

编译原理

laisg

2019 年 4 月 15 日

Part I

语法分析 _1

1 S_ 文法

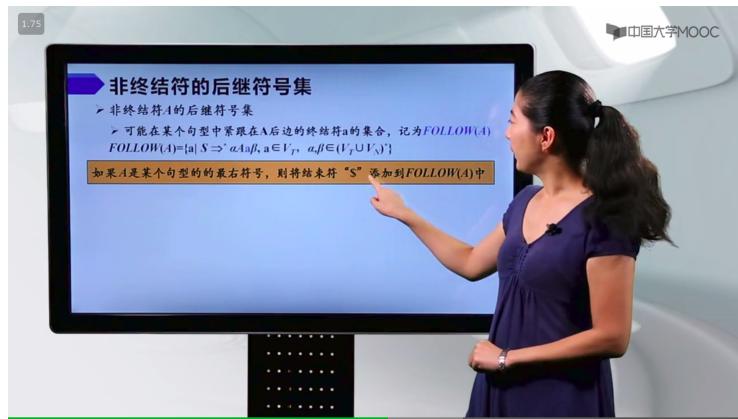
175 中国大学MOOC

S_文法

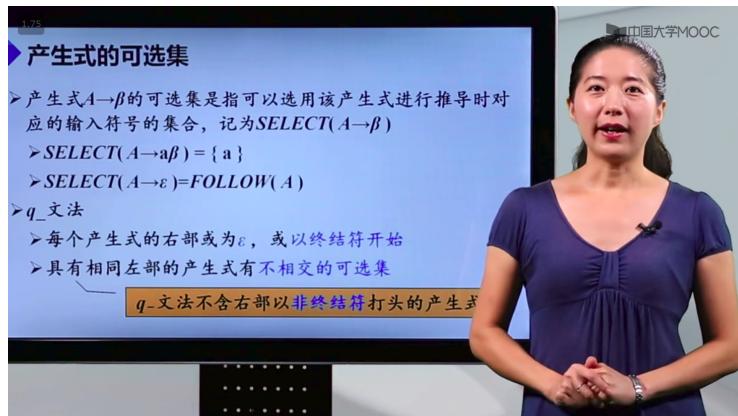
- 预测分析法的工作过程
- 从文法开始符号出发，在每一步推导过程中根据当前句型的最左非终结符 A 和当前输入符号 a ，选择正确的 A -产生式。为保证分析的确定性，选出的候选式必须是唯一的。
- S _文法（简单的确定性文法，Korenjak & Hopcroft, 1966）
 - 每个产生式的右部都以**终结符**开始
 - 同一非终结符的各个候选式的**首终结符**都不同

S _文法 不含 ϵ 产生式

2 FOLLOW 集



3 产生式的 SELECT 集【划重点】



4 FIRST 集

串首终结符集

- 串首终结符
 - 串首第一个符号，并且是终结符。简称首终结符
 - 给定一个文法符号串 α , α 的串首终结符集 $FIRST(\alpha)$ 被定义为可以从 α 推导出的所有串首终结符构成的集合。如果 $\alpha \Rightarrow^* \varepsilon$, 那么 ε 也在 $FIRST(\alpha)$ 中
 - 对于 $\forall \alpha \in (V_T \cup V_N)^*$, $FIRST(\alpha) = \{ a \mid \alpha \Rightarrow^* a\beta, a \in V_T, \beta \in (V_T \cup V_N)^* \}$
 - 如果 $\alpha \Rightarrow^* \varepsilon$, 那么 $\varepsilon \in FIRST(\alpha)$
 - 产生式 $A \rightarrow a$ 的可选集 $SELECT$
 - 如果 $\varepsilon \notin FIRST(\alpha)$, 那么 $SELECT(A \rightarrow a) = FIRST(\alpha)$
 - 如果 $\varepsilon \in FIRST(\alpha)$, 那么 $SELECT(A \rightarrow a) = (FIRST(\alpha) - \{\varepsilon\}) \cup FOLLOW(A)$

5 LL (1) 文法

LL(1)文法

- 文法 G 是LL(1)的，当且仅当 G 的任意两个具有相同左部的产生式 $A \rightarrow \alpha \mid \beta$ 满足下面的条件：
 - 不存在终结符 a 使得 α 和 β 都能够推导出以 a 开头的串
 - α 和 β 至多有一个能推导出 ε
 - 如果 $\beta \Rightarrow^* \varepsilon$, 则 $FIRST(\alpha) \cap FOLLOW(\beta) = \emptyset$;
如果 $\alpha \Rightarrow^* \varepsilon$, 则 $FIRST(\alpha) \cap FOLLOW(\beta) = \emptyset$;

Part II

语法分析 _2

6 FIRST 集和 FOLLOW 集的计算

6.1 FIRST 集的计算

算法

- 不断应用下列规则，直到没有新的终结符或 ε 可以被加入到任何FIRST集合中为止
- 如果 X 是一个终结符，那么 $FIRST(X) = \{X\}$
- 如果 X 是一个非终结符，且 $X \rightarrow Y_1 \dots Y_k \in P$ ($k \geq 1$)，那么如果对于某个 i , a 在 $FIRST(Y_i)$ 中且 ε 在所有的 $FIRST(Y_1), \dots, FIRST(Y_{i-1})$ 中(即 $Y_1 \dots Y_{i-1} \Rightarrow^* \varepsilon$)，就把 a 加入到 $FIRST(X)$ 中。如果对于所有的 $j = 1, 2, \dots, k$, ε 在 $FIRST(Y_j)$ 中，那么将 ε 加入到 $FIRST(X)$
- 如果 $X \rightarrow \varepsilon \in P$ ，那么将 ε 加入到 $FIRST(X)$ 中

计算串 $X_1X_2\dots X_n$ 的FIRST集合

- 向 $FIRST(X_1X_2\dots X_n)$ 加入 $FIRST(X_1)$ 中所有的非 ε 符号
- 如果 ε 在 $FIRST(X_1)$ 中，再加入 $FIRST(X_2)$ 中的所有非 ε 符号；如果 ε 在 $FIRST(X_1)$ 和 $FIRST(X_2)$ 中，再加入 $FIRST(X_3)$ 中的所有非 ε 符号，以此类推
- 最后，如果对所有的 i , ε 都在 $FIRST(X_i)$ 中，那么将 ε 加入到 $FIRST(X_1X_2\dots X_n)$ 中

6.2 FOLLOW 集的计算

计算非终结符 A 的 $FOLLOW(A)$

- $FOLLOW(A)$: 可能在某个句型中紧跟在 A 后边的终结符 a 的集合
 $FOLLOW(A) = \{a | S \xrightarrow{*} \alpha A \beta, a \in V_T, \alpha, \beta \in (V_T \cup V_N)^*\}$
- 如果 A 是某个句型的最右符号，则将结束符“ $\$$ ”添加到 $FOLLOW(A)$ 中

例

- ① $E \rightarrow TE'$ $FIRST(E) = \{(\text{id})\}$ $FOLLOW(E) = \{\$ \}$
- ② $E' \rightarrow +TE'|\varepsilon$ $FIRST(E') = \{+\varepsilon\}$ $FOLLOW(E') = \{\$ \}$
- ③ $T \rightarrow FT'$ $FIRST(T) = \{(\text{id})\}$ $FOLLOW(T) = \{+\$ \}$
- ④ $T' \rightarrow *FT'|\varepsilon$ $FIRST(T') = \{* \varepsilon\}$ $FOLLOW(T') = \{+\$ \}$
- ⑤ $F \rightarrow (E) \mid \text{id}$ $FIRST(F) = \{(\text{id})\}$ $FOLLOW(F) = \{* + \$ \}$

算法

- 不断应用下列规则，直到没有新的终结符可以被加入到任何 $FOLLOW$ 集合中为止
- 将 S 放入 $FOLLOW(S)$ 中，其中 S 是开始符号， S 是输入右端的结束标记
- 如果存在一个产生式 $A \rightarrow aB\beta$ ，那么 $FIRST(\beta)$ 中除 ε 之外的所有符号都在 $FOLLOW(B)$ 中
- 如果存在一个产生式 $A \rightarrow aB$ ，或存在产生式 $A \rightarrow aB\beta$ 且 $FIRST(\beta)$ 包含 ε ，那么 $FOLLOW(A)$ 中包含 $FOLLOW(B)$ 中

7 SELECT 集的计算

产生式的可选集

- 产生式 $A \rightarrow \beta$ 的可选集是指可以选用该产生式进行推导时对应的输入符号的集合，记为 $SELECT(A \rightarrow \beta)$
- $SELECT(A \rightarrow a\beta) = \{a\}$
- $SELECT(A \rightarrow \varepsilon) = FOLLOW(A)$

➢ q -文法

- 每个产生式的右部或为 ε ，或以终结符开始
- 具有相同左部的产生式有不相交的可选集

q -文法不含右部以非终结符打头的产生式

1.75

中国大学MOOC

例：表达式文法各产生式的SELECT 集

X	FIRST(X)	FOLLOW(X)
E	(id)	\$)
E'	+ ε	\$)
T	(id)	+) \$
T'	* ε	+ \$)
F	(id)	* +) \$

(1) $E \rightarrow TE' \quad \text{SELECT}(1)=\{ (\text{id}) \}$
(2) $E' \rightarrow + TE' \quad \text{SELECT}(2)=\{ + \}$
(3) $E' \rightarrow \epsilon \quad \text{SELECT}(3)=\{ \$ \}$
(4) $T \rightarrow FT' \quad \text{SELECT}(4)=\{ (\$) \}$
(5) $T' \rightarrow * FT' \quad \text{SELECT}(5)=\{ * \}$
(6) $T' \rightarrow \epsilon \quad \text{SELECT}(6)=\{ * \}$
(7) $F \rightarrow (E) \quad \text{SELECT}(7)=\{ \}$
(8) $F \rightarrow \text{id} \quad \text{SELECT}(8)=\{ \}$

8 通过 SELECT 集推出预测分析表

1.75

中国大学MOOC

非递归的预测分析法

非递归的预测分析不需要为每个非终结符编写递归下降过程，而是根据预测分析表构造一个自动机，也叫表驱动的预测分析。

输入：一个串 w 和文法 G 的分析表 M
输出：如果 w 在 $L(G)$ 中，输出 w 的最左推导；否则给出出错信息
方法：最初，语法分析器的格局如下：输入缓冲区中是 $w\$$ ， G 的开始符位于栈顶，其下面是 $\$$ 。下面的程序使用预测分析表 M 生成了处理这个输入的预测分析过程

```

输入 w 和文法 G 的分析表 M
输出: 如果 w 在 L(G) 中, 输出 w 的最左推导; 否则给出出错信息指示
方法: 最初, 语法分析器的格局如下: 输入缓冲区中是 w$, G 的开始符位于栈顶,
其下面是$. 下面的程序使用预测分析表 M 生成了处理这个输入的预测分析过程

设置 ip 指向输入 w 的第一个字符, 其中 ip 是输入指针,
令 X=栈顶字符;
while (X ≠ $) | /* 栈不空 */
    if (X 等于 ip 所指向的字符 a) 执行栈的弹出操作, 将 ip 向前移动一个位;
    else if (M[X, a] 是一个终结符) 则继续;
    else if (M[X, a] 是一个非终结符) error();
    else if (M[X, a] = X-Y1Y2...Yn)
        弹出产生式 X→Y1Y2...Yn;
        将 Y1, Y2,..., Yn 读入栈中, 其中 Y1 位于栈顶;
    }
    令 X=栈顶字符;
}

```

2.00

中国大学MOOC

表驱动的预测分析法

输入：一个串 w 和文法 G 的分析表 M
输出：如果 w 在 $L(G)$ 中，输出 w 的最左推导；否则给出出错信息指示
方法：最初，语法分析器的格局如下：输入缓冲区中是 $w\$$ ， G 的开始符位于栈顶，其下面是 $\$$ 。下面的程序使用预测分析表 M 生成了处理这个输入的预测分析过程

```

设置 ip 指向输入 w 的第一个字符, 其中 ip 是输入指针,
令 X=栈顶字符;
while (X ≠ $) | /* 栈不空 */
    if (X 等于 ip 所指向的字符 a) 执行栈的弹出操作, 将 ip 向前移动一个位;
    else if (M[X, a] 是一个终结符) 则继续;
    else if (M[X, a] 是一个非终结符) error();
    else if (M[X, a] = X-Y1Y2...Yn)
        弹出产生式 X→Y1Y2...Yn;
        将 Y1, Y2,..., Yn 读入栈中, 其中 Y1 位于栈顶;
    }
    令 X=栈顶字符;
}

```

The screenshot shows a video player interface. At the top right is the "中国大学MOOC" logo. The video content displays a comparison table titled "递归的预测分析法vs.非递归的预测分析法" (Recursive vs. Non-Recursive Parsing). The table has two columns: "递归的预测分析法" (Recursive) and "非递归的预测分析法" (Non-Recursive). The rows compare: 程序规模 (Program Size), 直观性 (Intuitiveness), 效率 (Efficiency), and 自动生成 (Automatic Generation). The non-recursive column generally includes a small smiley face icon next to its descriptions.

	递归的预测分析法	非递归的预测分析法
程序规模	程序规模较大，不需载入分析表	主控程序规模较小，需载入分析表（表较小）
直观性	较好	较差
效率	较低	分析时间大约正比于待分析程序的长度
自动生成	较难	较易

下方显示了讲师陈颖的信息：哈尔滨工业大学 陈颖 HARBIN INSTITUTE OF TECHNOLOGY CHENYIN. 播放进度条显示为 09:27 / 11:21。底部有音量、2x 放大、超高清和全屏按钮。

9 预测分析法实现步骤

The screenshot shows a video player interface. At the top right is the "中国大学MOOC" logo. The video content displays a list of five steps for parser implementation:

- 1) 构造文法
- 2) 改造文法：消除二义性、消除左递归、消除回溯
- 3) 求每个变量的FIRST集和 FOLLOW集，从而求得每个候选式的SELECT集
- 4) 检查是不是LL(1)文法。若是，构造...
- 5) 对于递归的预测分析，根据预测分析表为每一个非终结符编写一个过程；对于非递归的预测分析，实现表驱动的预测分析算法

下方显示了讲师陈颖的信息：哈尔滨工业大学 陈颖 HARBIN INSTITUTE OF TECHNOLOGY CHENYIN. 播放进度条显示为 09:27 / 11:21。底部有音量、2x 放大、超高清和全屏按钮。