

一.填空题

1.一个典型的编译程序，它一般包括八个方面的内容：

- ① 词法分析程序 ⑤ 代码优化程序
- ② 语法分析程序 ⑥ 目标代码生成
- ③ 语义分析程序 ⑦ 错误检查处理
- ④ 中间代码生成 ⑧ 信息表管理

2.编译执行和解释执行的区别在于：是否产生目标代码。

3.一个文法通常可表示成一个四元式 $G[S] = (V_N, V_T, P, S)$ 。

4.一个递归文法所产生的句子，其个数必然是无穷个。

5.设 $G[S]$ 为一文法，由文法的开始符号 S 推导出的符号串称为 G 的句型。

6.一个句型的最左直接短语（即规范分析中，最先被规约的子串）称为该句型的句柄。

7.Chomsky 定义了四类基本的文法，分别称之为：

- ① 0 型方法(短语结构) ③ 2 型方法(前后文无关)
- ② 1 型方法(前后文有关) ④ 3 型方法(正规)

8.词法分析的任务就在于依次扫描输入串中的各个字符并从其中识别出一系具有独立意义的单词，我们通常把构成各个单词的字符串称为该单词的词文。

9.一般来说各类单词的语法都能用相应的正规(3 型)文法来描述。

10 我们通常把一个有限自动机表示为 $M = (K, \Sigma, \delta, S_0, Z)$ 。

11 引入具有 ϵ 动作的 NFA 主要目的是把识别各类单词的有限自动机用 ϵ 失线连接

起来,组成一个单一的NFA,然后把所得的NFA确定化后再据此设计编译程序的词法分析器。

12 正规文法、正规式,在描述语言的意义下是等价的。

13 状态转换图、状态矩阵、有限自动机,在识别语言的意义下是等价的。

14 状态转换图中初态结点没有射入矢线,终态结点没有射出矢线。

15 通常构造词法分析程序的两种途径是:

- ①手工方式编程 ②借助工具自动生成。

二.判断题

1.文法和语言之间是一一对应关系。(×)

2.设 G_1 和 G_2 为两个文法,若它们所产生的语言相等,即 $L(G_1)=L(G_2)$,则称 G_1 和 G_2 等价。(√)

3.每棵语法树的叶从左到右排列组成一个句型。(√)

4.若文法 $G=(V_N, V_T, P, S)$ 的每一个产生式形如 $A \rightarrow \beta$, $A \in V_N$, $\beta \in V^*$,则称为上下文无关文法,通常用来描述计算机语言的语法结构。(×)

5.前后文无关文法的是否具有二义性是不可判定的。(√)

6.同一字母表上的 NFA 和 DFA 是等价的。(√)

7.若自动机 M_1 的状态数和自动机 M_2 的状态数不相等,则 M_1 与 M_2 一定不等价 (×)

8.状态转换图中终态结点只有一个。(×)

9.正规集与正规式之间并不存在一一对应关系。(√)

10 某些语言不能用正规式来描述,因此正规式的描述能力是有限的。(√)

三. 求解题

1. 试描述文法所产生的语言。 $S \rightarrow aSa | bSb | c$

答: $S \rightarrow aSa \rightarrow abSba \rightarrow abcbaba$ $L(G) = \{a^m b^n c b^n a^m | m, n \geq 0\}$

2. 试描述文法所产生的语言。 $S \rightarrow AB$ $A \rightarrow aAb | ab$ $B \rightarrow cBd | cd$

答: $S \rightarrow AB \rightarrow aAbcBd \rightarrow aabbccdd$ $L(G) = \{a^n b^n c^m d^m | m, n \geq 0\}$

3. 试构造所产生语言的文法。 $\{a^m b^n c^n d^m | m, n \geq 1\}$

答: $S \rightarrow AB$ $A \rightarrow aA | d$ $B \rightarrow bB | c$

4. 试构造所产生语言的文法。 $\{a^n \# b^n | n \geq 0\} \cup \{c^n \# d^n | n \geq 0\}$

答: $S \rightarrow E | F | \#$ $E \rightarrow aEb | \#$ $F \rightarrow cFd | \#$

5. 改写文法消除文法左递归。引入 $\{ \}$ 改写: $E \rightarrow EAT | T$, $T \rightarrow TMF | F$

答:

$E \rightarrow T \{AT\}$

$T \rightarrow F \{MF\}$

6. 改写文法消除文法左递归。引入 E', T' 改写: $E \rightarrow EAT | T$, $T \rightarrow TMF | F$

答:

$E \rightarrow TE'$ $E' \rightarrow ATE' | \epsilon$

$T \rightarrow FT'$ $T' \rightarrow MFT' | \epsilon$

四. 综合题

1. 文法: $S \rightarrow AB | c$ $A \rightarrow bA | a$ $B \rightarrow aSb | c$, 给出句子 $w = bbaacb$ 的最左推导和最右推导。这种推导是属于自顶向下还是自底向上的句型分析?

答:

① w 的最左推导

$S \rightarrow AB \rightarrow bAB \rightarrow bbAB \rightarrow bbaB \rightarrow bbaacb$

②W 的最右推导

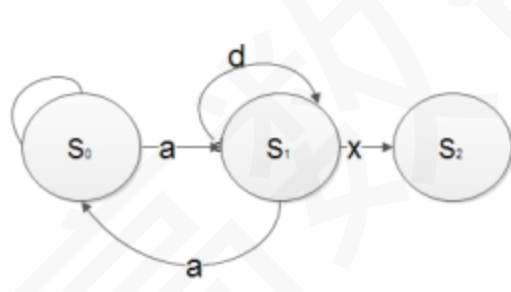
$S \rightarrow AB \rightarrow AaSb \rightarrow Aacb \rightarrow bAacb \rightarrow bbAacb \rightarrow bbaacb$

自底向上

2. 假设 $\langle \text{标识符} \rangle \rightarrow \langle \text{标识符} \rangle \text{字母} | \langle \text{标识符} \rangle \text{数字} | \text{字母} | \text{下划线}$ ，画出识别标识符的状态转换图。

答：

$S \rightarrow Sa | Sd | a | u$



3. 对下面给出的状态矩阵（初态 S，终态 C），画出相应的状态转换图，写出相应的右线性文法，并指出它接受的最短输入串。

	a	b
S	A	S
A	C	B
B	B	C
C	C	C

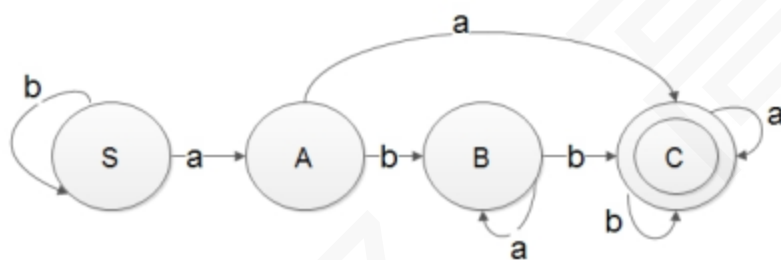
答：

$S \rightarrow aA | bS$

$A \rightarrow aC | bB$

$B \rightarrow aB | bC$

$C \rightarrow a | b | ac$



4. 试证明：方程 $X = rX + t$ ($X \rightarrow rX | t$) 有形如 $X = r^*t$ 的解。

方程 $X = Xr + t$ ($X \rightarrow Xr | t$) 有形如 $X = tr^*$ 的解。

答：

$X = rX + t \Rightarrow x \rightarrow rX | t (x \rightarrow rX, x \rightarrow t)$

$\therefore Lx = \{t, rt, rrt, \dots\}$

$X = r^*t$

$\therefore S = (a+ba)^*b \quad s = (a|ba)^*b$

亦可 $S \cdot a^*(baS + b) = a^*baS + a^*b = (a^*ba)^*a^*b$

$(a|ba)^*b$ 与 $(a^*ba)^*a^*b$ 等价

5. 由下面的正规文法构造相应的正规式。

$S \rightarrow aA \quad A \rightarrow bA | aB | b \quad B \rightarrow aA$

解：1. $S \rightarrow aA$ 2. $A \rightarrow bA | aB | b$ 3. $B \rightarrow aA$

将 3 代入 2 可得 $A = bA + aaA + b = (b+aa)A + b$

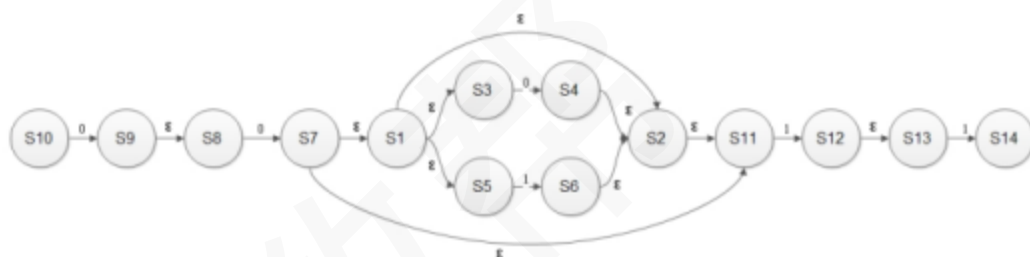
方程 $x = r^* + t$ 形有如同 $x = r^*t$ 的解

那么 $A = (b+aa)^*b$ 将 $A = (b+aa)^*b$ 代入得

$S = a(b+aa)^*b$

即文法所产生正规式 $a(b|aa)^*b$

6. 对给定的正规式 $r = 00(0|1)^*11$ 用 Thompson 法构造其 F.A.M.



五. 应用题

1. 假设小数具有 $dd\dots d.$, $.dd\dots d$ 和 $dd\dots d.dd\dots d$ 三种形式, 请用 20 行左右伪代码描述识别小数的过程, 并返回该小数的数值。

一. 填空题

1. 语法分析程序是以单词串形式的源程序作为输入和分析的对象。

2. 就产生语法树的方向而言, 大致将语法分析分为自顶向下(预测分析法)和自底向上(算符优先分析法)两大类。

3. 自顶向下的语法分析方法要回避的主要问题是左递归和回溯。

4. 在自顶向下的语法分析中, $FIRST(\gamma)$ 称为候选式的终结首符集, $FOLLOW(A)$ 称为符号 A 的后继终结符号集。

5. LL(1)分析法中, 第一个 L 的含义是从左到右扫描字符串, 第二个 L 的含义是最左推导, 1 的含义是向前查看一个符号。

6. 简单优先分析法每次所规约的是句型的句柄。算符优先分析法每次所规约的是句型的最左素短语。