

第四章 上下文无关文法

4.1

1. 为下列语言设计上下文无关文法。也请思考下列语言可不可以设计正规表达式?

(a) 满足这样条件的二进制串: 每个 0 之后都紧跟着至少一个 1;

上下文无关文法为: $S \rightarrow 01S \mid 1S$

正则表达式为: $((01)^*1^*)^*$

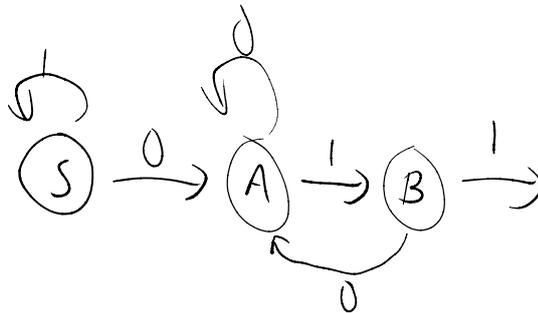
(b) 0 和 1 个数相等的二进制串;

上下文无关文法为: $S \rightarrow 0S1S \mid 1S0S \mid \epsilon$

无法设计正则表达式。

(c) 不含 011 子串的二进制串;

先设计 DFA:



上下文无关文法为:

$S \rightarrow 0A \mid 1S$

$A \rightarrow 0A \mid 1B$

$B \rightarrow 0A$

正则表达式为: $(0(?!11)|1(?!011))^*$

(d) 具有形式 xy 的二进制串, $x \neq y$;

无法设计上下文无关文法和正则表达式。

(e) 形如 xx 的二进制串;

上下文无关文法为: $S \rightarrow 0S0S \mid 1S1S \mid \varepsilon$

无法设计正则表达式。

2. 考虑文法

$$S \rightarrow (L) \mid a$$

$$L \rightarrow L, S \mid S$$

(a) 列出终结符、非终结符和开始符号;

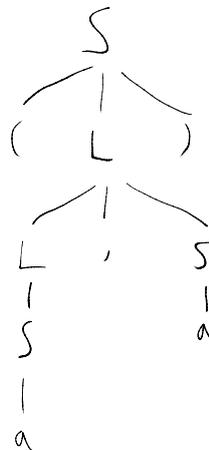
终结符为 (,)

非终结符为 L S

开始符号为 S

(b) 给出下列句子的语法树

(i) (a, a)



(ii) $(a, (a, a))$

$$\begin{aligned}
S &\Rightarrow (L) \Rightarrow (L, S) \Rightarrow (S, S) \\
&\Rightarrow (a, S) \Rightarrow (a, (L)) \Rightarrow (a, (L, S)) \\
&\Rightarrow (a, (S, S)) \Rightarrow (a, ((L, S), S)) \\
&\Rightarrow (a, ((S, S), S)) \Rightarrow (a, ((a, S), S)) \\
&\Rightarrow (a, ((a, a), S)) \Rightarrow (a, ((a, a), (L))) \\
&\Rightarrow (a, ((a, a), (L, S))) \Rightarrow (a, ((a, a), (S, S))) \\
\text{(iii)} \quad &\Rightarrow (a, ((a, a), (a, S))) \Rightarrow (a, ((a, a), (a, a)))
\end{aligned}$$

(d) 构造 (b) 中句子的最右推导:

$$\begin{aligned}
\text{(i)} \quad S &\Rightarrow (L) \Rightarrow (L, S) \Rightarrow (L, a) \Rightarrow (S, a) \Rightarrow (a, a) \\
S &\Rightarrow (L) \Rightarrow (L, S) \Rightarrow (L, (L, S)) \Rightarrow (L, (L, a)) \\
\text{(ii)} \quad &\Rightarrow (L, (S, a)) \Rightarrow (L, (a, a)) \Rightarrow (S, (a, a)) \Rightarrow (a, (a, a)) \\
S &\Rightarrow (L) \Rightarrow (L, S) \Rightarrow (L, (L, S)) \Rightarrow (L, (L, (L))) \\
&\Rightarrow (L, (L, (L, S))) \Rightarrow (L, (L, (L, a))) \\
&\Rightarrow (L, (L, (S, a))) \Rightarrow (L, (L, (a, a))) \\
&\Rightarrow (L, (S, (a, a))) \Rightarrow (L, ((L), (a, a))) \\
&\Rightarrow (L, ((L, S), (a, a))) \Rightarrow (L, ((L, a), (a, a))) \\
&\Rightarrow (L, ((S, a), (a, a))) \Rightarrow (L, ((a, a), (a, a))) \\
\text{(iii)} \quad &\Rightarrow (S, ((a, a), (a, a))) \Rightarrow (a, ((a, a), (a, a)))
\end{aligned}$$

(e) 该文法产生的语言是什么? (可以用自然语言描述出来, 也可以用集合的形式表示出来)

以 a 或 a 组成的元组递归组成的元组。

3. 考虑文法

$$S \rightarrow aSbS \mid bSaS \mid \varepsilon$$

为某个句子构造两个不同的最左推导, 以证明它是二义性的。

$$aabbab$$

$$S \Rightarrow aSbS \Rightarrow aaSbSbS \Rightarrow aabSbS \Rightarrow abbS \Rightarrow abbaSbS \Rightarrow abbabS \Rightarrow aabbab$$

$$S \Rightarrow aSbS \Rightarrow aaSbSbS \Rightarrow aabSbS \Rightarrow aabbsaSbS \Rightarrow abbaSbS \Rightarrow abbabS \Rightarrow aabbab$$

4.2

1. 考虑文法

$$S \rightarrow (L) \mid a$$

$$L \rightarrow L, S \mid S$$

消除左递归。

$$S \rightarrow (L) \mid a$$

$$L \rightarrow SL'$$

$$L' \rightarrow ,SL' \mid \varepsilon$$

2. 下面文法 G[S]

$$S \rightarrow AbB \mid A$$

$$A \rightarrow AB \mid caB \mid B$$

$$B \rightarrow Aa \mid b$$

消除左递归。

$$S \rightarrow AbB \mid A$$

$$A \rightarrow caBA' \mid BA'$$

$$A' \rightarrow BA' \mid \varepsilon$$

$$B \rightarrow ABaB' \mid caBaB' \mid bB'$$

$$B' \rightarrow aB' \mid \varepsilon$$

3. 下面文法 G[S]:

$$S \rightarrow aFbM \mid F$$

$$F \rightarrow M \mid abc$$

$$M \rightarrow abF \mid c$$

消除形如: $A \rightarrow B$ 这种单产生式, 并进行文法处理 (左递归和公共左因子)。

先消除单产生式：

$$S \rightarrow aFbM \mid abF \mid c \mid abc$$

$$F \rightarrow abF \mid abc$$

$$M \rightarrow abF \mid c$$

此时没有左递归，只需要消除公共左因子：

$$S \rightarrow aS' \mid c$$

$$S' \rightarrow FbM \mid bS''$$

$$S'' \rightarrow F \mid c$$

$$F \rightarrow abF'$$

$$F' \rightarrow F \mid c$$

$$M \rightarrow abF \mid c$$

即为结果。

4.3

1. 对于如下的文法，求出各个 *FIRST* 集和 *FOLLOW* 集。

$$S \rightarrow aAB \mid bA \mid \varepsilon$$

$$A \rightarrow aAb \mid \varepsilon$$

$$B \rightarrow bB \mid \varepsilon$$

$$FIRST(S) = \{a, b, \varepsilon\} \quad FOLLOW(S) = \{\$\}$$

$$FIRST(A) = \{a, \varepsilon\} \quad FOLLOW(A) = \{b, \$\}$$

$$FIRST(B) = \{b, \varepsilon\} \quad FOLLOW(B) = \{\$\}$$

2. 验证下列文法是否为 *LL(1)* 文法。

(1) 文法如下：

$$S \rightarrow AB \mid CDa$$

$$A \rightarrow ab \mid c$$

$$B \rightarrow dE$$

$$C \rightarrow eC \mid \varepsilon$$

$$D \rightarrow fD \mid f$$

$$E \rightarrow dE \mid \varepsilon$$

关于 D 的产生式有公共左因子，先提取公共左因子。

$$D \rightarrow fD'$$

$$D' \rightarrow D \mid \varepsilon$$

没有其他的公共左因子，也没有左递归，所以可以计算 *FIRST* 集和 *FOLLOW* 集了。

$$\begin{aligned} FIRST(S) &= \{a, c, e, \varepsilon\} & FOLLOW(S) &= \{\$\} \\ FIRST(A) &= \{a, c\} & FOLLOW(A) &= \{d\} \\ FIRST(B) &= \{d\} & FOLLOW(B) &= \{\$\} \\ FIRST(C) &= \{e, \varepsilon\} & FOLLOW(C) &= \{f\} \\ FIRST(D) &= \{f\} & FOLLOW(D) &= \{a\} \\ FIRST(D') &= \{f, \varepsilon\} & FOLLOW(D') &= \{a\} \\ FIRST(E) &= \{d, \varepsilon\} & FOLLOW(E) &= \{\$\} \end{aligned}$$

可以构造语法解析表如下：

	a	b	c	d	e	f	\$
S	$S \rightarrow AB$		$S \rightarrow AB$		$S \rightarrow CDa$		acc
A	$A \rightarrow ab$		$A \rightarrow c$				
B				$B \rightarrow dE$			
C						$C \rightarrow eC$	$C \rightarrow \varepsilon$
D						$D \rightarrow fD'$	
D'	$D' \rightarrow \varepsilon$					$D' \rightarrow D$	
E					$E \rightarrow dE$		$e \rightarrow \varepsilon$

语法解析表每个单元格不超过一个产生式，所以此文法是 LL(1) 文法。

(2) 文法如下：

$$\begin{aligned} S &\rightarrow aABbCD \mid \varepsilon \\ A &\rightarrow ASd \mid \varepsilon \\ B &\rightarrow Sac \mid eC \mid \varepsilon \\ C &\rightarrow Sf \mid Cg \mid \varepsilon \\ D &\rightarrow aBD \mid \varepsilon \end{aligned}$$

关于 A 的产生式有左递归，改写如下：

$$A \rightarrow A'$$

$$A' \rightarrow SdA \mid \varepsilon$$

又可以合并如下：

$$A \rightarrow SdA \mid \varepsilon$$

关于 C 的产生式也有左递归，改写如下：

$$C \rightarrow SfC' \mid C'$$

$$C' \rightarrow gC' \mid \varepsilon$$

没有公共左因子，也没有其他的左递归，所以可以计算 $FIRST$ 集和 $FOLLOW$ 集如下：

$$FIRST(S) = \{a, \varepsilon\} \quad FOLLOW(S) = \{d, a, f, \$\}$$

$$FIRST(A) = \{a, \varepsilon\} \quad FOLLOW(A) = \{a, e\}$$

$$FIRST(B) = \{a, \varepsilon, e\} \quad FOLLOW(B) = \{b, a, \$\}$$

$$FIRST(C) = \{a, \varepsilon, g\} \quad FOLLOW(C) = \{a, b, \$\}$$

$$FIRST(C') = \{g, \varepsilon\} \quad FOLLOW(C') = \{a, b, \$\}$$

$$FIRST(D) = \{a, \varepsilon\} \quad FOLLOW(D) = \{\$\}$$

可以构造语法解析表如下：

	a	b	c	d	e	f	g	\$
S	$S \rightarrow aABbCD$ $S \rightarrow \varepsilon$			$S \rightarrow \varepsilon$	$S \rightarrow \varepsilon$			acc
A	$A \rightarrow SdA$ $A \rightarrow \varepsilon$				$A \rightarrow \varepsilon$			
B	$B \rightarrow Sac$ $B \rightarrow \varepsilon$	$B \rightarrow \varepsilon$				$B \rightarrow eC$		$B \rightarrow \varepsilon$
C	$C \rightarrow SfC'$ $C \rightarrow C'$	$C \rightarrow C'$						$C \rightarrow C'$
C'	$C' \rightarrow \varepsilon$	$C' \rightarrow \varepsilon$					$C' \rightarrow gC'$	$C' \rightarrow \varepsilon$
D	$D \rightarrow aBD$							$D \rightarrow \varepsilon$

语法解析表有单元格超过一个产生式，所以不是 LL(1) 文法。