

编译原理

习题集

主 编 孙忠林
副主编 张 鹏
王 鸽
吴振寰

BIANYI
YUANLIXITUI



高等学校教材 GAO DENG XUE XIAO JIAO CAI

编译原理 习题集

主 编 孙忠林

副主编 张 鹏

王 鸽

吴振寰

主 审 郑永果

中国石油大学出版社

图书在版编目 (CIP) 数据

编译原理习题集/孙忠林主编. —东营: 中国石油大学出版社, 2008.3
ISBN 978-7-5636-2545-1

I. 编… II. 孙… III. 编译程序—程序设计—习题
IV. TP314-44

中国版本图书馆 CIP 数据核字 (2008) 第 019335 号

书 名: 编译原理习题集
主 编: 孙忠林

责任编辑: 刘 静 (电话 0546-8395937)
封面设计: 九天设计

出 版 者: 中国石油大学出版社 (山东 东营, 邮编 257061)
网 址: <http://www.uppbook.com.cn>
电子邮箱: cbs2006@163.com
印 刷 者: 沂南县汇丰印刷有限公司
发 行 者: 中国石油大学出版社 (电话 0546-8392062)
开 本: 185×260 印张: 8.875 字数: 215 千字
版 次: 2008 年 3 月第 1 版第 1 次印刷
定 价: 14.50 元

版权所有, 翻印必究。举报电话: 0546-8391810

本书封面覆有中国石油大学出版社标志的激光防伪膜。

本书封面贴有中国石油大学出版社标志的电码防伪标签, 无标签者不得销售。

序

编译原理是计算机科学与技术专业重要的专业基础课程，由山东省教育厅组编的山东省高校统编教材《编译原理》经多所高校使用，于2007年进行了修订再版。为配合《编译原理（修订版）》教材，解答学生在学习编译原理课程中所遇到的问题，使学生掌握课程内容和学习重点，经过山东科技大学信息科学与工程学院几位教师的努力，编写了这本《编译原理习题集》。

《编译原理习题集》的参编教师从事编译原理教学多年，有丰富的教学经验，了解学生在学习编译原理过程中的困难所在。在内容组织上，习题集采用了便于学习的方式，每章由本章要点、典型例题分析和习题与解答三部分组成，通过典型例题解析来辅助知识的学习和能力的培养。在例题设计时，从其科学性、应用性及开放性出发，营造贴近实际应用的环境，激发学生的学习兴趣，使学生在面对一个实际问题时，不仅要知其然，还要知其所以然，不仅使学生理解编译原理的基本理论，还要使学生掌握实际编译系统开发的基本能力。

本书是学习和掌握编译原理的一本很好的辅助材料，相信会对学习编译原理的广大师生有很大的帮助。

郑永果

2007年12月于青岛

前 言

《编译原理习题集》是《编译原理（修订版）》的配套用书。编译原理课程内容对本科生来说难度较大，许多同学在学习过程中希望有一本配套习题集作为学习的辅助资料。山东科技大学信息科学与工程学院的几位任课教师依据《编译原理（修订版）》的内容和多年的教学经验，组织编写了此习题集，以满足广大师生的需要。

本书除第1章外，其余各章均由3部分组成：本章要点、典型例题分析和习题与解答。“本章要点”给出了配套教材对应章节的内容重点、要点和难点，便于同学们归纳所学内容。“典型例题分析”选择了部分典型的示例并进行了必要的分析，给出了解题的思路和答案。“习题与解答”依据配套教材的章后习题，给出了习题的解题过程和答案，部分章节还增加了相应的习题。另外，习题集中还对部分院校的考研试题进行了分析解答。

本书共分为9章，前8章与配套教材的章节对应，第9章是教材附录课程设计的内容。第1、3、7章由吴振寰编写，第2、6、8章由王鸽编写，第4、5章由张鹏编写，第9章由孙忠林编写。山东科技大学信息科学与工程学院的郑永果教授对习题集作了全面审校，提出了许多宝贵意见，在此表示衷心的感谢。

由于编者水平有限，加之成书时间比较仓促，书中难免有不妥和错误之处，敬请读者批评指正。

编 者

2007年12月于青岛

目 录

第1章 绪 论.....	1
1.1 本章要点.....	1
1.2 习题与解答.....	1
第2章 文法和语言.....	3
2.1 本章要点.....	3
2.2 典型例题分析.....	4
2.3 习题与解答.....	6
第3章 词法分析.....	16
3.1 本章要点.....	16
3.2 典型例题分析.....	19
3.3 习题与解答.....	26
第4章 语法分析.....	38
4.1 本章要点.....	38
4.2 典型例题分析.....	44
4.3 习题与解答.....	50
第5章 语法制导翻译与中间代码生成.....	63
5.1 本章要点.....	63
5.2 典型例题分析.....	64
5.3 习题与解答.....	66
第6章 代码优化.....	83
6.1 本章要点.....	83
6.2 典型例题分析.....	85
6.3 习题与解答.....	90
第7章 运行时存储空间管理.....	106
7.1 本章要点.....	106
7.2 典型例题分析.....	109
7.3 习题与解答.....	114
第8章 目标代码生成.....	118
8.1 本章要点.....	118
8.2 典型例题分析.....	120
8.3 习题与解答.....	123
第9章 课程设计示例及指导.....	127

9.1 课程设计概述.....	127
9.1.1 设计目的.....	127
9.1.2 设计内容.....	127
9.1.3 组织实施.....	128
9.1.4 时间安排.....	128
9.2 课程设计示例.....	128
9.2.1 题目及要求.....	128
9.2.2 设计过程.....	129
9.3 课程设计参考题目.....	131

第1章 绪 论

1.1 本章要点

通过编写程序实现各种功能是计算机应用的主要特点。

计算机程序设计语言经历了由机器语言、汇编语言到高级语言的发展过程，编译器（编译程序）已经成为计算机系统中的一个重要系统软件。

高级语言编写的源程序是按一定规则排列的符号集合，翻译加工成机器指令程序才能被计算机所执行。

源程序翻译的方法有两种：编译程序把源程序翻译成可以执行的目标程序；解释程序则不生成目标程序，而是边翻译边执行。

本章的要点是掌握源程序翻译过程的各个阶段及其任务和编译器各部分的功能。一般来说，编译器被划分为6个阶段、8个功能部分。6个阶段是词法分析、语法分析、语义分析、中间代码生成、代码优化和目标代码生成。此外，表格管理和错误检测是贯穿整个编译过程的两个重要功能。

1.2 习题与解答

1-1 什么叫源程序？什么叫目标程序？

答：

源程序是待编译的高级语言程序，而目标程序就是经编译后可以在计算机上运行的机器语言程序。

1-2 在高级语言的参数传递方式中，传地址方式和传值方式有何区别？

答：

传地址方式和传值方式的主要区别在于：过程（或函数）的形式参数对应的形式单元内存放的是实参的地址还是实参值的拷贝。如果形式单元内存放实参的地址，则是传地址方式，任何对形参的访问都间接地访问到实参，所以实参的值可以被过程（或函数）所改变。如果形式单元内存放实参值的拷贝，则是传值方式，对形参的访问不会影响实参的值。

1-3 分别指出翻译程序、汇编程序、编译程序和解释程序的含义。

答：

完成把高级语言编写的源程序转化为可以运行的机器语言程序这一过程的是翻译程序。

用汇编语言编写的程序是汇编程序，汇编语言是用易读的符号来描述机器指令的。

把高级语言源程序翻译成可以运行的目标程序的是编译程序。

对高级语言源程序翻译一句执行一句的是解释程序。

1-4 编译程序通常由哪几个主要部分组成？各部分的主要功能是什么？

答：

编译程序通常有八大部分：表格管理、错误检测、词法分析、语法分析、语义分析、中间代码生成、代码优化和目标代码生成。

表格管理的功能是：收集编译过程中的各种符号及其属性信息，以备编译程序的其他部分快速地存储和检索符号。

错误检测的功能是：将各个阶段出现的错误进行适当校正，使得编译器能够继续运行，以便尽可能多的发现源程序中的错误。

词法分析的功能是：一个一个地读入源程序的字符，识别出单词，组成单词序列。

语法分析的功能是：把单词序列分解成各类语法单位，表示成语法树，并按照语法规则确认程序的语法是否正确。

语义分析的功能是：审查各语法单位的语义，发现错误，为代码生成阶段收集语义信息，如类型检查等。

中间代码生成的功能是：用一种结构简单、含义明确的记号系统产生源程序的一个中间表示，易于优化，易于翻译成目标程序。

中间代码优化的功能是：对中间代码进行改造，以使目标程序具有更高的存储和时间效率。

目标代码生成的功能是：生成可重定位的机器指令代码或汇编指令代码。

1-5 什么叫单词？什么是单词的属性？

答：

单词是源程序中逻辑上紧密相连的一组字符。单词的属性是指单词的一些信息，如单词的值、类别、所在的层次等。

1-6 常用的查填表技术有哪些？它们各自的特点是什么？

答：

常用的查填表技术有 3 种：线性查找、二叉树和杂凑技术。

线性查找的结构简单，占用存储空间少，但效率较低。

二叉树的查找速度比线性查找要快得多，但要边填表边顺序化，因此填表速度慢。

杂凑技术的查填表速度都很快，但需要预先准备相当大的空间。

1-7 错误处理的主要任务是什么？

答：

错误处理的任务是：把检测到的错误局部化，尽可能多地编译源程序代码，以便发现更多的错误，有可能的话进行适当的错误校正。

第2章 文法和语言

2.1 本章要点

1. 基本概念和运算

字母表是符号元素的非空有限集合, 用 Σ 表示。字母表中的元素称为符号, 常用小写字母 a、b、c 等表示。

符号串和符号串集合的运算有: 符号串的连接(乘积)、符号串的长度、符号串的前缀与后缀、符号串集合的乘积、符号串的方幂、串集合的方幂、正闭包及闭包。

2. 文法和 Chomsky 分类

文法是一个四元组: $G[S] = (V_N, V_T, P, S)$ 。其中:

V_N 为非终结符集合或称变量集; V_T 为终结符集合; P 为产生式集合; S 为开始符号(或称根符号, 识别符号)。其中: $V_N \cap V_T = \Phi$, 一般令 $V = V_N \cup V_T$, V 中的符号称为文法符号。

文法的 Chomsky 分类: 设文法 $G = (V_N, V_T, P, S)$ 。

0 型文法(短语文法), 其产生式形式为: $\alpha \rightarrow \beta$ 。其中: $\alpha \in V^* V_N V^*$, $\beta \in V^*$ 。

1 型文法(上下文有关文法), 其产生式形式为: $\alpha \rightarrow \beta$ 。其中: α, β 除满足 0 型文法的定义以外, 还要求 $|\alpha| \leq |\beta|$, 根符号产生式 $S \rightarrow \varepsilon$ 除外。

2 型文法(上下文无关文法), 其产生式形式为: $A \rightarrow \beta$ 。其中: $A \in V_N$, $\beta \in V^+$, 根符号产生式 $S \rightarrow \varepsilon$ 除外。

3 型文法(正则文法), 其产生式形式为: $A \rightarrow \alpha B \mid \beta$ 或 $A \rightarrow B\alpha \mid \beta$ 。其中: $A, B \in V_N$, $\alpha, \beta \in V_T^+$ 。

3. 推导和归约

已给文法 $G = (V_N, V_T, P, S)$, $V = V_N \cup V_T$, $x, y, \alpha, \beta \in V^*$, 且 $\alpha \rightarrow \beta \in P$, 此时, 符号串 $x\alpha y$ 能够直接推导出符号串 $x\beta y$, 记作 $x\alpha y \xRightarrow{G} x\beta y$, 或记作 $x\alpha y \Rightarrow x\beta y$ 。

若符号串 $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in V^*$ 且 $\alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{n-1} \Rightarrow \alpha_n$, 则 α_1 可推导出 α_n 来, 记作 $\alpha_1 \Rightarrow^* \alpha_n$ 。

若推导关系 $\alpha_1 \Rightarrow^* \alpha_n$ 中允许 $\alpha_1 = \alpha_n$, 则 α_1 和 α_n 的推导关系记作 $\alpha_1 \xRightarrow{*} \alpha_n$ 。归约是推导的逆过程, 若存在 $\alpha_1 \xRightarrow{*} \alpha_n$, 则认为 α_n 能够归约成 α_1 。

若在推导关系 $\alpha_1 \Rightarrow^* \alpha_n$ 中每次最先替换最右(左)的非终结符, 则称为最右(左)推导; 若在归约过程中每次最先归约最左(右)的非终结符, 则称为最左(右)归约。

句型的最右推导称为规范推导, 其逆过程最左归约称为规范归约。

4. 语法树

掌握语法树的画法, 从语法树上能够找出文法的短语、直接短语和句柄。能够利用语

法树判断文法是否为二义性文法。

文法句型的短语：设文法 $G=(V_N, V_T, P, S)$ ，若有 $S \xRightarrow{*} xAy \xRightarrow{*} x\beta y$ ，则 β 称为句型 $x\beta y$ 的相对于 A 的短语。若有 $S \xRightarrow{*} xAy \Rightarrow x\beta y$ ，则 β 称为句型 $x\beta y$ 的相对于 A 的直接短语。

句型的最左直接短语称为此句型的句柄。

语法树的叶结点符号自左至右所组成的符号串就是所给句型；剪下任何子树，子树叶结点符号自左至右所组成的符号串是相对于子树根结点的短语；二层子树的叶结点符号所组成的符号串是直接短语；最左二层子树的叶结点符号所组成的符号串是句柄。

若一个文法存在某个句型对应两棵不同的语法树，则称此文法是二义性文法。

若文法是非二义的，则此文法的规范句型的句柄是唯一的。

5. 文法等价变换

掌握文法等价变换的三种方法：消除无用符号和无用产生式、消除单一产生式、消除或规范空符产生式。

能够利用相应的算法对文法进行等价变换。

本章的难点是语言和文法之间的转换，能够根据文法写出该文法所表示的语言，也能够根据语言写出与之等价的文法。

2.2 典型例题分析

例 2-1 写出不包含无效前导零的无符号偶数所对应的文法 G 。

分析：

如下文法 $G[S]$ ：

```

S → <非 0 符><偶数>
<偶数> → <整><偶符>
<整> → <整><数> | ε
<偶符> → 0|2|4|6|8
<非 0 符> → 1|2|3|4|5|6|7|8|9
<数> → 1|3|5|7|9|<偶符>

```

在这个文法中引出了一个丢失语言中的句子的经典错误：该文法可以导出任意 2 位及更多位数的不包含无效前导零的无符号偶数，但排斥了一位偶数。正确的处理方法是，把 $G[S]$ 的第一行改为 $S \rightarrow \text{<非 0 符><偶数>|<偶符>}$ 。

答：

文法 $G[S]$ 的正确写法为：

```

S → <非 0 符><偶数>|<偶符>
<偶数> → <整><偶符>
<整> → <整><数> | ε
<偶符> → 0|2|4|6|8
<非 0 符> → 1|2|3|4|5|6|7|8|9
<数> → 0|<非 0 符>

```

例 2-2 文法 $G[S]: S \rightarrow S(S)S | \varepsilon$ 对应的语言 $L(G[S])$ 是什么?

分析:

首先这个文法的终结符集为 $\{ (,) \}$, 这就决定了句子只能由左右圆括号组成。S 的每一次非 ε 推导会产生一对左右圆括号, 由于这是 $S(S)S$ 的形式, 因此在已产生的 $()$ 的前面、后面和里面都可以产生圆括号对, 而且能递归分析下去, 同时由于 $S \rightarrow \varepsilon$ 的推导, 又可以使某一位置的 S 化去。这样, 句子中的左右圆括号的配对和嵌套具有随意性。由此可以得到: $L(G[S]) = \{ \alpha | \alpha \text{ 为可以任意嵌套的配对左右圆括号序列} \}$ 。做一些句子实例可以验证这一点。但是, 上述结论忽略了 $S \rightarrow \varepsilon$ 对空串的认可。

答:

正确结论应该是: $L(G[S]) = \{ \alpha | \alpha \text{ 为空串或可以任意嵌套的配对左右圆括号序列} \}$ 。

例 2-3 试写一文法, 使其描述的语言 $L(G)$ 是能被 5 整除的整数集合。

分析:

能被 5 整除的整数, 其末位数必定是 0 或 5, 所以, 只要保证生成的整数末位数字是 0 或 5 即可。

先凑出产生非负整数的产生式。

(1) 首先产生满足要求的一位整数: $Z \rightarrow 0|5$

(2) 再产生满足要求的二位整数: $Z \rightarrow A(0|5)$, $A \rightarrow 0|1|2|3|4|5|6|7|8|9$

(3) 最后产生满足要求的多位整数, 只需要添加产生式 $A \rightarrow AA$ 即可:

$Z \rightarrow A0|A5|0|5$

$A \rightarrow 0|1|2|3|4|5|6|7|8|9|AA$

加上符号, 即可得到构造能被 5 整除的整数集合的文法如下:

$G[Z]: Z \rightarrow [+|-](A0|A5|0|5)$

$A \rightarrow 0|1|2|3|4|5|6|7|8|9|AA$

如果还要求整数除 0 外, 均不以 0 开头, 则文法为:

$G[Z]: Z \rightarrow [+|-](A0|A5|0|5)$

$A \rightarrow AB|C$

$B \rightarrow 0|C$

$C \rightarrow 1|2|3|4|5|6|7|8|9$

答:

$G[Z]: Z \rightarrow [+|-](A0|A5|0|5)$

$A \rightarrow AB|C$

$B \rightarrow 0|C$

$C \rightarrow 1|2|3|4|5|6|7|8|9$

例 2-4 设文法 G 的产生式集为:

$E \rightarrow E+T | E-T | T$

$T \rightarrow T * F | T / F | F$

$F \rightarrow (E) | i$

(1) 给出 $i*(i-i)$ 的最右推导;

(2) 给 $i+i/i$ 的语法树, 找出此句子的全部短语和句柄;

(3) 给出它们的语言集合。

分析:

本题主要涉及最右推导与最左推导的定义以及语法树的画法。本题解题的关键在于画出语法树, 利用句型的语法树可以非常直观地确定句型的短语和句柄。

答:

(1) 由于 $i*(i-i)$ 中含有 “*”、“-” 和括号, 所以在推导过程中应该选择含有这 3 个符号的产生式。推导过程如下:

$E \Rightarrow T \Rightarrow T*F \Rightarrow T*(E) \Rightarrow T*(E-T) \Rightarrow T*(E-F) \Rightarrow T*(E-i) \Rightarrow T*(T-i) \Rightarrow T*(F-i) \Rightarrow T*(i-i) \Rightarrow F*(i-i) \Rightarrow i*(i-i)$

(2) 句型 $i+i/i$ 的语法树如图 2-1 所示。

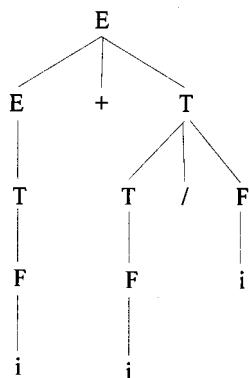


图 2-1 句型 $i+i/i$ 的语法树

由此语法树可知, 句型 $i+i/i$ 的短语有: i , i/i , $i+i/i$; 句柄是 i 。

(3) $L(G)=\{\text{可以带括号的加、减、乘、除四则运算表达式}\}$ 。

2.3 习题与解答

2-1 设字母表 $A=\{a,b,c\}$, $B=\{0,1,2\}$, 试回答如下问题:

(1) $AB=?$; $A^0=?$; $A^2=?$

(2) $AB=BA?$

(3) AB 中的符号长度等于多少?

答:

本题难度不大, 主要涉及符号串集合的乘积、符号串集合的方幂等运算。

(1) $AB=\{a0,a1,a2,b0,b1,b2,c0,c1,c2\}$

$A^0=\{\epsilon\}$

$A^2=AA=\{aa,ab,ac,ba,bb,bc,ca,cb,cc\}$

(2) 因为 $BA=\{0a,0b,0c,1a,1b,1c,2a,2b,2c\}$

所以 $AB \neq BA$

(3) AB 中的符号长度等于 2。

2-2 试构造产生如下语言的文法:

- (1) $\{a^n b^n \mid n \geq 0\}$;
- (2) 不以 0 打头的所有奇数组成的集合;
- (3) 由偶数个 0 和偶数个 1 组成的所有符号串的集合;
- (4) $\{anbn \mid n \geq 1, i \geq 0\}$ 。

答:

必须熟悉题目中所描述的语言的特点, 找到规律, 用文法的产生式将这种规律描述出来。

- (1) $G[S] = (\{S\}, \{a, b\}, \{S \rightarrow aSb \mid \epsilon\}, S)$ 。
- (2) $G[S] = (\{S, A, B, I, J\}, \{+, -, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$
 $\{S \rightarrow [+|-](J|IBJ), B \rightarrow 0B|1B| \epsilon, I \rightarrow J|2|4|6|8, J \rightarrow 1|3|5|7|9\}, S)$ 。
- (3) $G[S] = (\{S, A, B, C\}, \{0, 1\}, \{S \rightarrow 0A|1B| \epsilon, A \rightarrow 0S|1C, B \rightarrow 0C|1S, C \rightarrow 1A|0B\}, S)$ 。
- (4) $G[S] = (\{S, A, B\}, \{a, b, c\}, \{S \rightarrow AB, A \rightarrow aAb|ab, B \rightarrow Bc| \epsilon\}, S)$ 。

2-3 说明如下文法所产生的语言(集合)的特点:

- (1) $S \rightarrow SS \mid 1A0$
 $A \rightarrow 1A0 \mid \epsilon$
- (2) $S \rightarrow bAdc$
 $A \rightarrow AGS \mid a$
 $G \rightarrow \epsilon$

答:

(1) 该语言的特点是: 产生的句子中, 0、1 个数相同, 并且若干相接的 1 后必然紧接数量相同连续的 0。

(2) 该语言的特点是: 产生的句子中, 是以 ba 开头 dc 结尾的串, 且 ba、dc 个数相同。

2-4 下列各文法, 按要求做等价变换:

(1) 消去单一产生式

$$\begin{aligned} S &\rightarrow aAB \mid E \\ A &\rightarrow dDA \mid e \\ B &\rightarrow bE \mid f \\ D &\rightarrow eA \\ E &\rightarrow fA \mid g \end{aligned}$$

(2) 消除空符产生式

$$\begin{aligned} S &\rightarrow aAS \mid b \\ A &\rightarrow cS \mid \epsilon \end{aligned}$$

(3) 消除无用符号和无用产生式

$$\begin{aligned} S &\rightarrow SA \mid SB \mid A \\ A &\rightarrow B \mid (S) \mid () \end{aligned}$$

答:

掌握文法等价变换的原则, 熟悉消除无用符号和无用产生式、消除单一产生式和消除或规范空符产生式的算法。

(1) 由消除单一产生式算法 2.3 得:

$W(S)=\{S, E\}$ $W(A)=\{A\}$ $W(B)=\{B\}$

$W(D)=\{D\}$ $W(E)=\{E\}$

由 $W(S)$ 得: $S \rightarrow aAB \mid fA \mid g$

由 $W(A)$ 得: $A \rightarrow dDA \mid e$

由 $W(B)$ 得: $B \rightarrow bE \mid f$

由 $W(D)$ 得: $D \rightarrow eA$

由 $W(E)$ 得: $E \rightarrow fA \mid g$

所以等价变换后的文法为:

$S \rightarrow aAB \mid fA \mid g, A \rightarrow dDA \mid e$

$B \rightarrow bE \mid f$

$D \rightarrow eA$

$E \rightarrow fA \mid g$

(2) 由消除空符产生式算法 2.5 得:

$S \rightarrow aAS \mid aS \mid b$

$A \rightarrow cS$

(3) 由消除无用符号和无用产生式算法 2.1、2.2 得:

$S \rightarrow SA \mid A$

$A \rightarrow (S) \mid ()$

2-5 (上海交通大学 1996) 生成非 0 开头的正偶数集的文法是_____。

a. $Z \rightarrow ABC$

b. $Z \rightarrow ABC$

$C \rightarrow 0|2|4|6|8$

$C \rightarrow 0|2|4|6|8$

$B \rightarrow BA|B0|\epsilon$

$B \rightarrow BA|B0|0$

$A \rightarrow 1|2|3|\dots|9$

$A \rightarrow 1|2|3|\dots|9$

c. $Z \rightarrow ABC|2|4|6|8$

d. $Z \rightarrow ABC|2|4|6|8$

$C \rightarrow 0|2|4|6|8$

$C \rightarrow 0|2|4|6|8$

$B \rightarrow BA|B0|0$

$B \rightarrow BA|B0|\epsilon$

$A \rightarrow 1|2|3|\dots|9$

$A \rightarrow 1|2|3|\dots|9$

答:

答案是 d。

a. 不能推导出一位偶数。

b. 不能推导出所有满足条件的一位、二位偶数, 如 12。

c. 不能推导出满足条件的二位偶数, 如 12。

d. 从 Z 正好推导出语言: $\{x|x \text{ 是非 0 开头的正偶数}\}$ 。

2-6 (上海交通大学 2000) 写出不能被 5 整除的偶整数的文法。

答:

可依次构造产生一位、二位、多位满足要求的整数的产生式。

该文法为 $G[Z]$:

$Z \rightarrow [+](AB|B)$

$$A \rightarrow 0|1|2|3|4|5|6|7|8|9|AA$$

$$B \rightarrow 1|2|3|4|5|6|7|8|9$$

如果还要求整数除0外,均不以0开头,则文法为:

$$G[Z]: Z \rightarrow [+|-](AD|D)$$

$$A \rightarrow AB|C$$

$$B \rightarrow 0|5|D$$

$$C \rightarrow D|5$$

$$D \rightarrow 1|2|3|4|5|6|7|8|9$$

2-7 (中国科学院软件所 1993) 构造文法, 产生任意长的 a 、 b 串, 使得 $|a| \leq |b| \leq 2|a|$, 其中: “ $|a|$ ” 表示 a 字符的个数; “ $|b|$ ” 表示 b 字符的个数。

答:

由于要求符号串中 a 与 b 的个数满足 $|a| \leq |b| \leq 2|a|$, 所以既要用递归保证产生的任意长的字符串, 又要保证递归过程满足 a 、 b 的个数要求。所以:

$$G[Z]: Z \rightarrow aAb|bAa|abA|Aab|baA|Aba$$

$$A \rightarrow bZ|Zb|ab|ba|aZb|bZa|abZ|Zab|Zba|baZ$$

第1条规则和 $A \rightarrow ab|ba|aZb|bZa|abZ|Zab|Zba|baZ$ 保证 a 、 b 的个数一样多, 加上 $A \rightarrow bZ|Zb$, 保证递归产生任意长度的符号串, 且一旦产生 b , 就马上同时产生 a 和 b , 保证 b 的个数不超过 a 的个数的两倍。 $A \rightarrow \epsilon$ 的使用, 保证可以终止推导。

也可以将文法构造为:

$$G[S]: S \rightarrow aSBS|BSaS|\epsilon$$

$$B \rightarrow bb|b$$

规则 $S \rightarrow aSBS|BSaS$ 保证每产生一个 a , 就产生一个 B , 而 B 可以是一个 b (保证 a 和 b 个数一样多) 或者是两个 b (保证 b 的个数最多是 a 的个数的两倍)。 $S \rightarrow \epsilon$ 与 $S \rightarrow aSBS|BSaS$ 结合使用, 可以使得 a 与 b 或 b 与 a 之间有任何个 a 或 b 。

2-8 (华中科技大学 1991) 请给出描述语言 $L = \{a^{2m+1}b^m+1|m \geq 0\} \cup \{a^{2m}b^{m+2}|m \geq 0\}$ 的文法。

答:

将语言句子的描述稍作变形, 得: $a^{2m}abb^m$ 和 $a^{2m}bbb^m$, 这样一来, 发现句子中前后的 a 和 b 的个数就有了倍数关系。于是得到文法如下:

$$G[Z]: Z \rightarrow aaZb|ab|bb。$$

2-9 (北京工业大学 1985) 已知文法 $G1[A] = (\{A, B, C\}, \{a, b, c\}, P, A)$, 其中 P 有以下组成:

$$A \rightarrow abc \quad A \rightarrow aBbc$$

$$Bb \rightarrow bB \quad Bc \rightarrow Cbcc$$

$$bC \rightarrow Cb \quad aC \rightarrow aaB$$

$$aC \rightarrow aa$$

此文法所产生的语言是什么?

答:

分析文法的规则:

每使用 $Bc \rightarrow Cbcc$ 一次, b 、 c 的个数就增加一个; 每使用 $aC \rightarrow aaB$ 或 $aC \rightarrow aa$ 一次, a

的个数就增加一个；规则 $Bb \rightarrow bB$ 和 $bC \rightarrow Cb$ 起连接转换作用，并不产生新的符号 a 、 b 或 c 。

由于 A 是开始符号，由规则 $A \rightarrow aBbc$ 推导得到 B 后，每当使用规则 Bb 、 Bc 、 bC 、 aC ，就会通过 $aC \rightarrow aaB$ 递归调用 B 一次，所产生的 a 、 b 、 c 的个数分别多一个，推导得到的终结符号串为 $aabbcc$ 、 $aaabbbccc$ 、 $\dots$

所以，此文法描述的语言为： $L(G_1[A]) = \{a^n b^n c^n | n > 0\}$ 。

2-10（清华大学 1985）已知语言 $L = \{x | x \in \{a, b, c\}^*, \text{且 } x \text{ 中符号的排列是对称的（例如：} aabcbaa \text{ 或 } aabbaa \text{ 等）}\}$ ，试写出产生该语言的文法。

答：

因为 x 中的符号是对称出现的，所以，只要保证推导过程中能对称地产生出每一个符号就可以了。

因此，构造文法如下：

$G[Z]: Z \rightarrow aZa | bZb | cZc | a | b | c | \epsilon$

2-11（清华大学 1997）已知文法 $G[S]$ 为：

$S \rightarrow dAB$

$A \rightarrow Aa | a$

$B \rightarrow Bb | \epsilon$

$G[S]$ 产生的语言是什么？ $G[S]$ 能否改写为等价的正规文法？

答：

首先分析 $G[S]$ 产生的语言，得：

$L(G[S]) = \{da^n b^m | n \geq 1, m \geq 0\}$

根据语言的特点： a 、 b 的个数 n 与 m 没有相互制约关系，所以将 a^n 与 b^m 分别构造，得到正规文法如下：

$G[S]: S \rightarrow dA$

$A \rightarrow aA | aB | a$

$B \rightarrow bB | b$

2-12（中国科学院软件所 1999）构造一个文法 G ，它产生的语言为：

$L(G) = \{w | w \in \{a, b\}^*, w \text{ 中 } a \text{ 和 } b \text{ 的个数相等}\}$

答：

由于文法的句子 $w \in \{a, b\}^*$ ，所以文法的句子可分解为 3 种情形：以 a 打头的符号串、以 b 打头的符号串、 ϵ 。

首先构造以 a 打头的符号串。因为句子的长度可以无限，所以必然用到递归规则，构造规则如下：

$S \rightarrow aBS$

S 进入递归后推出符合要求的剩余部分，所以关键在于 aB 。由于 aB 中已有一个 a ，所以 B 必须保证产生一个 b ，或者产生 a 的个数比 b 的个数少 1 的符号串。因此构造 B 对应的规则如下：

$B \rightarrow b | aBB$

所以，以 a 打头的符号串的产生规则为：