Лекция 2

Ilya Yaroshevskiy

9 сентября

Содержание

1 $LL(1)$ грамматика

Лемма 1.

$$\alpha = c\beta \implies \text{FIRST}(\alpha) = \{c\}$$

$$\alpha = \varepsilon \implies \text{FIRST}(\alpha) = \{\varepsilon\}$$

$$\alpha = A\beta \implies \text{FIRST}(\alpha) = \text{FIRST}(\alpha) \setminus \varepsilon \cup (\text{FIRST}(\beta) \textit{ если } \varepsilon \in \text{FIRST}(A))$$

- *Массив* `FIRST[]: map<NonTerm, Set<Term1>>`: $N \to 2^{\Sigma \cup \{\varepsilon\}}$

```
while(change):
    for A → α ∈ Γ:
        FIRST[A] ∪= FIRST(α)
```

Доказательство. Исправить

$$\text{FIRST}[A] \subset \text{FIRST}(A) \implies c \in \text{FIRST}(A) \text{ и } c \notin \text{FIRST}[A]$$

Тогда $A \Rightarrow^k c\varepsilon$, $k \rightsquigarrow \min$. Предположим, что на k-той итерации ничего не поменялось, но мы прошли правило $A \Rightarrow^k \alpha \Rightarrow c\varepsilon$, **противоречие**. □

Лемма 2.

- `FOLLOW[]`

```
FOLLOW[S] = \{$\}
while(change):
    for A → α ∈ Γ:
        for β: α = βBγ:
            FOLLOW[B] ∪= FIRST(γ)\ε ∪ (FOLLOW(A) если ε ∈ FIRST(γ))
```

Пример.

$$\begin{aligned}
&E \to E + T \\
&E \to T \\
&T \to T \cdot E \\
&T \to F \\
&F \to (E) \\
&F \to n
\end{aligned}$$

	FIRST	FOLLOW
E	$($ n	$\$ +)$
T	$($ n	$\$ + \cdot)$
F	$($ n	$\$ + \cdot)$

Проверим, что $LL(1)$:

$$\text{FIRST}(E) \cap \text{FIRST}(T) = \{(, n\} \neq \varnothing$$

$\implies$ **не** $LL(1)$

Доказательство Теоремы. Пусть не $LL(1)$, тогда $\xi \neq \eta$

1. ξ не порождает ε, η не порождает $\varepsilon \implies c \in \text{FIRST}(\xi)$ и $c \in \text{FIRST}(\eta)$
2. ξ порождает ε, η порождает $\varepsilon \implies \varepsilon \in \text{FIRST}(\xi)$ и $\varepsilon \in \text{FIRST}(\eta)$
3. ξ порождает ε, η не порождает $\varepsilon \implies \varepsilon \in \text{FIRST}(\xi),\ c \in \text{FOLLOW}(A),\ c \in \text{FIRST}(\eta)$

□

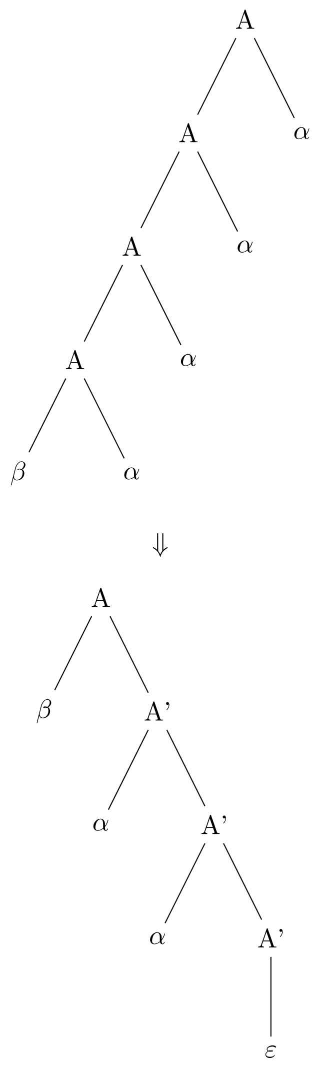
Определение. Γ — **леворекурсивная**, если

$$\exists A : A \Rightarrow^* A\alpha$$

Утверждение. *Леворекурсивная* $\implies$ *не* $LL(1)$ *(почти всегда)*

Примечание. Непосредственная левая рекурсия: $A \to A\alpha$. A не порождающий, значит есть $A \to \beta$.

- $c \in \text{FIRST}(\beta)$
- $c \in \text{FIRST}(A\alpha)$



$$
\begin{aligned}
&A \to \beta A' \\
&A' \to \alpha A' \\
&A' \to \varepsilon
\end{aligned}
$$

Определение. $A \to \alpha\beta$, $A \to \alpha\gamma$ — **правое ветвление**

Утверждение. *Правое ветвление* $\implies$ *не* $LL(1)$ *(почти всегда)*

$$
\begin{aligned}
&A \to \alpha A' \\
&A' \to \beta \\
&A' \to \gamma
\end{aligned}
$$

Пример. Перестроим грамматику для выражений:

$$
\begin{aligned}
&E \to TE' \\
&E \to +TE' \\
&E' \to \varepsilon \\
&T \to FT' \\
&T' \to \cdot FT' \\
&T' \to \varepsilon \\
&F \to (E) \\
&F \to n
\end{aligned}
$$

	FIRST	FOLLOW
E	n (	\$)
E'	ε +	\$)
T	n c	+ \$)
T'	ε $\cdot$	+ \$)
F	n c	+ $\cdot$ \$)

Это $LL(1)$ грамматика.

2 Конструирование парсеров

- Сверху (Top Down). Знаем корень, начинаем с него
- Снизу (Bottom Up). Знаем крону, начинаем с токенов

2.1 Рекурсивный спуск

Контекст:

- Текущий токен (`token`)
- `nextToken()` — следующий токен
- `FIRST1`$(\alpha) = \text{FIRST}(\alpha) \setminus \varepsilon \cup (\text{FOLLOW}(A)$ если $\varepsilon \in \text{FIRST}(\alpha))$

```
Tree A():
    r = Tree(A)
    switch(token):
        case FIRST1(α_1):
            // α_1
        case FIRST1(α_2):
            // α_2
        ⋮
        case FIRST1(α_j):
            for X_j : α_i = X_1 X_2 ...:
                // X_i = c - терминал
                ensure X_j = token
                nextToken()
                r.addChild(X_j)
                // X_j = A - нетерминал
                r.addChild(X_j())
```