

LL(k)-грамматики

SLL(1), $FIRST_k^G$, $FOLLOW_k^G$, LL(1), strong LL(1), LL(k)

В. С. Полозов

Кафедра системного программирования СПбГУ



Теория автоматов и формальных языков

Неформально

Цель: построить детерминированный разбор.

- Плюсы: скорость, однозначность
- Минусы: ограниченный набор языков (не допускаются неоднозначные)

Например, можно ограничить количество альтернатив для каждого нетерминала одной.

Неформально

Цель: построить детерминированный разбор.

- Плюсы: скорость, однозначность
- Минусы: ограниченный набор языков (не допускаются неоднозначные)

Например, можно ограничить количество альтернатив для каждого нетерминала одной.

BFS

Рассмотрим грамматики $G = (V_N, V_T, P, S)$ с правилами вида $A \rightarrow x\alpha, x \in V_T$.

Например:

$$\begin{aligned} S &\rightarrow aB \mid bA \\ A &\rightarrow a \mid aS \mid bAA \\ B &\rightarrow b \mid bS \mid aBB \end{aligned}$$

И нисходящий разбор поиском в ширину цепочки $aabb$.

Parse table

Построим *таблицу разбора* (управляющую таблицу):

	a	b	#
S	$S_1 : aB$	$S_2 : bA$	
A	$A_1 : a$ $A_2 : aS$	$A_3 : bAA$	
B	$B_3 : aBB$	$B_1 : b$ $B_2 : bS$	

Если в ячейке несколько записей - это поиск.

SLL(1)

Если в ячейке таблицы разбора не более одной записи - получим детерминированный разбор.

Назовем такую грамматику *простой* $LL(1)$ или $SLL(1)$.

Переформулируем:

Определение

Говорят, что контекстно-свободная грамматика G является простой $LL(1)$ -грамматикой, если в ней нет ϵ -правил, и все альтернативы для каждого нетерминала начинаются с терминалов и притом различных.

SLL(1)

Если в ячейке таблицы разбора не более одной записи - получим детерминированный разбор.

Назовем такую грамматику *простой $LL(1)$* или $SLL(1)$.

Переформулируем:

Определение

Говорят, что контекстно-свободная грамматика G является простой $LL(1)$ -грамматикой, если в ней нет ε -правил, и все альтернативы для каждого нетерминала начинаются с терминалов и притом различных.

SLL(1)

Например:

$$\begin{aligned} S &\rightarrow aB \\ B &\rightarrow b \mid aBb \end{aligned}$$

Таблица разбора:

	a	b	#
S	$S_1 : aB$		
B	$B_2 : aBb$	$B_1 : b$	

И вход: $aabb$ (с добавленным символом #)

SLL(1)

Например:

$$\begin{aligned} S &\rightarrow aB \\ B &\rightarrow b \mid aBb \end{aligned}$$

Таблица разбора:

	a	b	#
S	$S_1 : aB$		
B	$B_2 : aBb$	$B_1 : b$	

И вход: $aabb$ (с добавленным символом #)

SLL(1)

Например:

$$\begin{aligned} S &\rightarrow aB \\ B &\rightarrow b \mid aBb \end{aligned}$$

Таблица разбора:

	a	b	#
S	$S_1 : aB$		
B	$B_2 : aBb$	$B_1 : b$	

И вход: $aabb$ (с добавленным символом #)

Грамматики без ϵ -правил

Определим множество $FIRST_1^G$ ($FIRST_1$, $FIRST$):

Определение

$G = (V_N, V_T, P, S)$ - КС-грамматика

$$FIRST_1^G(\alpha) = \{w \in V_T \mid \alpha \xrightarrow[G]{*} wx \text{ для некоторой цепочки } x \in V_T^*\}$$

где $\alpha \in V^*$.

Вычисление $FIRST_1$

Вычисление:

- 1 Положим $FIRST(A) = \emptyset$ для $\forall A \in V_N$
- 2 Для каждого правила $A \rightarrow x\alpha, x \in V$
 - $x \in V_T \Rightarrow FIRST(A) \leftarrow FIRST(A) \cup x$
 - $x \in V_N \Rightarrow FIRST(A) \leftarrow FIRST(A) \cup FIRST(x)$
- 3 Повторить шаг 2, если какое-либо множество $FIRST$ изменилось

Грамматики без ϵ -правил: пример

Пример:

Session → *Fact Session*

Session → *Question*

Session → (*Session*) *Session*

Fact → ! *STRING*

Question → ? *STRING*

Посчитаем множества *FIRST* и таблицу анализа

Грамматики с ϵ -правилами

ϵ -правила удобны, например, для списков:

$Session \rightarrow Facts Question \mid (Session) Session$
 $Facts \rightarrow Fact Facts \mid \epsilon$
 $Fact \rightarrow ! STRING$
 $Question \rightarrow ? STRING$

Дополним множество $FIRST$ ϵ -цепочками

Определение

$$FIRST_1^G(\alpha) = \{w \in V_T^* \mid \text{либо } w = \epsilon \text{ и } \alpha \xrightarrow[G]{*} \epsilon, \\ \text{либо } |w| = 1 \text{ и } \alpha \xrightarrow[G]{*} wx \\ \text{для некоторой цепочки } x \in V_T^*\}.$$

Грамматики с ε -правилами

ε -правила удобны, например, для списков:

Session \rightarrow *Facts Question* | (*Session*) *Session*

Facts \rightarrow *Fact Facts* | ε

Fact \rightarrow ! *STRING*

Question \rightarrow ? *STRING*

Дополним множество *FIRST* ε -цепочками

Определение

$$FIRST_1^G(\alpha) = \{w \in V_T^* \mid \text{либо } w = \varepsilon \text{ и } \alpha \xrightarrow{*}_G \varepsilon,$$

$$\text{либо } |w| = 1 \text{ и } \alpha \xrightarrow{*}_G wx$$

для некоторой цепочки $x \in V_T^*\}$.

Множество *FOLLOW*

Определим множество $FOLLOW_1^G$ ($FOLLOW_1$, $FOLLOW$) как возможное множество терминальных символов после данной цепочки при выводе из стартового символа.

Определение

Для КС-грамматики

$G' = (V'_N = V_N \cup \{S'\}, V'_T = V_T \cup \{\#\}, P' = P \cup (S' \rightarrow S\#), S')$
построенной по $G = (V_N, V_T, P, S)$ и цепочки $\beta \in V^*$

$$FOLLOW_1^G(\beta) = \{w \in V_T \mid S \xRightarrow{*} \gamma\beta\alpha, w \in FIRST_1^G(\alpha)\}$$

Вычисление $FOLLOW_1$

Вычисление:

- 1 Положим $FOLLOW(A) = \emptyset$ для $\forall A \in V_N$
- 2 Для каждого правила $B \rightarrow \beta A \alpha$ где $\alpha, \beta \in V^*$, $A, B \in V'_N$
 - $FOLLOW(A) \leftarrow FOLLOW(A) \cup (FIRST(\alpha) \setminus \{\epsilon\})$
 - Если $\alpha \xrightarrow{*} \epsilon$ то $FOLLOW(A) \leftarrow FOLLOW(A) \cup FOLLOW(B)$
- 3 Повторить шаг 2, если какое-либо множество $FOLLOW$ изменилось

Сильный LL(1) анализатор

При построении таблицы анализатора для правила $A \rightarrow \alpha$:
 $\forall x \in FIRST(\alpha FOLLOW(A))$ положим ячейку $T(A, x) \leftarrow A \rightarrow \alpha$

Вопросы