Грамматики вида LL(1)

Романенко Владимир Васильевич, к.т.н., доцент каф. АСУ ТУСУР

Вид таблицы разбора LL(1)

```
declare 1 TABLE,
          2 terminals LIST,
          2 jump int,
          2 accept bool,
          2 stack bool,
          2 return bool,
          2 error bool;
где
declare 1 LIST,
          2 term string,
          2 next pointer;
```

Nº	terminals	jump	accept	stack	return	error

• Поиск символов-предшественников (START) $a \in S(\alpha) \Leftrightarrow \alpha \Rightarrow^* a\beta$,

где:

a – терминал или пустая цепочка, $a \in \Sigma \cup \{e\}$;

α и β – произвольные цепочки терминалов и/или нетерминалов, α, $β ∈ (N ∪ Σ)^*$;

 $S(\alpha)$ — множество символов-предшественников цепочки α .

• Поиск символов-предшественников (START) Пусть $\alpha = X_1 X_2 ... X_n$, где $X_i \in N \cup \Sigma \cup \{e\}$. Тогда:

$$S(\alpha) = \bigcup_{i=1}^{k} (S(X_i) - \{e\}) \cup \Delta,$$

где:
$$k = \begin{cases} j \mid e \notin S(X_j) \land e \in S(X_i), i < j, \\ n \mid e \in S(X_i), i = 1, 2, ..., n, \end{cases}$$

$$\Delta = \begin{cases} \{e\} \mid e \in S(X_i), i = 1, 2, ..., n, \\ \emptyset \mid e \notin S(X_j) \land e \in S(X_i), i \neq j. \end{cases}$$

5

$$A \rightarrow B C D F$$

$$B \rightarrow 1 \mid 2 \mid e$$

$$C \rightarrow a \mid b \mid e$$

$$D \rightarrow x$$

$$F \rightarrow y$$

• Поиск символов-предшественников (START) Частные случаи:

 $-S(a\beta) = \{a\}$, где $a \in \Sigma$, а β − произвольная цепочка;

 $-S(e) = \{e\}.$

• Поиск символов-предшественников (START)

Алгоритм:

- 1. Для всех правил $(A \rightarrow \alpha) \in P$ положить $S(A) = \emptyset$.
- 2. Для каждого правила $(A \rightarrow \alpha) \in P$

$$S(A) = S(A) \cup S(\alpha)$$
.

При этом
$$S(X_i) = \begin{cases} \{X_i\} \mid X_i \in \Sigma \cup \{e\}, \\ \bigcup_j S(X_j) \mid X_i \in N \land (X_j \rightarrow \beta) \in P \land X_j = X_i. \end{cases}$$

3. Если были изменения – вернуться на шаг 2.

8

Пример:

$$E \rightarrow TE'$$
 $E' \rightarrow + TE'$
 $E' \rightarrow e$
 $T \rightarrow FT'$
 $T' \rightarrow *FT'$
 $T' \rightarrow e$
 $F \rightarrow (E)$

 $F \rightarrow a$



• Поиск последующих символов (FOLLOW) $a \in F(A) \Leftrightarrow \alpha A\beta \Rightarrow^* \alpha Aa\gamma$,

где:

a — терминал или признак конца цепочки, $a \in \Sigma \cup \{\bot\};$

α, β и γ — произвольные цепочки терминалов и/или нетерминалов, α, β, $γ ∈ (N ∪ Σ)^*$;

A – нетерминал, $A \in \mathbb{N}$;

F(A) — множество последующих символов для нетерминала A.



$$A \rightarrow B C D F$$

$$B \rightarrow 1 \mid 2 \mid e$$

$$C \rightarrow a \mid b \mid e$$

$$D \rightarrow x$$

$$F \rightarrow y$$

11

• Поиск последующих символов (FOLLOW) Для правила грамматики вида $B \to \alpha A \beta$

$$F(A) = (S(\beta) - \{e\}) \cup \begin{cases} F(B) \mid \beta = e \lor e \in S(\beta), \\ \emptyset \mid \beta \neq e \land e \notin S(\beta). \end{cases}$$



• Поиск последующих символов (FOLLOW)

Алгоритм:

- 1. Для всех нетерминалов $A \in N$ положить $F(A) = \emptyset$. Для стартового нетерминала положить $F(S) = \{\bot\}$.
- 2. Для каждого вхождения нетерминала A в правую часть порождающих правил грамматики вида ($B \to \alpha A\beta$) $\in P$ добавить к множеству F(A) новые элементы.
- 3. Если были изменения вернуться на шаг 2.



$$E \rightarrow TE'$$

$$E' \rightarrow + TE'$$

$$E' \rightarrow e$$

$$T \rightarrow F T'$$

$$T' \rightarrow F'$$

$$T' \rightarrow e$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

Поиск направляющих символов



• Поиск направляющих символов

Для нетерминалов в левой части правила:

$$T(A) = (S(A) - \{e\}) \cup \begin{cases} F(A) \mid e \in S(A), \\ \emptyset \mid e \notin S(A). \end{cases}$$

Для нетерминалов в правой части правила:

$$T(A) = \begin{cases} \left(S(A\beta) - \{e\} \right) \cup F(B) \mid e \in S(A\beta), \\ S(A\beta) \mid e \notin S(A\beta). \end{cases}$$

Поиск направляющих символов



$$E \rightarrow TE'$$

$$E' \rightarrow + TE'$$

$$E' \rightarrow e$$

$$T \rightarrow F T'$$

$$T' \rightarrow F'$$

$$T' \rightarrow e$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

Поиск направляющих символов

16

• Поиск направляющих символов

Проверка грамматики:

$$A_1 \to \alpha_1$$

$$A_2 \to \alpha_2$$
.....
$$A_n \to \alpha_n$$

где $A_1=A_2=...=A_n$, тогда $T(A_i) {\smallfrown} T(A_j)=\varnothing \text{ при } i\neq j.$

17

1. Разметить грамматику. Порядковые номера $i \in M$ присваиваются всем элементам грамматики, от первого правила к последнему, от левого символа к правому. При наличии у порождающего правила альтернатив сначала нумеруются левые части всех альтернативных правил, а уже затем их правые части. Еще одним важным требованием является то, что все альтернативные правила должны следовать друг за другом в списке правил.



$$E \rightarrow TE'$$

$$E' \rightarrow + TE'$$

$$E' \rightarrow e$$

$$T \rightarrow F T'$$

$$T' \rightarrow F'$$

$$T' \rightarrow e$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

19

2. Построить два вспомогательных

множества $M_L \subset M$ и $M_R \subset M$. В первое включить порядковые номера $i \in M$ элементов грамматики, расположенных в левой части порождающего правила, во второе — порядковые номера элементов, которыми заканчиваются правые части порождающих правил:

$$M_L = \{i \mid (X_i \rightarrow \alpha) \in P\},$$

$$M_R = \{j \mid (A \rightarrow \alpha X_j) \in P\}.$$



$$\begin{split} E_{1} &\to T_{2} \, E'_{3} \\ E'_{4} &\to +_{6} \, T_{7} \, E'_{8} \\ E'_{5} &\to e_{9} \\ T_{10} &\to F_{11} \, T'_{12} \\ T'_{13} &\to ^{*}_{15} \, F_{16} \, T'_{17} \\ T'_{14} &\to e_{18} \\ F_{19} &\to (_{21} \, E_{22} \,)_{23} \\ F_{20} &\to a_{24} \end{split}$$

21

3. Построить таблицу, состоящую из столбцов **terminals**, **jump**, **accept**, **stack**, **return**, **error**. Количество строк таблицы определяется количеством элементов во множестве M — по одной строке на каждый элемент грамматики.

Nº	terminals	jump	accept	stack	return	error



4. Заполнить ячейки таблицы

4.1. Определение множества направляющих символов **terminals**:

$$\mathbf{terminals}_{i} = \begin{cases} T(X_{i}) | X_{i} \in N, \\ \left\{X_{i}\right\} | X_{i} \in \Sigma, \\ T(X_{j}) | X_{i} = e \land (X_{j} \rightarrow X_{i}) \in P. \end{cases}$$



$$E_{1} \rightarrow T_{2} E'_{3}$$

$$E'_{4} \rightarrow +_{6} T_{7} E'_{8}$$

$$E'_{5} \rightarrow e_{9}$$

$$T_{10} \rightarrow F_{11} T'_{12}$$

$$T'_{13} \rightarrow *_{15} F_{16} T'_{17}$$

$$T'_{14} \rightarrow e_{18}$$

$$F_{19} \rightarrow (_{21} E_{22})_{23}$$

$$F_{20} \rightarrow a_{24}$$

terminals_i =
$$\begin{cases} T(X_i) | X_i \in N, \\ \{X_i\} | X_i \in \Sigma, \\ T(X_j) | X_i = e \land (X_j \to X_i) \in P. \end{cases}$$

$$T(A) = (S(A) - \{e\}) \cup \begin{cases} F(A) | e \in S(A), \\ \emptyset | e \notin S(A). \end{cases}$$

$$T(E_1) = \{ (, a \} \}$$

$$T(E'_4) = \{ + \}$$

$$T(E'_5) = \{), \bot \}$$

$$T(T_{10}) = \{ (, a \} \}$$

$$T(T'_{13}) = \{ * \}$$

$$T(F_{19}) = \{ (\} \}$$



$$\begin{split} E_{1} &\to T_{2} \, E'_{3} \\ E'_{4} &\to +_{6} \, T_{7} \, E'_{8} \\ E'_{5} &\to e_{9} \\ T_{10} &\to F_{11} \, T'_{12} \\ T'_{13} &\to ^{*}_{15} \, F_{16} \, T'_{17} \\ T'_{14} &\to e_{18} \\ F_{19} &\to (_{21} \, E_{22} \,)_{23} \\ F_{20} &\to a_{24} \end{split}$$

$$\mathbf{terminals}_{i} = \begin{cases} T(X_{i}) \mid X_{i} \in N, \\ \left\{X_{i}\right\} \mid X_{i} \in \Sigma, \\ T(X_{j}) \mid X_{i} = e \land \left(X_{j} \rightarrow X_{i}\right) \in P. \end{cases}$$

$$T(A) = \begin{cases} \left(S(A\beta) - \left\{e\right\}\right) \cup F(B) \mid e \in S(A\beta), \\ S(A\beta) \mid e \notin S(A\beta). \end{cases}$$



4. Заполнить ячейки таблицы

4.2. Определение номера строки для перехода **jump**:

$$\mathbf{jump}_{i} = \begin{cases} k \mid i \in M_{L} \land \left(X_{i} \rightarrow X_{k} \alpha\right) \in P, \\ k \mid X_{i} \in N \land i \not\in M_{L} \land X_{k} = X_{i} \land k \in M_{L} \land k - 1 \not\in M_{L}, \\ i + 1 \mid i \in \Sigma \land i \not\in M_{R}, \\ 0 \mid i \in \Sigma \cup \{e\} \land i \in M_{R}. \end{cases}$$

(26)

$$E_{1} \rightarrow T_{2} E'_{3}$$

$$E'_{4} \rightarrow +_{6} T_{7} E'_{8}$$

$$E'_{5} \rightarrow e_{9}$$

$$T_{10} \rightarrow F_{11} T'_{12}$$

$$T'_{13} \rightarrow *_{15} F_{16} T'_{17}$$

$$T'_{14} \rightarrow e_{18}$$

$$F_{19} \rightarrow (_{21} E_{22})_{23}$$

$$F_{20} \rightarrow a_{24}$$

$$\mathbf{jump}_{i} = \begin{cases} k \mid i \in M_{L} \land \left(X_{i} \rightarrow X_{k} \alpha\right) \in P, \\ k \mid X_{i} \in N \land i \notin M_{L} \land X_{k} = X_{i} \land k \in M_{L} \land k - 1 \notin M_{L}, \\ i + 1 \mid i \in \Sigma \land i \notin M_{R}, \\ 0 \mid i \in \Sigma \cup \left\{e\right\} \land i \in M_{R}. \end{cases}$$

27

4. Заполнить ячейки таблицы

4.3. Определение флага приема текущего символа **accept**:

$$\mathbf{accept}_i = \begin{cases} \text{true} \mid X_i \in \Sigma, \\ \text{false} \mid X_i \notin \Sigma. \end{cases}$$



$$E_{1} \rightarrow T_{2} E'_{3}$$

$$E'_{4} \rightarrow +_{6} T_{7} E'_{8}$$

$$E'_{5} \rightarrow e_{9}$$

$$T_{10} \rightarrow F_{11} T'_{12}$$

$$T'_{13} \rightarrow *_{15} F_{16} T'_{17}$$

$$T'_{14} \rightarrow e_{18}$$

$$F_{19} \rightarrow (_{21} E_{22})_{23}$$

$$F_{20} \rightarrow a_{24}$$

$$\mathbf{accept}_i = \begin{cases} \text{true} \mid X_i \in \Sigma, \\ \text{false} \mid X_i \notin \Sigma. \end{cases}$$



4. Заполнить ячейки таблицы

4.4. Определение флага помещения текущего состояния в стек **stack**:

$$\mathbf{stack}_{i} = \begin{cases} \text{true} \mid X_{i} \in N \land i \notin M_{L} \land i \notin M_{R}, \\ \text{false} \mid X_{i} \notin N \lor i \in M_{L} \lor i \in M_{R}. \end{cases}$$



$$E_{1} \to T_{2} E'_{3}$$

$$E'_{4} \to +_{6} T_{7} E'_{8}$$

$$E'_{5} \to e_{9}$$

$$T_{10} \to F_{11} T'_{12}$$

$$T'_{13} \to *_{15} F_{16} T'_{17}$$

$$T'_{14} \to e_{18}$$

$$F_{19} \to (_{21} E_{22})_{23}$$

$$F_{20} \to a_{24}$$

$$\mathbf{stack}_{i} = \begin{cases} \text{true} \mid X_{i} \in N \land i \notin M_{L} \land i \notin M_{R}, \\ \text{false} \mid X_{i} \notin N \lor i \in M_{L} \lor i \in M_{R}. \end{cases}$$



4. Заполнить ячейки таблицы

4.5. Определение флага возврата к предыдущему состоянию **return**:

$$\mathbf{return}_{i} = \begin{cases} \text{true} \middle| i \in \Sigma \cup \{e\} \land i \in M_{R}, \\ \text{false} \middle| i \notin \Sigma \cup \{e\} \lor i \notin M_{R}. \end{cases}$$



$$E_{1} \rightarrow T_{2} E'_{3}$$

$$E'_{4} \rightarrow +_{6} T_{7} E'_{8}$$

$$E'_{5} \rightarrow e_{9}$$

$$T_{10} \rightarrow F_{11} T'_{12}$$

$$T'_{13} \rightarrow *_{15} F_{16} T'_{17}$$

$$T'_{14} \rightarrow e_{18}$$

$$F_{19} \rightarrow (_{21} E_{22})_{23}$$

$$F_{20} \rightarrow a_{24}$$

$$\mathbf{return}_{i} = \begin{cases} \text{true} \ \middle| \ i \in \Sigma \cup \{e\} \land i \in M_{R} \ , \\ \text{false} \ \middle| \ i \notin \Sigma \cup \{e\} \lor i \notin M_{R} . \end{cases}$$



4. Заполнить ячейки таблицы

4.5. Определение флага ошибки **error**:

$$\mathbf{error}_{i} = \begin{cases} \text{false} \mid i \in M_{L} \land i + 1 \in M_{L}, \\ \text{true} \mid i \notin M_{L} \lor i + 1 \notin M_{L}. \end{cases}$$



$$E_{1} \rightarrow T_{2} E'_{3}$$
 $E'_{4} \rightarrow +_{6} T_{7} E'_{8}$
 $E'_{5} \rightarrow e_{9}$
 $T_{10} \rightarrow F_{11} T'_{12}$
 $T'_{13} \rightarrow *_{15} F_{16} T'_{17}$
 $T'_{14} \rightarrow e_{18}$
 $F_{19} \rightarrow (_{21} E_{22})_{23}$
 $F_{20} \rightarrow a_{24}$

$$\mathbf{error}_{i} = \begin{cases} \text{false} \mid i \in M_{L} \land i + 1 \in M_{L}, \\ \text{true} \mid i \notin M_{L} \lor i + 1 \notin M_{L}. \end{cases}$$

35

i	X	terminals	jump	accept	stack	return	error
1	E	(, a	2				
2	T	(, a	10		true		
3	E'	+,), ⊥	4				
4	E'	+	6				false
5	E'), ⊥	9				
6	+	+	7	true			
7	T	(, a	10		true		
8	E'	+,), ⊥	4				
9	e), ⊥	0			true	
10	T	(, a	11				
11	F	(, a	19		true		
12	T'	+, *,), ⊥ *	13				
13	T'		15				false
14	T'	+ ,), ⊥	18				
15	*	*	16	true			
16	F	(, a	19		true		
17	T'	+, *,), ⊥	13				
18	e	+,), ⊥	0			true	2.1
19	F	(21				false
20	F	a	24				
21	(22	true			
22	E	(, a	1		true		
23)	J	0	true		true	
24	a	a	0	true		true	



Исходные данные:

- таблица разбора;
- анализируемая цепочка $\alpha = a_1 a_2 ... a_n \bot$.

Драйвер (ядро компилятора):

- k текущий символ цепочки;
- M магазин (стек), для которого определена операция помещения символа в магазин ($M \leftarrow x$) и извлечения символа из магазина ($M \rightarrow x$);
- і текущая строка таблицы разбора.



Алгоритм работы драйвера:

- 1. Положить i := 1.
- 2. Положить k := 1.
- $3. M \leftarrow 0.$
- 4. Если $a_k \in \mathbf{terminals}_i$, то:
- 4.1. Если **accept**_i = true, то k := k + 1.
- 4.2. Если **stack** $_{i}$ = true, то $M \leftarrow i$.

38)

Алгоритм работы драйвера:

4.3. Если **return** $_i$ = true, то:

 $4.3.1. M \rightarrow i;$

4.3.2. Если i = 0, то перейти на шаг 6;

4.3.3.i := i + 1;

4.3.4. Вернуться на шаг 4.

4.4. Если **jump** $_{i} \neq 0$, то:

4.4.1. $i := \mathbf{jump}_i$;

4.4.2. Вернуться на шаг 4.

39

Алгоритм работы драйвера:

5. Иначе если **error** $_i$ = false, то:

- 5.1. i := i + 1;
- 5.2. Вернуться на шаг 4.
- 6. В противном случае разбор окончен. Если при этом стек M пуст, а $a_k = \bot$, то разбор завершен успешно. Иначе цепочка содержит синтаксическую ошибку и k позиция этой ошибки.



Пример: $\alpha = (a+a)*a\bot$.

i	M	a_ka_n
1	0	(a+a)*a⊥
2	0	(a+a)*a⊥
10	0 2	(a+a)*a⊥
11	0 2	(a+a)*a⊥
19	0 2 11	(a+a)*a⊥
21	0 2 11	(a+a)*a⊥
22	0 2 11	a+a)*a⊥
1	0 2 11 22	a+a)*a⊥
2	0 2 11 22	a+a)*a⊥
10	0 2 11 22 2	a+a)*a⊥
11	0 2 11 22 2	a+a)*a⊥
19	0 2 11 22 2 11	a+a)*a⊥
20	0 2 11 22 2 11	a+a)*a⊥
24	0 2 11 22 2 11	a+a)*a⊥

i	terminals	jump	accept	stack	return	error
1	(, a	2				
2	(, a	10		true		
3	+,), ⊥	4				
4	+	6				false
5), ⊥	9				
6	+	7	true			
7	(, a	10		true		
8	+,), ⊥	4				
9), ⊥	0			true	
10	(, a	11				
11	(, a	19		true		
	+, *,), ⊥	13				
13		15				false
	+,), ⊥	18				
15		16	true			
	(, a	19		true		
	+, *,), ⊥	13				
	+,), ⊥	0			true	
19	(21				false
20		24				
21	•	22	true			
	(, a	1		true		
23)	0	true		true	
24	a	0	true		true	



Пример: $\alpha = (a+a)*a\bot$.

i	M	a_ka_n
12	0 2 11 22 2	+a)*a⊥
13	0 2 11 22 2	+a)*a⊥
14	0 2 11 22 2	+a)*a⊥
18	0 2 11 22 2	+a)*a⊥
3	0 2 11 22	+a)*a⊥
4	0 2 11 22	+a)*a⊥
6	0 2 11 22	+a)*a⊥
7	0 2 11 22	a)*a⊥
10	0 2 11 22 7	a)*a⊥
11	0 2 11 22 7	a)*a⊥
19	0 2 11 22 7 11	a)*a⊥
20	0 2 11 22 7 11	a)*a⊥
24	0 2 11 22 7 11	a)*a⊥
12	0 2 11 22 7)*a⊥

i	terminals	jump	accept	stack	return	error
1	(, a	2				
	(, a	10		true		
3	+,), ⊥	4				
4	+	6				false
5), ⊥	9				
	+	7	true			
	(, a	10		true		
	+,), ⊥	4				
), ⊥	0			true	
	(, a	11				
	(, a	19		true		
12	$+, *,), \bot$	13				
13		15				false
	+,), ⊥	18				
15		16	true			
	(, a	19		true		
	+, *,), ⊥	13				
	+,), ⊥	0			true	
19		21				false
20		24				
21	-	22	true			
	(, a	1		true		
23)	0	true		true	
2 4	a	0	true		true	



Пример: $\alpha = (a+a)^*a\bot$.

i	M	a_ka_n
13	0 2 11 22 7)*a⊥
14	0 2 11 22 7)*a⊥
18	0 2 11 22 7)*a⊥
8	0 2 11 22)*a⊥
4	0 2 11 22)*a⊥
5	0 2 11 22)*a⊥
9	0 2 11 22)*a⊥
23	0 2 11)*a⊥
12	0 2	*a⊥
13	0 2	*a⊥
15	0 2	*a⊥
16	0 2	a⊥
19	0 2 16	a⊥
20	0 2 16	a⊥

i	terminals	jump	accept	stack	return	error
1	(, a	2				
2	(, a	10		true		
3	+,), ⊥	4				
4	+	6				false
), ⊥	9				
6	+	7	true			
7	(, a +,), ⊥	10		true		
8	+,), ⊥	4				
9), ⊥	0			true	
10	(, a	11				
11	(, a	19		true		
12	$+, *,), \bot$	13				
13		15				false
14	+,), ⊥	18				
15		16	true			
	(, a	19		true		
	+, *,), ⊥	13				
	+,), ⊥	0			true	
19	(21				false
20		24				
21		22	true			
	(, a	1		true		
23)	0	true		true	
24	a	0	true		true	



Пример: $\alpha = (a+a)*a\bot$.

i	M	a_ka_n
24	0 2 16	a⊥
17	0 2	Τ
13	0 2	Т
14	0 2	Τ
18	0 2	Τ
3	0	Τ
4	0	Τ
5	0	Τ
9	0	Τ
	HALT	

i	terminals	jump	accept	stack	return	error
1	(, a	2				
2	(, a	10		true		
3	+,), ⊥	4				
4	+),⊥	6				false
5), ⊥	9				
6	+	7	true			
7	(, a	10		true		
	+,), ⊥	4				
), ⊥	0			true	
	(, a	11				
	(, a	19		true		
12	+, *,), ⊥	13				
13		15				false
	+,), ⊥	18				
15		16	true			
	(, a	19		true		
	+, *,), ⊥	13				
	+,), ⊥	0			true	
19		21				false
20		24				
21		22	true			
	(, a	1		true		
23		0	true		true	
24	a	0	true		true	