Derivatoare LR

Derivatoare LR

Algoritmul LR(k) - poate fi folosit atat pt a verifica daca o gramatica este LR(1) cat si pt construirea derivatorului sau
Algoritmul LR(k): numarul de stari este foarte mare

- similar cazului strong LL(k), multe tranzitii din LR(1) sunt independente de simbolul de lookahead
- $\Rightarrow$ putem construi un parser cu mai putine stari care implementeaza analiza LR(1) dar cu tranzitii mai putine, foosind lookahead doar cand este necesar

- LR(k)
- *simple* LR(k) : SLR(k)
- *lookahead* LR(k): LALR(k)

- se porneste cu LR(0): nu examineaza deloc simbolurile dinainte
- si se folosete lookahead doar la nevoie - simple LR(1) (SLR(1))
- Obs: nu toate LR(1) sunt SLR(1)
- LALR(1) - lookahead aplicat la SLR(1)

Reprezentarea tranzitiilor automatului intr-o forma convenabila - determinarea rapida a tranzitiei urmatoarea + limite rezonabile pentru totalul de memorie

Cum construim matricea de tranzitie?
prin Functia de tranzitie $f(q, \nu)$

- actiunile posibile in starea q cand sirul de intrare incepe cu elementul $\nu \in T$
- $+ f(q, \nu), \nu \in N$ - actiunea din starea q dupa o reducere la ν

Algoritm de parsare

Fie o gramatica LR(k) $G = (T, N, P, Z)$ si automatul stiva $A = (T, Q, R, q_0, \{q_0\}, Q, q_0)$ construit cu alg LR(k).

$$f(q, \nu) = \begin{cases} q' & daca\ \nu\gamma \in T^*\ si\ q\nu\gamma \to qq'\gamma \in R\ sau \\ & daca\ \nu \in N\ si\ q' = next(q, \nu) (shift\ \ transition) \\ X \to \chi & daca\ [X \to \chi.; \nu] \in q\ (reduce\ \ transition) \\ STOP & daca\ \nu = \#\ si\ [Z \to S.; \#] \in q \\ ERROR & altfel \end{cases}$$

Legatura LR(k)

- toate tranzitiile obtinute in pasul 5 sunt shift transitions
- tranzitiile de la pasul 7 $q_1 \ldots q_n q \omega \rightarrow q_1 q' \omega$ sunt rupte in doi pasi:
  - pt ca $[X \rightarrow \chi.; \nu]$ este in q, trebuie aplicata reducere conform $X \rightarrow \chi$ si sterse $|\chi|$ stari de pe stiva
  - noua stare $f(q_1, X) = q' = next(q_1, X)$
- daca $\omega = \#$ si $[Z \rightarrow S.; \#] \in q$ automatul se opreste

$$f(q,\nu) = \begin{cases} q' & daca\ \nu\gamma \in T^*\ si\ q\nu\gamma \rightarrow qq'\gamma \in R\ sau \\ & daca\ \nu \in N\ si\ q' = next(q,\nu) (shift\ transition) \\ X \rightarrow \chi & daca\ [X \rightarrow \chi.; \nu] \in q\ (reduce\ transition) \\ STOP & daca\ \nu = \#\ si\ [Z \rightarrow S.; \#] \in q \\ ERROR & altfel \end{cases}$$

Exemplu k=0

$$
\begin{aligned}
&(1) Z \rightarrow E\\
&(2) E \rightarrow E + F\\
&(3) E \rightarrow F\\
&(4) F \rightarrow i\\
&(5) F \rightarrow (E)
\end{aligned}
$$

Stare	E	F	i	+	(	)
q_0 $H([Z \to .E])$ $[E \to .E + F]$ $[E \to .F])$ $[F \to .i])$ $[F \to .(E)]$	$[Z \to E.]$ $[E \to E. + F]$	$[E \to F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	
$next(q_0 \ldots)$	q_1	q_2	q_3		q_4	
R			$q_0 i \to q_0 q_3$		$q_0 (\to q_0 q_4$	
q_1 $H([Z \to E.])$ $[Z \to E.]$ $[E \to E. + F]$	$\emptyset$	$\emptyset$	$\emptyset$	$[E \to E + .F]$	$\emptyset$	
$next(q_1 \ldots)$				q_5		
R				$q_1 + \to q_1 q_5$		
q_2 $H([E \to F.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
q_3 $H([F \to i.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
q_4 $H([F \to (.E)])$ $[E \to .E + F]$ $[E \to .F])$ $[F \to .i])$ $[F \to .(E)]$	$[F \to (E.)]$ $[E \to E. + F]$	$[E \to F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	
$next(q_1 \ldots)$	q_6	q_2	q_3		q_4	
R			$q_4 i \to q_4 q_3$		$q_4 (\to q_4 q_4$	

Stare	E	F	i	+	(	)
q_5 $H([E \to E + .F])$ $[F \to .i])$ $[F \to .(E)]$		$[E \to E + F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	
$next(q_5 \ldots)$		q_7	q_3		q_4	
R			$q_5 i \to q_5 q_3$		$q_5 (\to q_5 q_4$	
q_6 $H([F \to (E.)],$ $[E \to E. + F])$	$\emptyset$	$\emptyset$	$\emptyset$	$[E \to E + .F]$		$F \to (E).$
$next(q_6 \ldots)$				q_5		q_8
R				$q_6 + \to q_6 q_5$		$q_6) \to q_6 q_8$
q_7 $H([E \to E + F.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
q_8 $H([F \to (E).])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	

urmeaza pasul 7

Stare	E	F	i	$+$	(	)
q_0 $H([Z \to .E])$ $[E \to .E + F]$ $[E \to .F])$ $[F \to .i])$ $[F \to .(E)]$	$[Z \to E.]$ $[E \to E. + F]$	$[E \to F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	
$next(q_0 \ldots)$	q_1	q_2	q_3		q_4	
R			$q_0 i \to q_0 q_3$		$q_0(\to q_0 q_4$	
q_1 $H([Z \to E.])$ $[Z \to E.]$ $[E \to E. + F]$	$\emptyset$	$\emptyset$	$\emptyset$	$[E \to E + .F]$	$\emptyset$	
$next(q_1 \ldots)$				q_5		
R				$q_1 + \to q_1 q_5$		
q_2 $H([E \to F.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
step 7:R	$q_0 q_2 \to q_0 q_1,\ q_4 q_2 \to q_4 q_6$ reducere cu 3; $f(q_0, E) = q_1,\ f(q_4, E) = q_6$					
q_3 $H([F \to i.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
step 7:R	$q_0 q_3 \to q_0 q_2,\ q_4 q_3 \to q_4 q_2,\ q_5 q_3 \to q_5 q_7$ reducere cu 4; $f(q_0, F) = q_2,\ f(q_4, F) = q_2,\ f(q_5, F) = q_7$					
q_4 $H([F \to (.E)])$ $[E \to .E + F]$ $[E \to .F])$ $[F \to .i])$	$[F \to (E.)]$ $[E \to E. + F]$	$[E \to F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	

Stare	E	F	i	+	(	)
q_5 $H([E \to E+.F])$ $[F \to .i])$ $[F \to .(E)]$		$[E \to E+F.]$	$[F \to i.]$	$\emptyset$	$[F \to (.E)]$	
$next(q_5 \dots)$		q_7	q_3		q_4	
R			$q_5 i \to q_5 q_3$		$q_5(\to q_5 q_4$	
q_6 $H([F \to (E.)],$ $[E \to E.+F])$	$\emptyset$	$\emptyset$	$\emptyset$	$[E \to E+.F]$		$F \to (E).$
$next(q_6 \dots)$				q_5		q_8
R				$q_6+ \to q_6 q_5$		$q_6) \to q_6 q_8$
q_7 $H([E \to E+F.])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
step 7:R	$q_0 q_1 q_5 q_7 \to q_0 q_1$ reducere cu 2; $f(q_0, E) = q_1$					
q_8 $H([F \to (E).])$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	
step 7:R	$q_0 q_4 q_6 q_8 \to q_0 q_2$, $q_5 q_4 q_6 q_8 \to q_5 q_7$ reducere cu 5					

Exemplu k=0

$$(1)Z \rightarrow E \quad (2)E \rightarrow E+F \quad (3)E \rightarrow F (4)F \rightarrow i \quad (5)F \rightarrow (E)$$

Stare (q)		Situaþie	v	$f(q, v)$
0	*	$[Z \rightarrow .E]$	E	1
		$[E \rightarrow .E+F]$		
		$[E \rightarrow .F]$	F	2
		$[F \rightarrow .i]$	i	3
		$[F \rightarrow .(E)]$	(	4
1	*	$[Z \rightarrow E.]$	#	STOP
	*	$[E \rightarrow E.+F]$	+	5
2	*	$[E \rightarrow F.]$		red. 3
3	*	$[F \rightarrow i.]$		red. 4
4	*	$[F \rightarrow (.E)]$	E	6
		$[E \rightarrow .E+F]$		
		$[E \rightarrow .F]$	F	2
		$[F \rightarrow .i]$	i	3
		$[F \rightarrow .(E)]$	(	4
5	*	$[E \rightarrow E+.F]$	F	7
		$[F \rightarrow .i]$	i	3
		$[F \rightarrow .(E)]$	(	4
6	*	$[F \rightarrow (E.)]$	)	8
	*	$[E \rightarrow E.+F]$	+	5
7	*	$[E \rightarrow E+F.]$		red. 2
8	*	$[F \rightarrow (E).]$		red. 5

pas 7. Pentru fiecare $[X \to \chi.;\omega] \in q$, unde $\chi = x_1..x_n$ se face

$$R = R \cup \{q_1..q_n q\omega \to q_1 q'\omega | [X \to .\chi;\omega] \in q_1,$$
$$q_{i+1} = next(q_i, x_i)(i \in 1..n-1),$$
$$q = next(q_n, x_n),$$
$$q' = next(q_1, X)\}$$

Pt q_3: $H([F \to i.;\#])$

q_1 $[F \to .i;\#] \in ?$

q $? = next(?, i)$, $? = next(?, i)$, $? = next(?, i)$

q' $? = next(?, F)$

q' $? = next(?, F)$

q' $? = next(?, F)$

pas 7. Pentru fiecare $[X \to \chi.;\omega] \in q$, unde $\chi = x_1..x_n$ se face

$$\begin{aligned} R = R \cup \{q_1..q_n q\omega \to q_1 q'\omega | [X \to .\chi;\omega] \in q_1, \\ q_{i+1} = next(q_i, x_i)(i \in 1..n-1), \\ q = next(q_n, x_n), \\ q' = next(q_1, X)\} \end{aligned}$$

Pt q_3: $H([F \to i.;\#])$

$$\begin{array}{ll} q_1 & [F \to .i;\#] \in q_0, q_4, q_5 \\ q & q_3 = next(q_0, i),\ q_3 = next(q_4, i),\ q_3 = next(q_5, i) \\ q' & q_2 = next(q_0, F) \\ q' & q_2 = next(q_4, F) \\ q' & q_7 = next(q_5, F) \end{array}$$

pas 7. Pentru fiecare $[X \to \chi.; \omega] \in q$, unde $\chi = x_1..x_n$ se face

$$\begin{aligned} R = R \cup \{q_1..q_n q\omega \to q_1 q'\omega | [X \to .\chi; \omega] \in q_1, \\ q_{i+1} = next(q_i, x_i)(i \in 1..n-1), \\ q = next(q_n, x_n), \\ q' = next(q_1, X)\} \end{aligned}$$

Pt q_3: $H([F \to i.; \#])$

$$\begin{array}{lll} q_1 & [F \to .i; \#] \in q_0, q_4, q_5 & \\ q & q_3 = next(q_0, i),\ q_3 = next(q_4, i),\ q_3 = next(q_5, i) & \\ q' & q_2 = next(q_0, F) & q_0 q_3 \to q_0 q_2 \\ q' & q_2 = next(q_4, F) & q_4 q_3 \to q_4 q_2 \\ q' & q_7 = next(q_5, F) & q_5 q_3 \to q_5 q_7 \end{array}$$

in forma matriciala: scot q_3 de pe stiva, se aplica reducerea 4, starea urm e q_2

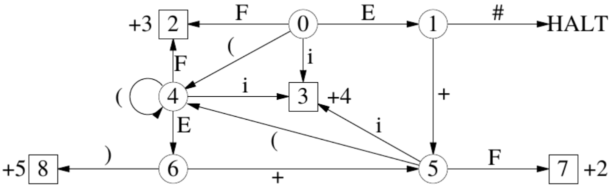
Tabel de tranzitii

$(1) Z \to E \quad (2) E \to E + F \quad (3) E \to F (4) F \to i \quad (5) F \to (E)$

	i	(	)	$+$	#	E	F
0	3	4	.	.	.	1	2
1	.	.	.	5	*		
2	+3	+3	+3	+3	+3		
3	+4	+4	+4	+4	+4		
4	3	4	.	.	.	6	2
5	3	4	.	.	.		7
6	.	.	8	5	.		
7	+2	+2	+2	+2	+2		
8	+5	+5	+5	+5	+5		

Diagrama de tranzitii

- suprapunerea gramaticilor regulate care corespund claselor stiva k
- starile prin care trece sunt inarcate pe stiva pana cand se ajunge intr-o stare finala
- in stare finala se face reducerea pe baza productiei $X \to \chi$, se elimina $|\chi|$ stari de pe stiva si se continua ca si cand s-ar fi citit simbolul X



(1)$Z \to E$ (2)$E \to E + F$ (3)$E \to F$(4)$F \to i$ (5)$F \to (E)$

Stiva	Derivare dreapta	simbol urmator	reducere cu productia	stare urmatoare
0	.i+(i+i)#	i		3
0 3	i.+(i+i)#		4	2
0 2	F.+(i+i)#		3	1
0 1	E.+(i+i)#	+		5
0 1 5	E+.(i+i)#	(		4
0 1 5 4	E+(.i+i)#	i		3
0 1 5 4 3	E+(i.+i)#		4	2
0 1 5 4 2	E+(F.+i)#		3	6
0 1 5 4 6	E+(E.+i)#	+		5
0 1 5 4 6 5	E+(E+.i)#	i		3
0 1 5 4 6 5 3	E+(E+i.)#		4	7
0 1 5 4 6 5 7	E+(E+F.)#		2	6
0 1 5 4 6	E+(E.)#	)		8
0 1 5 4 6 8	E+(E).#		5	7
0 1 5 7	E+F.#		2	1
0 1	E.#			

Stiva	Derivare dreapta	simbol urmator	reducere cu productia	stare urmatoare
0	.i+(i+i)#	i		3
0 3	i.+(i+i)#		4	2
0 2	F.+(i+i)#		3	1
0 1	E.+(i+i)#	+		5
0 1 5	E+.(i+i)#	(		4
0 1 5 4	E+(.i+i)#	i		3
0 1 5 4 3	E+(i.+i)#		4	2
0 1 5 4 2	E+(F.+i)#		3	6
0 1 5 4 6	E+(E.+i)#	+		5
0 1 5 4 6 5	E+(E+.i)#	i		3
0 1 5 4 6 5 3	E+(E+i.)#		4	7
0 1 5 4 6 5 7	E+(E+F.)#		2	6
0 1 5 4 6	E+(E.)#	)		8
0 1 5 4 6 8	E+(E).#		5	7
0 1 5 7	E+F.#		2	1
0 1	E.#			

Singura distinctie dintre modul de functionare a unui derivator (parser) LR(K) cu $k > 0$ si derivatorul LR(0) este faptul ca reducerile pot depinde de simboluri lookahead: in starile finale reducerile se realizeaza doar daca contextul permite

Non LR(0)

$$(1) Z \to E \quad (2) E \to E + T \quad (3) E \to T$$
$$(4) T \to T * F \quad (5) T \to F \quad (6) F \to i \quad (7) F \to (E)$$

De ce nu este LR(0)? daca se aplica LR(0): in starea 2 si stare 9 se poate aplica atat reducere cat si shift

Stare (q)	Situaþie	v	$f(q, v)$
0 *	$[Z \to .E]$	E	1
	$[E \to .E+T]$		
	$[E \to .T]$	T	2
	$[T \to .T*F]$		
	$[T \to .F]$	F	3
	$[F \to .i]$	i	4
	$[F \to .(E)]$	(	5
1 *	$[Z \to E.]$	#	STOP
*	$[E \to E.+T]$	+	6
2 *	$[E \to T.]$	#,), +	reducere 3
*	$[T \to T.*F]$	*	7
3 *	$[T \to F.]$		reducere 5
4 *	$[F \to i.]$		reducere 6

Non LR(0)

$$(1) Z \to E \quad (2) E \to E + T \quad (3) E \to T$$
$$(4) T \to T * F \quad (5) T \to F \quad (6) F \to i \quad (7) F \to (E)$$

	i	(	)	+	*	#	E	T	F
0	4	5	.	.	.	.	1	2	3
1	.	.	.	6	.	*			
2	.	.	+3	+3	7	+3			
3	+5	+5	+5	+5	+5	+5			
4	+6	+6	+6	+6	+6	+6			
5	4	5	.	.	.	.	8	2	3
6	4	5	.	.	.	.		9	3
7	4	5	.	.	.	.			10
8	.	.	11	6	.	.			
9	.	.	+2	+2	7	+2			
10	+4	+4	+4	+4	+4	+4			
11	+7	+7	+7	+7	+7	+7			

Right derivation before transition	Stack	Next Symbol	Reduce by Production	Next State
$.i + i * (i + i)\#$	0	i		4
$i. + i * (i + i)\#$	0,4		6	3
$F. + i * (i + i)\#$	0,3		5	2
$T. + i * (i + i)\#$	0,2	$+$	3	1
$E. + i * (i + i)\#$	0,1	$+$		6
$E + .i * (i + i)\#$	0,1,6	i		4
$E + i. * (i + i)\#$	0,1,6,4		6	3
$E + F. * (i + i)\#$	0,1,6,3		5	9
$E + T. * (i + i)\#$	0,1,6,9	$*$		7
$E + T * .(i + i)\#$	0,1,6,9,7	(		5
$E + T * (.i + i)\#$	0,1,6,9,7,5	i		4
$E + T * (i. + i)\#$	0,1,6,9,7,5,4		6	3
$E + T * (F. + i)\#$	0,1,6,9,7,5,3		5	2
$E + T * (T. + i)\#$	0,1,6,9,7,5,2	$+$	3	8
$E + T * (E. + i)\#$	0,1,6,9,7,5,8	$+$		6
$E + T * (E + .i)\#$	0,1,6,9,7,5,8,6	i		4
$E + T * (E + i.)\#$	0,1,6,9,7,5,8,6,4		6	3
$E + T * (E + F.)\#$	0,1,6,9,7,5,8,6,3		5	9
$E + T * (E + T.)\#$	0,1,6,9,7,5,8,6,9	)	2	8
$E + T * (E.)\#$	0,1,6,9,7,5,8	)		11
$E + T * (E).\#$	0,1,6,9,7,5,8,11		7	10
$E + T * F.\#$	0,1,6,9,7,10		4	9
$E + T.\#$	0,1,6,9	$\#$	2	1
$E.\#$	0,1	$\#$		HALT

- coloana next symbol este libera cand derivatorul nu examineaza simbol inainte
- exemplul arata ca un derivator LR(0)poate fi totusi folosit pt o gramatica non LR(0) prin considerarea ocazionala a simbolului lookahead
- starile in care lookahead este necesar: stari inadecvate: sunt caracterizate de $[X \to \chi.]$ + inca o situatie

Gramatici SLR(1)

O gramatica independenta de context $G = (T, N, P, Z)$ este SLR(1) daca algoritmul SLR(1) conduce la un automat stiva determinist.

Automatul stiva $A(T, Q, R, q_0, \{q_0\}, Q, q_0)$ se poate defini prin functia $f(q, \nu)$ in locul productiilor R. Algoritmul e similar LR(k) cu inchiderea:

$$H(M) = M \cup \{[Y \to .\beta] | \exists [X \to \mu.Y\gamma] \in H(M), Y \to \beta \in P\}$$

Algoritm SLR(1)

1. $Q = \{q_0\}$ cu $q_0 = H([Z \to .S])$
2. pt orice $q \in Q$ se efectueaza pasii 3-4 pt fiecare $\nu \in V$
3. fie $basis(q, \nu) = \{[X \to \mu\nu.\gamma] | [X \to \mu.\nu\gamma] \in q\}$
4. daca $basis(q, \nu) \neq \emptyset$ atunci $next(q, \nu) = H(basis(q, \nu))$. Se include $q' = next(q, \nu)$ in Q
5. daca $basis(q, \nu) \neq \emptyset$ si $\nu \in T$ se actualizeaza
6. daca toate elementele lui Q au fost tratate se executa pasul 6 pt fiecare $q \in Q$ si alg se termina; altfel se continua pasul 2
7. pentru fiecare $\nu \in V$ se defineste $f(q, \nu)$

$$f(q, \nu) = \begin{cases} next(q, \nu) & daca\ [X \to \mu.\nu\gamma \in q \\ X \to \chi & daca\ [X \to \chi.] \in q\ si\ \nu \in FOLLOW(X) \\ STOP & daca\ \nu = \#\ si\ [Z \to S.; \#] \in q \\ ERROR & altfel \end{cases}$$

Observatie: este aproape identic cu algoritmul LR(k), cu k = 0, singura diferenta fiind restrictia aditionala $\nu \in FOLLOW(X)$ in cazul al doilea, al reducerii.

SLR(1)

$$f(q,\nu) = \begin{cases} next(q,\nu) & daca\ [X \to \mu.\nu\gamma] \in q \\ X \to \chi & daca\ [X \to \chi.] \in q\ si\ \nu \in FOLLOW(X) \\ STOP & daca\ \nu = \#\ si\ [Z \to S.;\#] \in q \\ ERROR & altfel \end{cases}$$

vs forma matriciala a LR(k):

$$f(q,\nu) = \begin{cases} q' & daca\ \nu\gamma \in T^*\ si\ q\nu\gamma \to qq'\gamma \in R\ sau \\ & daca\ \nu \in N\ si\ q' = next(q,\nu) \\ X \to \chi & daca\ [X \to \chi.;\nu] \in q \\ STOP & daca\ \nu = \#\ si\ [Z \to S.;\#] \in q \\ ERROR & altfel \end{cases}$$

- gramaticile SLR(1) acopera multe constructii importante ale limbajelor ce nu pot fi exprimate in LR(0)
- comparativ cu LR(1), automatul SLR(1) are mai putine stari: pt exemplul anterior, raportul este 22:12

Comparatie derivatoare

Complexitatea computationala a generarii derivatorului

Tip gramatica	Test	Generare derivator
LL(1)	n^2	n^2
LL(k) tare	n^{k+1}	n^{k+1}
LL(k)	n^{2k}	$2^{n^k+(k+1)log\ \ n}$
SLR(1)	n^2	$2^{n+log\ \ n}$
SLR(k)	n^{k+2}	$2^{n+klog\ \ n}$
LR(k)	$n^{2(k+1)}$	$2^{n^{k+1}+klogn}$

n suma lungimilor partilor dreapta ale tuturor productiilor

- Instrumente pentru reprezentare
  - Siruri de rescriere
  - Gramatici - ierarhia lui Chomsky
  - Derivari si arbori de derivare
- Gramatici regulate si automate finite
  - Automate finite
  - Diagrame de stare si expresii regulate
- Gramatici independente de context si automate stiva: Automate stiva
- Analiza sintatica descendenta:
  - LL(k)
  - eliminare recursivitate stanga
  - Factorizare stanga
  - gramatici LL(k) tari
  - Derivator LL(1) - segmente de program
- Analiza sintatica ascendenta
  - LR(k)
  - Derivator LR(0) - functia de tranzitie
  - SLR(1)