύμβολο εισόδου;

Λύση

- 1. Τα στοιχεία είναι 20.
- 2. (a) shift 4
 - (b) shift 5
 - (c) accept
 - (d) shift 6
 - (e) reduce $E \to T$ kai goto 0
 - (f) error
 - (g) error
 - (h) shift 7
- 3. Η συντακτική ανάλυση για τη συμβολοσειρά id * id παρουσιάζεται στον πίνακα 8.

stack	input	action
0	id*id\$	shift 5
0id5	*id\$	$reduce\: F \to id$
olus	Ius	goto 3
0F3	*id\$	$\mathrm{reduce}\ T \to F$
01.3	щ	goto 2
0T2	*id\$	shift 7
0T2*7	id\$	shift 5
0T2*7id5	\$	$\operatorname{reduce} F \to \operatorname{id}$
012 /103	, p	goto 10
0T2*7F10	\$	$\operatorname{reduce} T \to T * F$
012 /110	, p	goto 2
0T2	\$	$\operatorname{reduce} E \to T$
012	D D	goto 1
0E1	\$	accept

Πίνακας 8: Συντακτική ανάλυση της εισόδου id*id

- 4. Θα σήμαινε ότι ο συντακτικός αναλυτής έχει φτάσει σε κατάσταση accept και ότι η συμβολοσειρά είναι έγκυρη συντακτικά.
- 5. Θα σήμαινε ότι ο συντακτικός αναλυτής έχει φτάσει σε κατάσταση error και ότι η συμβολοσειρά δεν είναι έγκυρη συντακτικά.

Ερώτημα 7

Για το LR(0) αυτόματο και τη γραμματική του προηγούμενου ερωτήματος.

- 1. Κατασκευάστε τον πίνακα συντακτικής ανάλυσης SLR. Δίνεται ότι $FOLLOW(E')=\{\$\}$, $FOLLOW(E)=\{+,)$, $\$\}$, $FOLLOW(T)=\{*, +,)$, $\$\}$.
- 2. Πραγματοποιήστε συντακτική ανάλυση για τη συμβολοσειρά (id + id) * id.
- 3. Τι θα σήμαινε εαν ο πίνακας συντακτικής ανάλυσης SLR σε κάποια από τα κελιά του είχε 2 ή περισσότερες εγγραφές;

Λ ύση

1. Ο πίνακας συντακτικής ανάλυσης είναι ο πίνακας 9.

state			:	action	1		goto		
	id	+	*	()	\$	E	T	F
0	s5			s4			g1	g2	g3
1		s6				accept			
2		r2	s7		r2	r2			
3		r4	r4		r4	r4			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r1	s7		r1	r1			
10		r3	r3		r3	r3			
11		r5	r5		r5	r5			

Πίνακας 9: Πίνακας συντακτικής ανάλυσης SLR

2. Η συντακτική ανάλυση για τη συμβολοσειρά (id+id)*id παρουσιάζεται στον πίνακα 10.

stack	input	action
0	(id+id)*id\$	s4
0(4	id+id)*id\$	s5
0(4id5	+id)*id\$	r6: $F \rightarrow \text{id}$ και g3
0(4F3	+id)*id\$	r4: $T o F$ και g2
0(4T2	+id)*id\$	$\mathbf{r2}$: $E \rightarrow T$ και $\mathbf{g8}$
0(4E8	+id)*id\$	s6
0(4E8+6	id)*id\$	s5
0(4E8+6id5)*id\$	r6: $F o id$ και g3
0(4E8+6F3)*id\$	r4: $T \rightarrow F$ και g9
0(4E8+6T9)*id\$	$r1: E \rightarrow E + T$ και g8
0(4E8)*id\$	s11
0(4E8)11	*id\$	$r5: F \rightarrow (E)$ και $g3$
0F3	*id\$	r4: $T o F$ και g2
0T2	*id\$	s7
0T2*7	id\$	s5
0T2*7id5	\$	r6: $F \rightarrow \text{id}$ και g10
0T2*7F10	\$	$r3: T \rightarrow T * F$ και $g2$
0T2	\$	accept

Πίνακας 10: Συντακτική ανάλυση της εισόδου $(\mathrm{id}+\mathrm{id})*\mathrm{id}$

3. Θα σήμαινε ότι η γραμματική δεν είναι SLR(1).

Δίνεται η ακόλουθη γραμματική:

 $T \to R$

 $T \to aTc$

 $R \to \epsilon$

 $R \to bR$

Απαντήστε στα ακόλουθα υποερωτήματα:

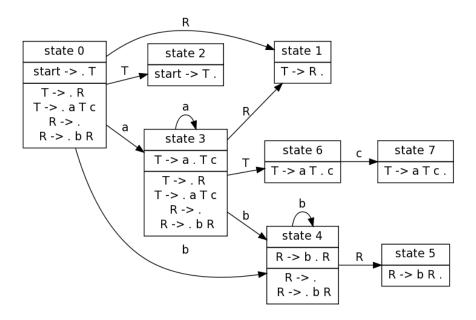
- 1. Υπολογίστε τα FIRST SETS και FOLLOW SETS των συμβόλων της γραμματικής.
- 2. Γράψτε την αντίστοιχη επαυξημένη γραμματική.
- 3. Σχεδιάστε το LR(0) αυτόματο της γραμματικής.
- 4. Συμπληρώστε τον πίνακα ελέγχου συντακτικής ανάλυσης SLR(1).
- 5. Χρησιμοποιώντας τον πίνακα ελέγχου SLR(1) πραγματοποιήστε συντακτική ανάλυση για την είσοδο abbc.
- 6. Σχεδιάστε το δένδρο συντακτικής ανάλυσης και καταγράψτε τη δεξιότερη παραγωγή που οδηγεί από το αρχικό σύμβολο στη συμβολοσειρά εισόδου abbc.

Λύση

1. Υπολογισμός των FIRST SETS και FOLLOW SETS

	FIRST	FOLLOW
a	$\{a\}$	
b	$\{b\}$	
c	$\{c\}$	
T	$\{a,b,\epsilon\}$	$\{\$, c\}$
R	$\{b,\epsilon\}$	$\{\$, c\}$

- 2. Επαυξημένη γραμματική
 - 0. $S \rightarrow T$ \$
 - 1. $T \rightarrow R$
 - 2. $T \rightarrow aTc$
 - 3. $R \rightarrow \epsilon$
 - 4. $R \rightarrow bR$
- 3. LR(0) για τη γραμματική



Σχήμα 5: LR(0) αυτόματο για τη γραμματική G

Η επαλήθευση της ορθότητας του LR(0) αυτομάτου μπορεί να γίνει στην ιστοσελίδα:

http://smlweb.cpsc.ucalgary.ca/start.html

Η γραμματική θα πρέπει να εισαχθεί με την ακόλουθη μορφή:

 $T \rightarrow R$.

 $T \rightarrow a T c$.

R -> .

 $R \rightarrow b R$.

4. Πίνακας ελέγχου SLR(1)

		AC	GO	ГО		
	a	b	С	\$	T	R
0	s3	s4	r3	r3	g2	g1
1			r1	r1		
2				accept		
3	s3	s4	r3	r3	g6	g1
4		s4	r3	r3		g5
5			r4	r4		
6			s7			
7			r2	r2		

Πίνακας 11: Πίνακας ελέγχου SLR(1)

Περιγραφή δημιουργίας περιεχομένων της γραμμής 0 στον πίνακα ελέγχου SLR(1)

- Από την κατάσταση 0 η μετάβαση με το τερματικό σύμβολο a οδηγεί στην κατάσταση 3, συνεπώς στη θέση του πίνακα [0,a] εισάγεται η τιμή s3.
- Από την κατάσταση 0 η μετάβαση με το τερματικό σύμβολο b οδηγεί στην κατάσταση 4, συνεπώς στη θέση του πίνακα [0,b] εισάγεται η τιμή s4.

- Από την κατάσταση 0 η μετάβαση με το μη τερματικό σύμβολο T οδηγεί στην κατάσταση 2, συνεπώς στη θέση του πίνακα [0,T] εισάγεται η τιμή g2.
- Από την κατάσταση 0 η μετάβαση με το μη τερματικό σύμβολο R οδηγεί στην κατάσταση 1, συνεπώς στη θέση του πίνακα [0,R] εισάγεται η τιμή g1.
- Στην κατάσταση 0 υπάρχει το στοιχείο $R \to \bullet$ που αντιστοιχεί στην παραγωγή 3 και στο οποίο το σύμβολο \bullet βρίσκεται στη δεξιότερη θέση του στοιχείου. Το Follow Set του R είναι το $\{\$, c\}$, άρα στις θέσεις του πίνακα [0,c] και [0,\$] εισάγεται η τιμή r3.
- 5. Συντακτική ανάλυση της εισόδου abbc.

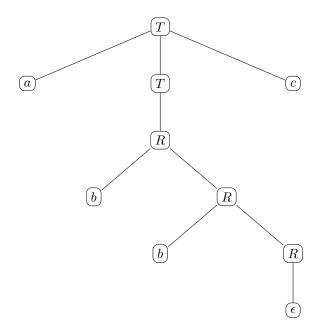
Αριθμός γραμμής	stack	input	action
1	0	abbc\$	s3
2	0a3	bbc\$	s4
3	0a3b4	bc\$	s4
4	0a3b4b4	c\$	r3: $R \to \epsilon$ g5
5	0a3b4b4R5	c\$	r4: $R \rightarrow bR$ g5
6	0a3b4R5	c\$	r4: $R \rightarrow bR$ g1
7	0a3R1	c\$	r1: $T \rightarrow R$ g6
8	0a3T6	c\$	s7
9	0a3T6c7	\$	$r2: T \to aTc$ $g2$
10	0T2	\$	accept

Περιγραφή συμπλήρωσης των περιεχομένων του πίνακα συντακτικού ελέγχου:

- Γραμμή 1: Αρχικά τοποθετείται στη στοίβα η κατάσταση 0 και εξετάζεται το πλέον αριστερό στοιχείο της εισόδου, δηλαδή το a. Ο συνδυασμός [0,a] δίνει ως ενέργεια που θα πραγματοποιηθεί την s3, δηλαδή ωθούνται στη στοίβα το σύμβολο a και η κατάσταση 3 (νέο περιεχόμενο στοίβας: 0a3).
- Γραμμή 2: Εξετάζεται ο συνδυασμός [3,b] που δίνει ως ενέργεια που θα πραγματοποιηθεί την s4, δηλαδή ωθούνται στη στοίβα το σύμβολο b και η κατάσταση 4 (νέο περιεχόμενο στοίβας: 0a3b4).
- Γραμμή 3: Εξετάζεται ο συνδυασμός [4,b] που δίνει ως ενέργεια που θα πραγματοποιηθεί την s4, δηλαδή ωθούνται στη στοίβα το σύμβολο b και η κατάσταση 4 (νέο περιεχόμενο στοίβας: 0a3b4b4).
- Γραμμή 4: Εξετάζεται ο συνδυασμός [4,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την r3, δηλαδή γίνεται ελάττωση με βάση τον κανόνα παραγωγής 3 $(R \to \epsilon)$ που δεν έχει σύμβολα στο δεξί του μέρος (συνεπώς δεν απωθούνται σύμβολα από τη στοίβα) και ο συνδυασμός του τερματικού συμβόλου R με την κατάσταση που βρίσκεται στην κορυφή της στοίβας (δηλαδή την 4) καθορίζει ότι ωθούνται στη στοίβα το σύμβολο R και η κατάσταση 5 (νέο περιεχόμενο στοίβας: 0a3b4b4R5).
- Γραμμή 5: Εξετάζεται ο συνδυασμός [5,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την r4, δηλαδή γίνεται ελάττωση με βάση τον κανόνα παραγωγής 4 $(R \to bR)$ που έχει 2 σύμβολα στο δεξί του μέρος (συνεπώς απωθούνται 4 σύμβολα από τη στοίβα) και ο συνδυασμός του τερματικού συμβόλου R με την κατάσταση που πλέον βρίσκεται στην κορυφή της στοίβας (δηλαδή την 4) καθορίζει ότι ωθούνται στη στοίβα το σύμβολο R και η κατάσταση 5 (νέο περιεχόμενο στοίβας: 0a3b4R5).
- Γραμμή 6: Εξετάζεται ο συνδυασμός [5,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την r4, δηλαδή γίνεται ελάττωση με βάση τον κανόνα παραγωγής 4 $(R \to bR)$ που έχει 2 σύμβολα στο δεξί του μέρος (συνεπώς απωθούνται 4 σύμβολα από τη στοίβα) και ο συνδυασμός του τερματικού συμβόλου R με την

κατάσταση που πλέον βρίσκεται στην κορυφή της στοίβας (δηλαδή την 3) καθορίζει ότι ωθούνται στη στοίβα το σύμβολο R και η κατάσταση 1 (νέο περιεχόμενο στοίβας: 0a3R1).

- Γραμμή 7: Εξετάζεται ο συνδυασμός [1,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την r1, δηλαδή γίνεται ελάττωση με βάση τον κανόνα παραγωγής 4 $(T \to R)$ που έχει 1 σύμβολο στο δεξί του μέρος (συνεπώς απωθούνται 2 σύμβολα από τη στοίβα) και ο συνδυασμός του τερματικού συμβόλου T με την κατάσταση που πλέον βρίσκεται στην κορυφή της στοίβας (δηλαδή την 3) καθορίζει ότι ωθούνται στη στοίβα το σύμβολο T και η κατάσταση 6 (νέο περιεχόμενο στοίβας: 0a3T6).
- Γραμμή 8: Εξετάζεται ο συνδυασμός [6,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την s7, δηλαδή ωθούνται στη στοίβα το σύμβολο b και η κατάσταση 7 (νέο περιεχόμενο στοίβας: 0a3T6c7).
- Γραμμή 9: Εξετάζεται ο συνδυασμός [7,\$] που δίνει ως ενέργεια που θα πραγματοποιηθεί την r2, δηλαδή γίνεται ελάττωση με βάση τον κανόνα παραγωγής 2 ($T \to aTc$) που έχει 3 σύμβολα στο δεξί του μέρος (συνεπώς απωθούνται 6 σύμβολα από τη στοίβα) και ο συνδυασμός του τερματικού συμβόλου T με την κατάσταση που πλέον βρίσκεται στην κορυφή της στοίβας (δηλαδή την 0) καθορίζει ότι ωθούνται στη στοίβα το σύμβολο T και η κατάσταση 2 (νέο περιεχόμενο στοίβας: 0T2).
- Γραμμή 10: Εξετάζεται ο συνδυασμός [2,c] που δίνει ως ενέργεια που θα πραγματοποιηθεί την accept. Άρα η συμβολοσειρά έχει αναγνωριστεί επιτυχώς.
- 6. Δένδρο συντακτικής ανάλυσης και δεξιότερη παραγωγή από το αρχικό σύμβολο της γραμματικής στη συμβολοσειρά εισόδου.



$$T \xrightarrow{\operatorname{rm}^*} abbc$$

$$T \to aTc \to aRc \to abRc \to abbRc \to abbc$$