

Μεταγλωττιστές

Ανοδικές Μέθοδοι Συντακτικής Ανάλυσης

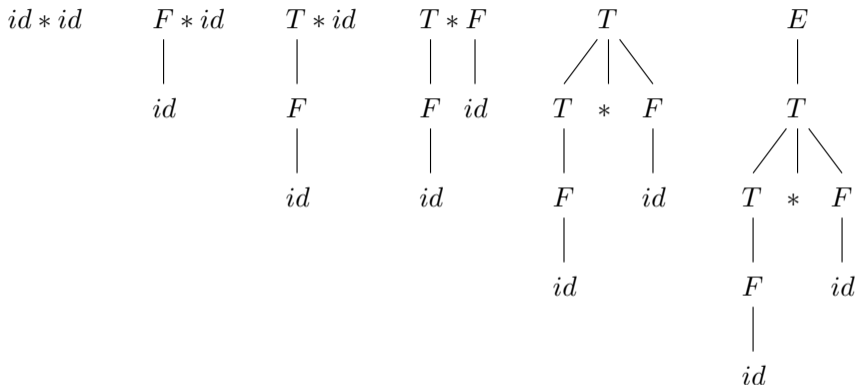
Δημήτρης Μιχαήλ



Τμήμα Πληροφορικής και Τηλεματικής
Χαροκόπειο Πανεπιστήμιο

Ανοδική Κατασκευή Συντακτικού Δέντρου

κατασκευή δέντρου που ξεκινάει από τα φύλλα προς την ρίζα



- θα δούμε μια γενική μορφή που ονομάζεται "shift-reduce"
- η κατασκευή τέτοιων συντακτικών αναλυτών γίνεται συνήθως αυτόματα
- είναι απαραίτητο όμως να ξέρουμε πως λειτουργούν ώστε να μπορούμε να γράφουμε κατάλληλες γραμματικές

LR Κατασκευή Συντακτικού Δέντρου

Η αδυναμία των $LL(k)$ τεχνικών είναι πως πρέπει να αποφασίσουν πιο κανόνα παραγωγής να ακολουθήσουν έχοντας δει μόνο τα πρώτα k σύμβολα του δεξιού μέρους του κανόνα.

Μία πιο αποτελεσματική στρατηγική, $LR(k)$, αναβάλλει την απόφαση μέχρι να έχει δει όλα τα σύμβολα του δεξιού μέρους και k σύμβολα επιπλέον.

$LR(k)$

- Left-to-right parse
- Rightmost-derivation in reverse
- k -token lookahead

Η ανοδική κατασκευή συντακτικού δέντρου είναι η διαδικασία "ελάττωσης" μιας συμβολοσειράς w στο αρχικό σύμβολο μιας γραμματικής.

Σε κάθε βήμα "ελάττωσης", ένα πρόθεμα της εισόδου που ταιριάζει με την δεξιά μεριά ενός κανόνας παραγωγής αντικαθίσταται με το μη-τερματικό σύμβολο που είναι στην αριστερή πλευρά του κανόνα.

Η ελάττωση είναι ουσιαστικά η αντίθετη λειτουργία από την παραγωγή.

Η μέθοδος "shift-reduce" είναι μια μορφή της ανοδικής συντακτικής ανάλυσης όπου

- μια στοίβα κρατά σύμβολα της γραμματικής

Κατά την διάρκεια ενός αριστερά-προς-δεξιά περάσματος της εισόδου

- κάνουμε ολίσθηση (shift) ενός ή περισσότερων συμβόλων στην στοίβα
- κάνουμε ελάττωση (reduce) με βάση μια παραγωγή, αφαιρώντας όλα τα σύμβολα της δεξιάς πλευράς από την στοίβα και προσθέτοντας το σύμβολο της αριστερής.

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift
\$ T^*	id \$	shift

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift
\$ T^*	id \$	shift
\$ T^* id	\$	reduce by $F \rightarrow \mathbf{id}$

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift
\$ T^*	id \$	shift
\$ T^* id	\$	reduce by $F \rightarrow \mathbf{id}$
\$ T^* F	\$	reduce by $T \rightarrow T^* F$

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift
\$ T^*	id \$	shift
\$ T^* id	\$	reduce by $F \rightarrow \mathbf{id}$
\$ T^* F	\$	reduce by $T \rightarrow T^* F$
\$ T	\$	reduce by $E \rightarrow T$

Ολίσθηση-Ελάττωση

Shift-Reduce

E	\Rightarrow	$E + T \mid T$
T	\Rightarrow	$T * F \mid F$
F	\Rightarrow	$(E) \mid \mathbf{id}$

STACK	INPUT	ACTION
\$	id * id \$	shift
\$ id	* id \$	reduce by $F \rightarrow \mathbf{id}$
\$ F	* id \$	reduce by $T \rightarrow F$
\$ T	* id \$	shift
\$ T^*	id \$	shift
\$ T^* id	\$	reduce by $F \rightarrow \mathbf{id}$
\$ T^* F	\$	reduce by $T \rightarrow T^* F$
\$ T	\$	reduce by $E \rightarrow T$
\$ E	\$	accept

4 δυνατές λειτουργίες:

- 1 *Shift*. Προσθήκη του επόμενο συμβόλου εισοδου στην κορυφή της στοίβας
- 2 *Reduce*. Επιλογή κανόνα $X \rightarrow A B C$, αφαίρεση των C , B και A από την κορυφή της στοίβας, προσθήκη του X στην στοίβα
- 3 *Accept*. Επιτυχής συντακτική ανάλυση
- 4 *Error*. Ανακάλυψη λάθους και κλήση συνάρτησης διαχείρισης λαθών

Κατά την διάρκεια συντακτικής ανάλυσης ολίσθησης-ελάττωσης μπορεί να προκύψουν δύο προβλήματα:

1 σύγκρουση ελάττωσης-ελάττωσης (reduce-reduce conflict)

Σε μια κατάσταση του συντακτικού αναλυτή είναι δυνατή η ελάττωση με παραπάνω από ένα κανόνες.

Κατά την διάρκεια συντακτικής ανάλυσης ολίσθησης-ελάττωσης μπορεί να προκύψουν δύο προβλήματα:

1 σύγκρουση ελάττωσης-ελάττωσης (reduce-reduce conflict)

Σε μια κατάσταση του συντακτικού αναλυτή είναι δυνατή η ελάττωση με παραπάνω από ένα κανόνες.

2 σύγκρουση ολίσθησης-ελάττωσης (shift-reduce conflict)

Σε μια κατάσταση του συντακτικού αναλυτή είναι δυνατή είτε η ελάττωση με βάση κάποιο κανόνα είτε η ολίσθηση του επόμενου συμβόλου της εισόδου στην στοίβα.

Κατά την διάρκεια συντακτικής ανάλυσης ολίσθησης-ελάττωσης μπορεί να προκύψουν δύο προβλήματα:

1 σύγκρουση ελάττωσης-ελάττωσης (reduce-reduce conflict)

Σε μια κατάσταση του συντακτικού αναλυτή είναι δυνατή η ελάττωση με παραπάνω από ένα κανόνες.

2 σύγκρουση ολίσθησης-ελάττωσης (shift-reduce conflict)

Σε μια κατάσταση του συντακτικού αναλυτή είναι δυνατή είτε η ελάττωση με βάση κάποιο κανόνα είτε η ολίσθηση του επόμενου συμβόλου της εισόδου στην στοίβα.

Στις περιπτώσεις αυτές ο συντακτικός αναλυτής δεν ξέρει ποια απόφαση να πάρει.

Επίσης θέλουμε να αποφύγουμε την χρήση οπισθοδρόμησης (backtracking) αφού είναι ιδιαίτερα δαπανηρή.

Συγκρούσεις

Shift/Reduce

Μια διαφορούμενη γραμματική δεν μπορεί να είναι $LR(1)$.

$stmt$	\rightarrow	if $expr$ then $stmt$ else $stmt$
		if $expr$ then $stmt$
		other

Ένας συντακτικός αναλυτής ελάττωσης-ολίσθησης στην κατάσταση

STACK	INPUT
... if $expr$ then $stmt$	else ... \$

ανάλογα με τι ακολουθεί το **else** στην είσοδο πρέπει να κάνει

- 1 ελάττωση με τον κανόνα $stmt \rightarrow \mathbf{if\ expr\ then\ stmt}$
- 2 ολίσθηση του **else**

επειδή έχει όμως ένα (1) token lookahead υπάρχει shift/reduce σύγκρουση.

Έστω η παρακάτω γραμματική μιας γλώσσας όπου και οι συναρτήσεις αλλά και οι πίνακες χρησιμοποιούν παρενθέσεις και έστω πως ο λεκτικός αναλυτής επιστρέφει την λεκτική μονάδα **id** για όλα τα αναγνωριστικά.

<i>stmt</i>	→	id (<i>parameter_list</i>)
<i>stmt</i>	→	<i>expr</i> := <i>expr</i>
<i>parameter_list</i>	→	<i>parameter_list</i> , <i>parameter</i> <i>parameter</i>
<i>parameter</i>	→	id
<i>expr</i>	→	id
<i>expr</i>	→	id (<i>expr_list</i>)
<i>expr_list</i>	→	<i>expr_list</i> , <i>expr</i> <i>expr</i>

STACK

... **id** (**id**

INPUT

, **id**) ... \$

σίγουρα πρέπει να γίνει ελάττωση, αλλά με ποιόν κανόνα; Αυτή είναι μια σύγκρουση reduce/reduce.

Συντακτικοί Αναλυτές $LR(k)$

LR Parsing

Η ονομασία $LR(k)$ προκύπτει από τα παρακάτω:

- scanning input from **L**eft to right
- **R**eftmost derivation in reverse
- using **k** input symbol of lookahead at each step to make parsing action decisions

Με LR συμβολίζουμε την περίπτωση $k = 1$, δηλαδή $LR(1)$.

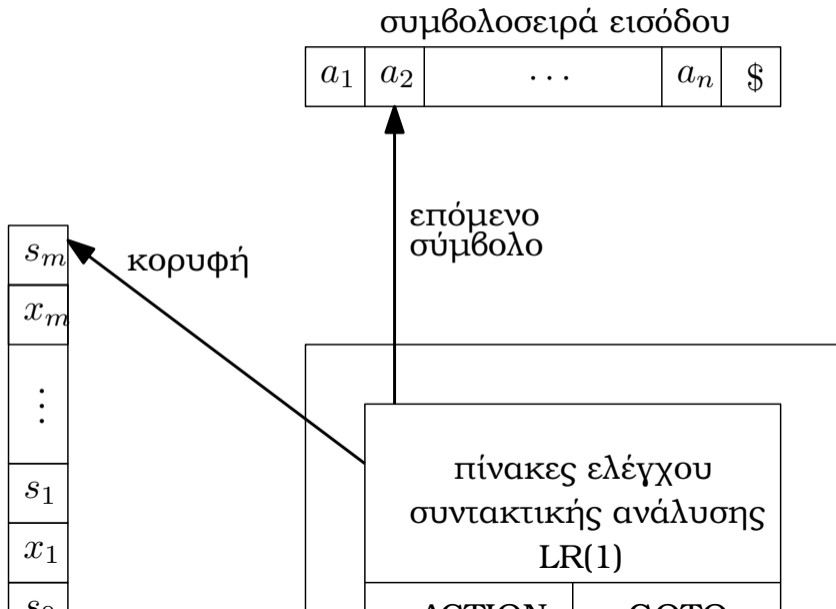
Κατασκευή Συντακτικών Αναλυτών LR

- Η κατασκευή ενός LR συντακτικού αναλυτή είναι πολύ δουλειά για να γίνει με το χέρι.
- Χρησιμοποιούμε συνήθως κάποιο αυτοματοποιημένο εργαλείο.
- Πρέπει όμως να γνωρίζουμε μερικές βασικές έννοιες (όπως οι καταστάσεις) ακόμη και αν χρησιμοποιούμε κάποιο αυτόματο εργαλείο.
- Τα μηνύματα εξόδου μιας γεννήτριας LR συντακτικών αναλυτών εμπεριέχουν τις καταστάσεις του συντακτικού αναλυτή.

- Ένας συντακτικός αναλυτής $LR(k)$ είναι ένα ντετερμινιστικό αυτόματο με μια επιπλέον στοίβα.
- Ένας συντακτικός αναλυτής $LR(k)$ λειτουργεί και αυτός με έναν πίνακα ελέγχου.
- Ο αναλυτής κοιτάει k σύμβολα της εισόδου ώστε να πάρει απόφαση.
- Ο πίνακας ελέγχου περιέχει αποφάσεις ανάλογα με τις πληροφορίες που υπάρχουν στην στοίβα και τα k σύμβολα της εισόδου.

Συντακτικοί Αναλυτές LR(1)

LR Parsing



Πίνακας Ελέγχου Συντακτικής Ανάλυσης

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

Πίνακας Ελέγχου Συντακτικής Ανάλυσης

Έχει μια γραμμή για κάθε κατάσταση και μια στήλη για κάθε τερματικό, μη-τερματικό καθώς και το σύμβολο \$.

Περιέχει τα εξής στοιχεία:

sn	ολίσθηση στην κατάσταση n (s hift)
rn	ελάττωση με τον κανόνα n (r educe)
gn	πήγαινε στην κατάσταση n (g oto)
acc	αποδοχή (a cccept)
	λάθος

Ο Αλγόριθμος Συντακτικής Ανάλυσης LR(1)

Input: Συμβολοσειρά εισόδου w και πίνακας LR με συναρτήσεις *ACTION* και *GOTO* γραμματικής G

Output: Αν $w \in L(G)$ τότε σύνολο παραγωγών για ανοδική συντακτική ανάλυση της w αλλιώς λάθος.

πρόσθεσε την αρχική κατάσταση στην κορυφή της στοίβας;

while αλήθεια **do**

 έστω s η κατάσταση στην κορυφή της στοίβας;

 έστω a το πρώτο σύμβολο της $w\$,$

if $ACTION[s, a] = shift\ t$ **then**

 πρόσθεση του συμβόλου a στην κορυφή της στοίβας;

 πρόσθεση της κατάστασης t στην κορυφή της στοίβας;

 αφαίρεση του συμβόλου a από την συμβολοσειρά εισόδου $w,$

else if $ACTION[s, a] = reduce\ A \rightarrow \beta_1\beta_2 \cdots \beta_k$ **then**

 αφαίρεση $2k$ συμβόλων από την στοίβα;

 έστω s' η κατάσταση στην κορυφή της στοίβας;

 πρόσθεση του A στην κορυφή της στοίβας;

 πρόσθεση της κατάστασης $GOTO[s', A]$ στην κορυφή της στοίβας;

 τύπωσε τον κανόνα $A \rightarrow \beta_1\beta_2 \cdots \beta_k;$

else if $ACTION[s, a] = accept$ **then**

 break;

else

 error();

end

end

Παράδειγμα

0	S	\rightarrow	E
1	E	\rightarrow	T
2	E	\rightarrow	$E + T$
3	T	\rightarrow	F
4	T	\rightarrow	$T * F$
5	F	\rightarrow	(E)
6	F	\rightarrow	id

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

Ας δούμε πως λειτουργεί ο συντακτικός αναλυτής που φαίνεται παραπάνω στην είσοδο

$id * id + id$

Ο πίνακας αυτός κατασκευάστηκε με την μέθοδο SLR που θα δούμε παρακάτω. Η αρχική κατάσταση είναι η κατάσταση 0.

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα
0

είσοδος
 $id * id + id \$$

κίνηση
s5

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα

0
0 *id* 5

είσοδος

id * *id* + *id* \$
* *id* + *id* \$

κίνηση

s5
r6 και g3

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2
0 T 2	$* id + id \$$	s7

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2
0 T 2	$* id + id \$$	s7
0 T 2 * 7	$id + id \$$	s5

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 $id 5$	$* id + id \$$	r6 και g3
0 $F 3$	$* id + id \$$	r3 και g2
0 $T 2$	$* id + id \$$	s7
0 $T 2 * 7$	$id + id \$$	s5
0 $T 2 * 7 id 5$	$+ id \$$	r6 και g10

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2
0 T 2	$* id + id \$$	s7
0 T 2 * 7	$id + id \$$	s5
0 T 2 * 7 id 5	$+ id \$$	r6 και g10
0 T 2 * 7 F 10	$+ id \$$	r4 και g2

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 <i>id</i> 5	$* id + id \$$	r6 και g3
0 <i>F</i> 3	$* id + id \$$	r3 και g2
0 <i>T</i> 2	$* id + id \$$	s7
0 <i>T</i> 2 * 7	$id + id \$$	s5
0 <i>T</i> 2 * 7 <i>id</i> 5	$+ id \$$	r6 και g10
0 <i>T</i> 2 * 7 <i>F</i> 10	$+ id \$$	r4 και g2
0 <i>T</i> 2	$+ id \$$	r1 και g1

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 <i>id</i> 5	$* id + id \$$	r6 και g3
0 <i>F</i> 3	$* id + id \$$	r3 και g2
0 <i>T</i> 2	$* id + id \$$	s7
0 <i>T</i> 2 * 7	$id + id \$$	s5
0 <i>T</i> 2 * 7 <i>id</i> 5	$+ id \$$	r6 και g10
0 <i>T</i> 2 * 7 <i>F</i> 10	$+ id \$$	r4 και g2
0 <i>T</i> 2	$+ id \$$	r1 και g1
0 <i>E</i> 1	$+ id \$$	s6

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6			acc				
2		r1	s7		r1				
3		r3	r3		r3				
4	s5			s4			g8	g2	g3
5		r6	r6		r6				
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2				
10		r4	r4		r4				
11		r5	r5		r5				

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 <i>id</i> 5	$* id + id \$$	r6 και g3
0 <i>F</i> 3	$* id + id \$$	r3 και g2
0 <i>T</i> 2	$* id + id \$$	s7
0 <i>T</i> 2 * 7	$id + id \$$	s5
0 <i>T</i> 2 * 7 <i>id</i> 5	$+ id \$$	r6 και g10
0 <i>T</i> 2 * 7 <i>F</i> 10	$+ id \$$	r4 και g2
0 <i>T</i> 2	$+ id \$$	r1 και g1
0 <i>E</i> 1	$+ id \$$	s6
0 <i>E</i> 1 + 6	$id \$$	s5

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2
0 T 2	$* id + id \$$	s7
0 T 2 * 7	$id + id \$$	s5
0 T 2 * 7 id 5	$+ id \$$	r6 και g10
0 T 2 * 7 F 10	$+ id \$$	r4 και g2
0 T 2	$+ id \$$	r1 και g1
0 E 1	$+ id \$$	s6
0 E 1 + 6	$id \$$	s5
0 E 1 + 6 id 5	$\$$	r6 και g3

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 id 5	$* id + id \$$	r6 και g3
0 F 3	$* id + id \$$	r3 και g2
0 T 2	$* id + id \$$	s7
0 T 2 * 7	$id + id \$$	s5
0 T 2 * 7 id 5	$+ id \$$	r6 και g10
0 T 2 * 7 F 10	$+ id \$$	r4 και g2
0 T 2	$+ id \$$	r1 και g1
0 E 1	$+ id \$$	s6
0 E 1 + 6	$id \$$	s5
0 E 1 + 6 id 5	$\$$	r6 και g3
0 E 1 + 6 F 3	$\$$	r3 και g9

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 <i>id</i> 5	$* id + id \$$	r6 και g3
0 <i>F</i> 3	$* id + id \$$	r3 και g2
0 <i>T</i> 2	$* id + id \$$	s7
0 <i>T</i> 2 * 7	$id + id \$$	s5
0 <i>T</i> 2 * 7 <i>id</i> 5	$+ id \$$	r6 και g10
0 <i>T</i> 2 * 7 <i>F</i> 10	$+ id \$$	r4 και g2
0 <i>T</i> 2	$+ id \$$	r1 και g1
0 <i>E</i> 1	$+ id \$$	s6
0 <i>E</i> 1 + 6	$id \$$	s5
0 <i>E</i> 1 + 6 <i>id</i> 5	$\$$	r6 και g3
0 <i>E</i> 1 + 6 <i>F</i> 3	$\$$	r3 και g9
0 <i>E</i> 1 + 6 <i>T</i> 9	$\$$	r2 και g1

Παράδειγμα

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			g1	g2	g3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			g8	g2	g3
5		r6	r6		r6	r6			
6	s5			s4				g9	g3
7	s5			s4					g10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

στοίβα	είσοδος	κίνηση
0	$id * id + id \$$	s5
0 <i>id</i> 5	$* id + id \$$	r6 και g3
0 <i>F</i> 3	$* id + id \$$	r3 και g2
0 <i>T</i> 2	$* id + id \$$	s7
0 <i>T</i> 2 * 7	$id + id \$$	s5
0 <i>T</i> 2 * 7 <i>id</i> 5	$+ id \$$	r6 και g10
0 <i>T</i> 2 * 7 <i>F</i> 10	$+ id \$$	r4 και g2
0 <i>T</i> 2	$+ id \$$	r1 και g1
0 <i>E</i> 1	$+ id \$$	s6
0 <i>E</i> 1 + 6	$id \$$	s5
0 <i>E</i> 1 + 6 <i>id</i> 5	$\$$	r6 και g3
0 <i>E</i> 1 + 6 <i>F</i> 3	$\$$	r3 και g9
0 <i>E</i> 1 + 6 <i>T</i> 9	$\$$	r2 και g1
0 <i>E</i> 1	$\$$	acc

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Σημαντικός αλγόριθμος από θεωρητική σκοπιά, δεν προκύπτει συχνά στην πράξη.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Σημαντικός αλγόριθμος από θεωρητική σκοπιά, δεν προκύπτει συχνά στην πράξη.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που θα μελετήσουμε παρακάτω

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Σημαντικός αλγόριθμος από θεωρητική σκοπιά, δεν προκύπτει συχνά στην πράξη.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που θα μελετήσουμε παρακάτω
- 3 πλήρης αλγόριθμος $LR(1)$, κατασκευάζει αρκετά μεγάλους πίνακες ελέγχου

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Σημαντικός αλγόριθμος από θεωρητική σκοπιά, δεν προκύπτει συχνά στην πράξη.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που θα μελετήσουμε παρακάτω
- 3 πλήρης αλγόριθμος $LR(1)$, κατασκευάζει αρκετά μεγάλους πίνακες ελέγχου
- 4 αλγόριθμος $LALR(1)$ (lookahead $LR(1)$) που προκύπτει εννώνοντας καταστάσεις ενός $LR(1)$ αναλυτή που έχουν ίδια στοιχεία εκτός από τα σύνολα προεπισκόπησης. Το πιο γνωστό εργαλείο YACC και οι διάφορες υλοποιήσεις του είναι ένας $LALR(1)$ συντακτικός αναλυτής.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Σημαντικός αλγόριθμος από θεωρητική σκοπιά, δεν προκύπτει συχνά στην πράξη.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που θα μελετήσουμε παρακάτω
- 3 πλήρης αλγόριθμος $LR(1)$, κατασκευάζει αρκετά μεγάλους πίνακες ελέγχου
- 4 αλγόριθμος $LALR(1)$ (lookahead $LR(1)$) που προκύπτει εννώνοντας καταστάσεις ενός $LR(1)$ αναλυτή που έχουν ίδια στοιχεία εκτός από τα σύνολα προεπισκόπησης. Το πιο γνωστό εργαλείο YACC και οι διάφορες υλοποιήσεις του είναι ένας $LALR(1)$ συντακτικός αναλυτής.

Σε όλες τις περιπτώσεις ο αλγόριθμος της συντακτικής ανάλυσης είναι ο ίδιος, μόνο ο πίνακας ελέγχου αλλάζει.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Σε σχέση με την εκφραστικότητα των γλωσσών έχουμε

$$LR(0) \subset SLR(1) \subset LALR(1) \subset LR(1)$$

όπου π.χ μια γραμματική ανήκει στην κατηγορία $LR(1)$ εάν ο αλγόριθμος κατασκευής του πίνακα ελέγχου έχει ως έξοδο ένα πίνακα που δεν έχει παραπάνω από ένα στοιχεία σε κάποια θέση του.

Συντακτικοί Αναλυτές $SLR(1)$

Στοιχεία και Καταστάσεις

Περιγράψαμε την λειτουργία ενός συντακτικού αναλυτή $LR(1)$ αλλά δεν περιγράψαμε ακόμη τις καταστάσεις του.

Θα ασχοληθούμε με την ευκολότερη κατηγορία συντακτικών αναλυτών $SLR(1)$ που σημαίνει Simple $LR(1)$.

Για να περιγράψουμε τις καταστάσεις ενός SLR συντακτικού αναλυτή θα εξηγήσουμε πρώτα την έννοια του *στοιχείου*.

Συντακτικοί Αναλυτές *SLR*

Στοιχεία

Ένα στοιχείο μια γραμματικής είναι ένας κανόνας παραγωγής όπου ανάμεσα στα σύμβολα που αποτελούν το δεξιό μέλος υπάρχει και το σύμβολο \bullet .

π.χ στον κανόνα παραγωγής $A \rightarrow xy$ αντιστοιχούν τα τρία στοιχεία

1 $A \rightarrow \bullet xy$

2 $A \rightarrow x \bullet y$

3 $A \rightarrow xy \bullet$

ενώ στο κανόνα $A \rightarrow \epsilon$ αντιστοιχεί μόνο το στοιχείο $A \rightarrow \bullet$.

Συντακτικοί Αναλυτές *SLR*

Στοιχεία

Ένα στοιχείο μια γραμματικής είναι ένας κανόνας παραγωγής όπου ανάμεσα στα σύμβολα που αποτελούν το δεξιό μέλος υπάρχει και το σύμβολο \bullet .

π.χ στον κανόνα παραγωγής $A \rightarrow xy$ αντιστοιχούν τα τρία στοιχεία

1 $A \rightarrow \bullet xy$

2 $A \rightarrow x \bullet y$

3 $A \rightarrow xy \bullet$

ενώ στο κανόνα $A \rightarrow \epsilon$ αντιστοιχεί μόνο το στοιχείο $A \rightarrow \bullet$.

Τα στοιχεία δείχνουν ποιο τμήμα ενός κανόνα παραγωγής έχει ήδη χρησιμοποιηθεί για την αναγνώριση της συμβολοσειράς εισόδου.

Θα δούμε πως οι καταστάσεις ενός συντακτικού αναλυτή *SLR* είναι υποσύνολα στοιχείων.

Συντακτικοί Αναλυτές *SLR*

Καταστάσεις

- 1 Μία κατάσταση ενός συντακτικού αναλυτή *SLR* είναι ένα υποσύνολο στοιχείων της γραμματικής.
- 2 Μία κατάσταση μοντελοποιεί τι υπάρχει στην στοίβα και τι βλέπουμε μπροστά μας στην είσοδο.

Συντακτικοί Αναλυτές *SLR*

Καταστάσεις

Πως μπορούμε να περιγράψουμε την κατάσταση ενός συντακτικού αναλυτή *SLR* για την παρακάτω γραμματική;

$$\begin{array}{l} 0 \quad S \rightarrow E\$ \\ 1 \quad E \rightarrow T + E \\ 2 \quad E \rightarrow T \\ 2 \quad T \rightarrow x \end{array}$$

Αρχικά η στοίβα είναι άδεια και η είσοδος είναι μια πρόταση E ακολουθούμενη από $\$$.

Γράφουμε

$$S \rightarrow \bullet E\$$$

για να συμβολίσουμε πως ο συντακτικός αναλυτής βρίσκεται ακριβώς πριν από ένα E .

Συντακτικοί Αναλυτές SLR

Καταστάσεις

Αφού η είσοδος ξεκινά με ένα E , μπορεί να ξεκινά με οποιοδήποτε δεξιό μέρος μιας παραγωγής του E .

Επίσης αφού ξεκινά με ένα T λόγω την παραγωγής $E \rightarrow T$ μπορεί να ξεκινά με οποιοδήποτε δεξιό μέρος μιας παραγωγής του T .

S	\rightarrow	$\bullet E \$$
E	\rightarrow	$\bullet T + E$
E	\rightarrow	$\bullet T$
T	\rightarrow	$\bullet x$

0	S	\rightarrow	$E \$$
1	E	\rightarrow	$T + E$
2	E	\rightarrow	T
2	T	\rightarrow	x

Καταλήγουμε στο σύνολο στοιχείων στα αριστερά που αποτελεί την αρχική κατάσταση I_0 του SLR αυτόματου.

Συντακτικοί Αναλυτές *SLR*

Ολίσθηση

Έστω πως ο συντακτικός αναλυτής βρίσκεται στην κατάσταση I_0 .

S	\rightarrow	$\bullet E\$$
E	\rightarrow	$\bullet T + E$
E	\rightarrow	$\bullet T$
T	\rightarrow	$\bullet x$

0	S	\rightarrow	$E\$$
1	E	\rightarrow	$T + E$
2	E	\rightarrow	T
2	T	\rightarrow	x

Συντακτικοί Αναλυτές SLR

Ολίσθηση

Έστω πως ο συντακτικός αναλυτής βρίσκεται στην κατάσταση I_0 .

S	\rightarrow	$\bullet E \$$
E	\rightarrow	$\bullet T + E$
E	\rightarrow	$\bullet T$
T	\rightarrow	$\bullet x$

0	S	\rightarrow	$E \$$
1	E	\rightarrow	$T + E$
2	E	\rightarrow	T
2	T	\rightarrow	x

Τι θα συμβεί άμα κάνουμε ολίσθηση ενός x ; Η κορυφή της στοίβας θα έχει πλέον ένα x .

T	\rightarrow	$x \bullet$
-----	---------------	-------------

Αυτή είναι η κατάσταση I_1 του συντακτικού αναλυτή που συμβολίζει πως σε αυτή την κατάσταση είμαστε έτοιμοι να κάνουμε ελάττωση με βάση αυτή την παραγωγή.

Συντακτικοί Αναλυτές SLR

Μετάβαση

Φανταστείτε την μετάβαση από την κατάσταση I_0 αν ο συντακτικός αναλυτής διαβάσει μια συμβολοσειρά που παράχθηκε από το T .

S	\rightarrow	$\bullet E \$$
E	\rightarrow	$\bullet T + E$
E	\rightarrow	$\bullet T$
T	\rightarrow	$\bullet x$

0	S	\rightarrow	$E \$$
1	E	\rightarrow	$T + E$
2	E	\rightarrow	T
3	T	\rightarrow	x

Συντακτικοί Αναλυτές SLR

Μετάβαση

Φανταστείτε την μετάβαση από την κατάσταση I_0 αν ο συντακτικός αναλυτής διαβάσει μια συμβολοσειρά που παράχθηκε από το T .

S	\rightarrow	$\bullet E \$$
E	\rightarrow	$\bullet T + E$
E	\rightarrow	$\bullet T$
T	\rightarrow	$\bullet x$

0	S	\rightarrow	$E \$$
1	E	\rightarrow	$T + E$
2	E	\rightarrow	T
3	T	\rightarrow	x

Μετά την ελάττωση με βάση ένα κανόνα παραγωγής με αριστερό μέρος το T , η κορυφή της στοίβας θα περιέχει το T .

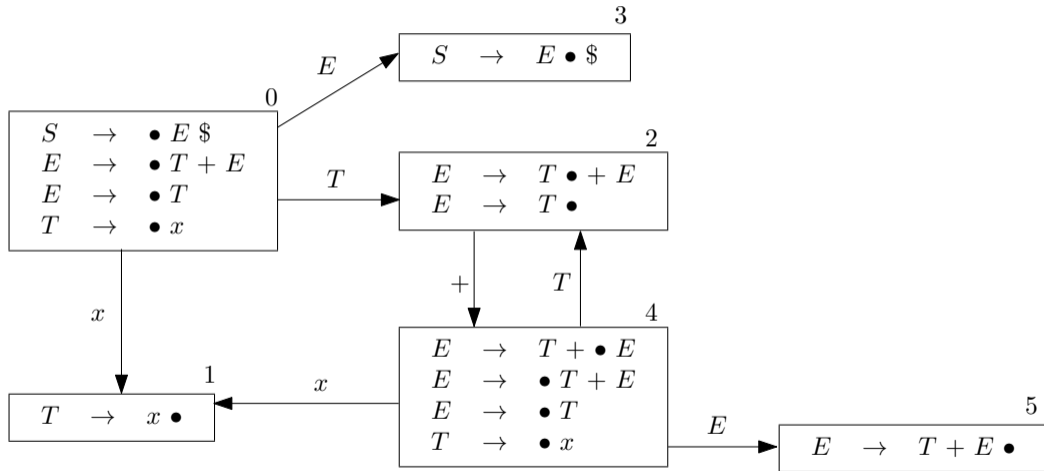
E	\rightarrow	$T \bullet + E$
E	\rightarrow	$T \bullet$

Αυτή είναι η κατάσταση I_2 του συντακτικού αναλυτή που συμβολίζει πως σε αυτή την κατάσταση μπορούμε

- να κάνουμε ολίσθηση ενός $+$,
- ή να κάνουμε ελάττωση με τον κανόνα $E \rightarrow T$.

Συντακτικοί Αναλυτές SLR

Αν συστηματοποιήσουμε αυτή την διαδικασία θα οδηγηθούμε στις παρακάτω καταστάσεις:



Συντακτικοί Αναλυτές *SLR*

Η αυτοματοποίηση της διαδικασίας κατασκευής των καταστάσεων του αυτόματου χρειάζεται να ορίσουμε δυο βασικές διαδικασίες.

Έστω I ένα σύνολο στοιχείων και X ένα σύμβολο της γραμματικής, τερματικό ή μη-τερματικό.

closure(I)

Προσθέτει περισσότερα στοιχεία σε ένα σύνολο στοιχείων όταν υπάρχει \bullet αριστερά από ένα μη-τερματικό.

goto(I, X)

Μετακινεί την \bullet μετά το σύμβολο X σε όλα τα στοιχεία του I .

Συντακτικοί Αναλυτές SLR

Η συνάρτηση closure

Input: Σύνολο στοιχείων I

Output: $closure(I)$

repeat

for κάθε στοιχείο $A \rightarrow \alpha \bullet X\beta$ στο I **do**

for κάθε κανόνα $X \rightarrow \gamma$ στην γραμματική **do**

$I \leftarrow I \cup \{X \rightarrow \bullet\gamma\};$

end

end

until το I δεν αλλάζει;

return το σύνολο I ;

Συντακτικοί Αναλυτές SLR

Η συνάρτηση closure

Έστω η παρακάτω γραμματική

- 0 $S' \rightarrow S\$$
- 1 $S \rightarrow (L)$
- 2 $S \rightarrow x$
- 3 $L \rightarrow S$
- 4 $L \rightarrow L, S$

Εάν $I = \{S \rightarrow (\bullet L)\}$ τότε το $closure(I)$ είναι

S	\rightarrow	$(\bullet L)$
L	\rightarrow	$\bullet S$
L	\rightarrow	$\bullet L, S$
S	\rightarrow	$\bullet x$
S	\rightarrow	$\bullet (L)$

Συντακτικοί Αναλυτές *SLR*

Η συνάρτηση *goto*

Input: Σύνολο στοιχείων I και σύμβολο X τερματικό ή μη-τερματικό

Output: $goto(I, X)$

έστω J το άδειο σύνολο

for κάθε στοιχείο $A \rightarrow \alpha \bullet X \beta$ στο I **do**

 | $J \leftarrow J \cup \{A \rightarrow \alpha X \bullet \beta\};$

end

return $closure(J);$

Συντακτικοί Αναλυτές SLR

Η συνάρτηση goto

Έστω η παρακάτω γραμματική

0	S'	\rightarrow	$S\$$
1	S	\rightarrow	(L)
2	S	\rightarrow	x
3	L	\rightarrow	S
4	L	\rightarrow	L, S

Εάν το σύνολο αριστερά είναι το I τότε βλέπουμε δεξιά το σύνολο $goto(I, ($).

S'	\rightarrow	$\bullet S\$$
S	\rightarrow	$\bullet (L)$
S	\rightarrow	$\bullet x$

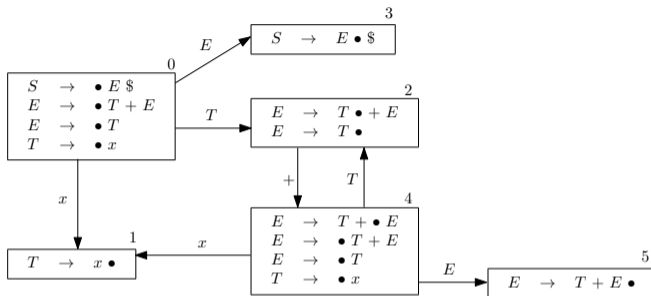
S	\rightarrow	$(\bullet L)$
L	\rightarrow	$\bullet S$
L	\rightarrow	$\bullet L, S$
S	\rightarrow	$\bullet x$
S	\rightarrow	$\bullet (L)$

Υπολογισμός Καταστάσεων

- πρόσθεσε τον κανόνα $S' \rightarrow S\$$
- εκκίνηση με $I_0 = closure(\{S' \rightarrow \bullet S\})$
- τα σύνολα στοιχείων είναι οι καταστάσεις
- για κάθε κατάσταση I και σύμβολο X τερματικό ή μη-τερματικό αν το σύνολο στοιχείων $goto(I, X)$ δεν είναι ήδη μια κατάσταση, τότε πρόσθεσε το ως μια νέα κατάσταση
- επανάληψη μέχρι να μην βρίσκουμε νέες καταστάσεις

Συντακτικοί Αναλυτές SLR

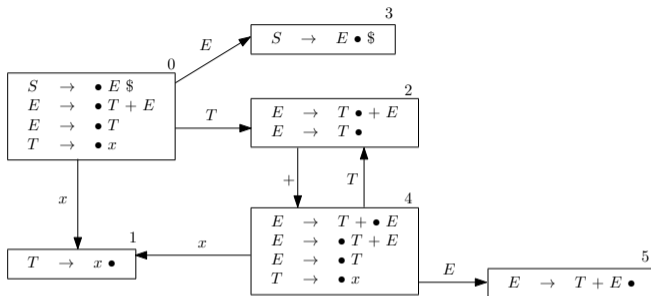
Για να γεμίσουμε τον πίνακα ακολουθούμε τους εξής κανόνες.



	ACTION			GOTO	
	x	+	\$	E	T
0					
1					
2					
3					
4					
5					

Συντακτικοί Αναλυτές SLR

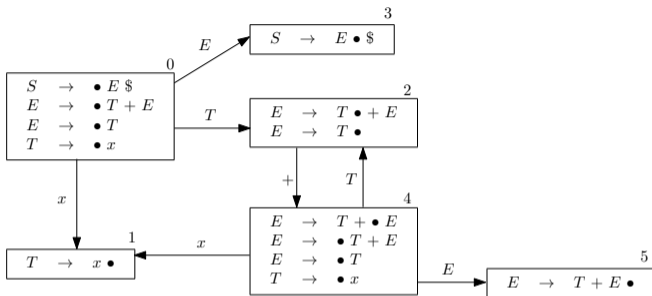
Για κάθε μετάβαση από I_i σε I_j με τερματικό σύμβολο X βάζουμε sj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0					
1					
2					
3					
4					
5					

Συντακτικοί Αναλυτές SLR

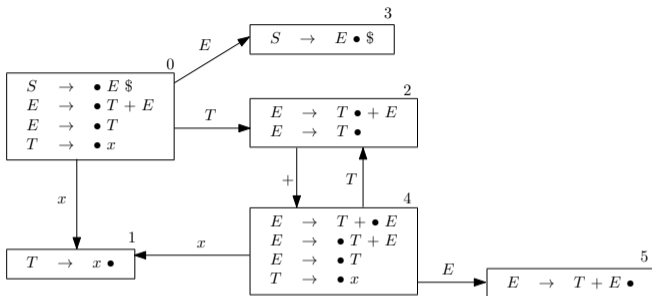
Για κάθε μετάβαση από I_i σε I_j με τερματικό σύμβολο X βάζουμε sj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1				
1					
2					
3					
4					
5					

Συντακτικοί Αναλυτές SLR

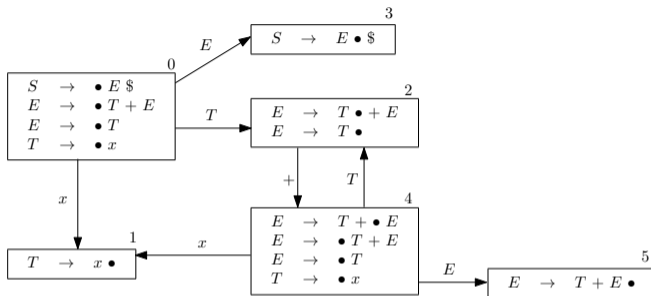
Για κάθε μετάβαση από I_i σε I_j με τερματικό σύμβολο X βάζουμε sj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1				
1					
2		s4			
3					
4					
5					

Συντακτικοί Αναλυτές SLR

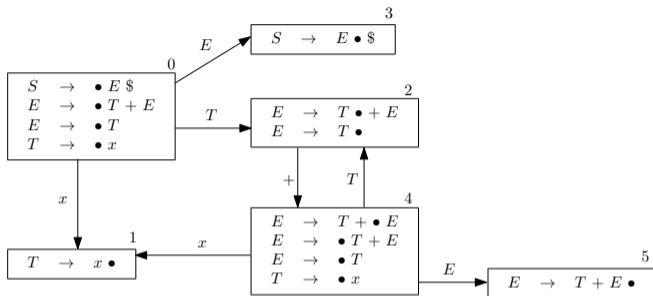
Για κάθε μετάβαση από I_i σε I_j με τερματικό σύμβολο X βάζουμε sj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1				
1					
2		s4			
3					
4	s1				
5					

Συντακτικοί Αναλυτές SLR

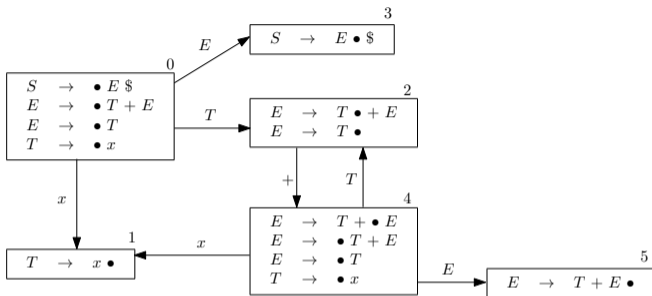
Για κάθε μετάβαση από I_i σε I_j με μη-τερματικό σύμβολο X βάζουμε gj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1				
1					
2		s4			
3					
4	s1				
5					

Συντακτικοί Αναλυτές SLR

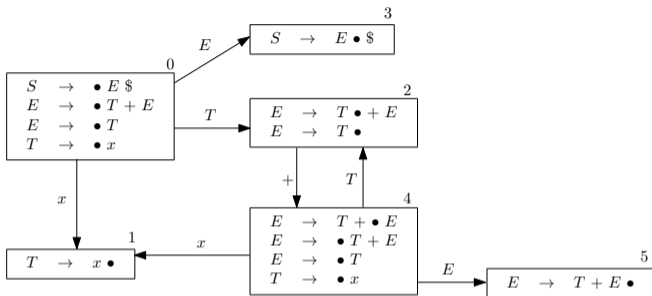
Για κάθε μετάβαση από I_i σε I_j με μη-τερματικό σύμβολο X βάζουμε gj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	
1					
2		s4			
3					
4	s1				
5					

Συντακτικοί Αναλυτές SLR

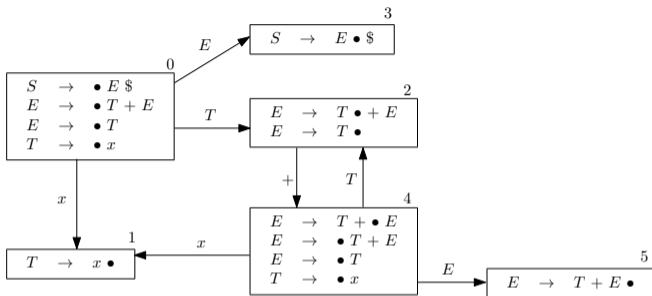
Για κάθε μετάβαση από I_i σε I_j με μη-τερματικό σύμβολο X βάζουμε gj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3					
4	s1				
5					

Συντακτικοί Αναλυτές SLR

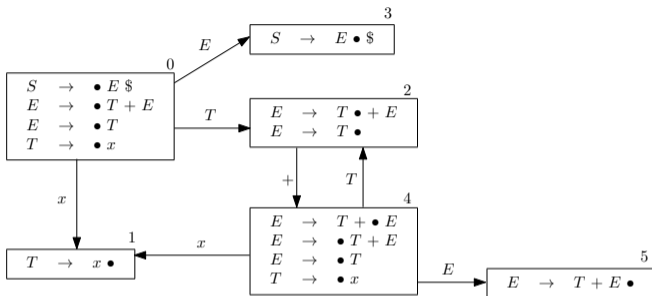
Για κάθε μετάβαση από I_i σε I_j με μη-τερματικό σύμβολο X βάζουμε gj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3					
4	s1				g2
5					

Συντακτικοί Αναλυτές SLR

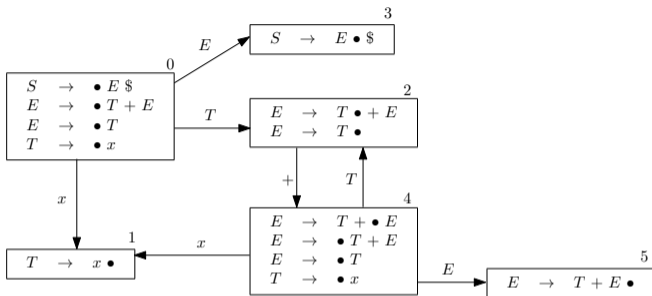
Για κάθε μετάβαση από I_i σε I_j με μη-τερματικό σύμβολο X βάζουμε gj στην θέση (i, X) του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3					
4	s1			g5	g2
5					

Συντακτικοί Αναλυτές SLR

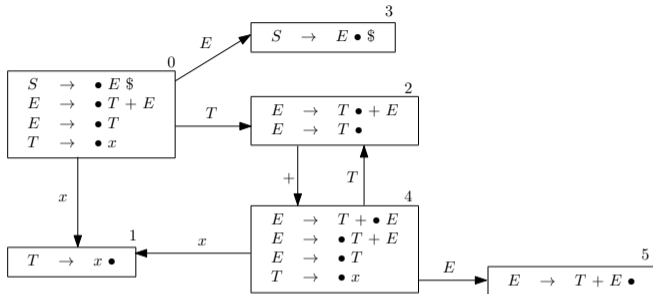
Για κάθε κατάσταση I_i που περιέχει τον κανόνα με το $\$$ και τελεία στο τέλος, π.χ $S \rightarrow E \bullet \$$ βάζουμε **acc** στην θέση $(i, \$)$ του πίνακα.



	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3					
4	s1			g5	g2
5					

Συντακτικοί Αναλυτές SLR

Για κάθε κατάσταση I_i που περιέχει τον κανόνα με το $\$$ και τελεία στο τέλος, π.χ $S \rightarrow E \bullet \$$ βάζουμε **acc** στην θέση $(i, \$)$ του πίνακα.



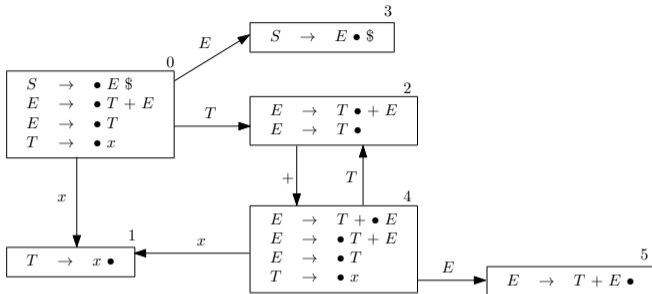
	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3			acc		
4	s1			g5	g2
5					

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



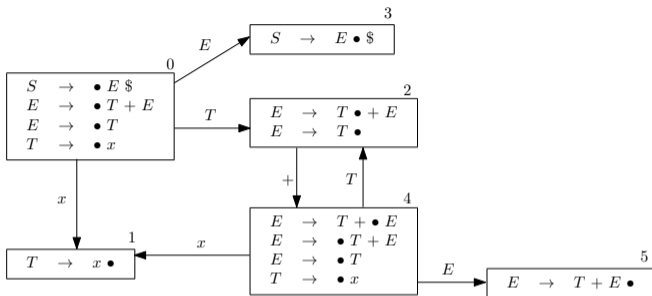
	ACTION			GOTO	
	x	$+$	$\$$	E	T
0	s1			g3	g2
1					
2		s4			
3			acc		
4	s1			g5	g2
5					

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



	ACTION			GOTO	
	x	+	\$	E	T
0	s1			g3	g2
1					
2		s4			
3			acc		
4	s1			g5	g2
5					

- 0 $S \rightarrow E \$$
- 1 $E \rightarrow T + E$
- 2 $E \rightarrow T$
- 3 $T \rightarrow x$

$$FOLLOW(T) = \{+, \$\}$$

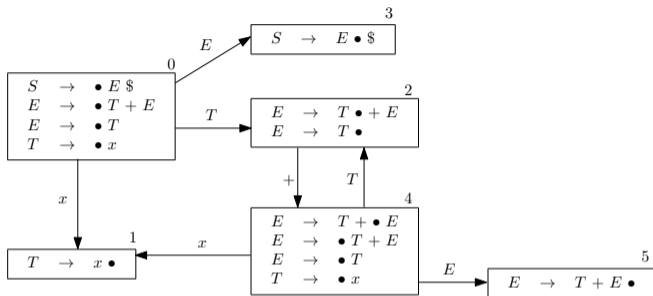
$$FOLLOW(E) = \{\$ \}$$

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



	ACTION			GOTO	
	x	+	\$	E	T
0	s1			g3	g2
1		r3			
2		s4			
3			acc		
4	s1			g5	g2
5					

0 $S \rightarrow E\$$
 1 $E \rightarrow T + E$
 2 $E \rightarrow T$
 3 $T \rightarrow x$

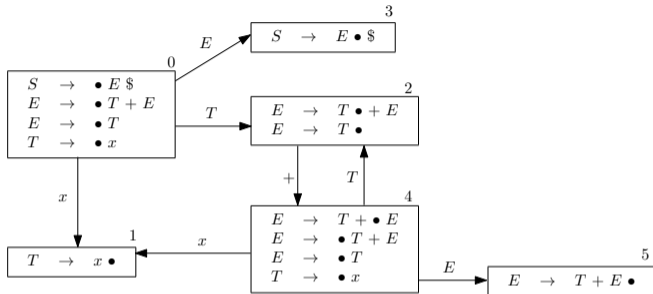
$FOLLOW(T) = \{+, \$\}$
 $FOLLOW(E) = \{\$ \}$

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



	ACTION			GOTO	
	x	+	\$	E	T
0	s1			g3	g2
1		r3	r3		
2		s4			
3			acc		
4	s1			g5	g2
5					

- 0 $S \rightarrow E\$$
- 1 $E \rightarrow T + E$
- 2 $E \rightarrow T$
- 3 $T \rightarrow x$

$$FOLLOW(T) = \{+, \$\}$$

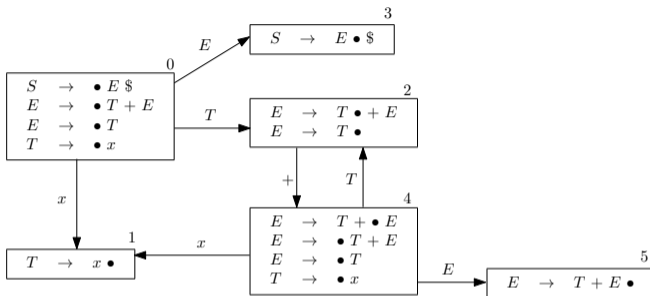
$$FOLLOW(E) = \{\$ \}$$

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



	ACTION			GOTO	
	x	+	\$	E	T
0	s1			g3	g2
1		r3	r3		
2		s4	r2		
3			acc		
4	s1			g5	g2
5					

0 $S \rightarrow E\$$
 1 $E \rightarrow T + E$
 2 $E \rightarrow T$
 3 $T \rightarrow x$

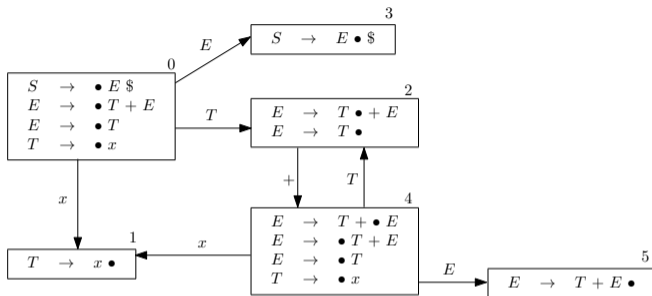
$FOLLOW(T) = \{+, \$\}$
 $FOLLOW(E) = \{\$ \}$

Συντακτικοί Αναλυτές SLR

Για μια κατάσταση I_i που περιέχει τον κανόνα n ως στοιχείο με την τελεία στο τέλος

$$A \rightarrow \gamma \bullet$$

βάζουμε n στην θέση (i, X) του πίνακα όπου $X \in FOLLOW(A)$.



	ACTION			GOTO	
	x	+	\$	E	T
0	s1			g3	g2
1		r3	r3		
2		s4	r2		
3			acc		
4	s1			g5	g2
5			r1		

0 $S \rightarrow E\$$
 1 $E \rightarrow T + E$
 2 $E \rightarrow T$
 3 $T \rightarrow x$

$FOLLOW(T) = \{+, \$\}$
 $FOLLOW(E) = \{\$ \}$

Άλλο Παράδειγμα SLR

0 $S \rightarrow E$
 1 $E \rightarrow T$
 2 $E \rightarrow E + T$
 3 $T \rightarrow F$
 4 $T \rightarrow T * F$
 5 $F \rightarrow (E)$
 6 $F \rightarrow id$

	ACTION					GOTO			
	<i>id</i>	+	*	()	\$	<i>E</i>	<i>T</i>	<i>F</i>
0	s5			s4			s1	s2	s3
1		s6				acc			
2		r1	s7		r1	r1			
3		r3	r3		r3	r3			
4	s5			s4			s8	s2	s3
5		r6	r6		r6	r6			
6	s5			s4				s9	s3
7	s5			s4					s10
8		s6			s11				
9		r2	s7		r2	r2			
10		r4	r4		r4	r4			
11		r5	r5		r5	r5			

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Δεν χρησιμοποιεί την συνάρτηση $FOLLOW(\cdot)$ για να θέσει τις ελαττώσεις.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Δεν χρησιμοποιεί την συνάρτηση $FOLLOW(\cdot)$ για να θέσει τις ελαττώσεις.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που είδαμε σε λεπτομέρεια.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Δεν χρησιμοποιεί την συνάρτηση $FOLLOW(\cdot)$ για να θέσει τις ελαττώσεις.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που είδαμε σε λεπτομέρεια.
- 3 πλήρης αλγόριθμος $LR(1)$, κατασκευάζει αρκετά μεγάλους πίνακες ελέγχου αφού έχει πιο πολύπλοκα στοιχεία που περιέχουν και το σύμβολο προεπισκόπησης.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Υπάρχουν διάφοροι αλγόριθμοι κατασκευής του πίνακα ελέγχου με διαφορετικά χαρακτηριστικά:

- 1 ο αλγόριθμος $LR(0)$ σε περίπτωση που η γραμματική ανήκει στην κατηγορία $LR(0)$ δηλαδή δεν χρειάζεται προεπισκόπηση λεκτικών μονάδων για την αναγνώριση της. Δεν χρησιμοποιεί την συνάρτηση $FOLLOW(\cdot)$ για να θέσει τις ελαττώσεις.
- 2 ο αλγόριθμος $SLR(1)$ ή Simple $LR(1)$ που είδαμε σε λεπτομέρεια.
- 3 πλήρης αλγόριθμος $LR(1)$, κατασκευάζει αρκετά μεγάλους πίνακες ελέγχου αφού έχει πιο πολύπλοκα στοιχεία που περιέχουν και το σύμβολο προεπισκόπησης.
- 4 αλγόριθμος $LALR(1)$ (lookahead $LR(1)$) που προκύπτει ενώνοντας καταστάσεις ενός $LR(1)$ αναλυτή που έχουν ίδια στοιχεία εκτός από τα σύνολα προεπισκόπησης. Το πιο γνωστό εργαλείο YACC και οι διάφορες υλοποιήσεις του είναι ένας $LALR(1)$ συντακτικός αναλυτής.

Σε όλες τις περιπτώσεις ο αλγόριθμος της συντακτικής ανάλυσης είναι ο ίδιος, μόνο ο πίνακας ελέγχου αλλάζει.

Συντακτικοί Αναλυτές $LR(1)$

Κατασκευή Πίνακα Ελέγχου

Σε σχέση με την εκφραστικότητα των γλωσσών έχουμε

$$LR(0) \subset SLR(1) \subset LALR(1) \subset LR(1)$$

όπου π.χ μια γραμματική ανήκει στην κατηγορία $LR(1)$ εάν ο αλγόριθμος κατασκευής του πίνακα ελέγχου έχει ως έξοδο ένα πίνακα που δεν έχει παραπάνω από ένα στοιχεία σε κάποια θέση του.

Συντακτικοί Αναλυτές LR

Πλεονεκτήματα

- 1 Μπορούμε να φτιάξουμε LR συντακτικούς αναλυτές για τις περισσότερες γλώσσες προγραμματισμού που μπορούν να γραφούν με γραμματικές χωρίς συμφραζόμενα.
- 2 Η μέθοδος συντακτικής ανάλυσης LR είναι η πιο γενική shift-reduce χωρίς οπισθοχώρηση μέθοδος, αλλά μπορεί να υλοποιηθεί πολύ αποδοτικά.
- 3 Ένας συντακτικός αναλυτής αναγνωρίζει λάθη στην σύνταξη όσο πιο γρήγορα γίνεται σε ένα πέρασμα από αριστερά προς τα δεξιά.
- 4 Οι γραμματικές LR μπορούν να περιγράψουν περισσότερες γλώσσες από τις LL.

Ιεραρχία Γραμματικών

