

○ 计算机专业数学辅导丛书

编译原理

—习题与解析

伍春香 编著



清华大学出版社



北京科海培训中心

- 计算机专业教学辅导丛书

编译原理——习题与解析

伍春香 编著

清华大学出版社

(京)新登字 158 号

内 容 提 要

本书根据作者多年的教学实践、收集、整理而成,目的是将编译原理这门理论性较强的课程通过对习题的分析与解答来帮助读者充分理解编译的基本原理,了解和掌握基本的编译方法,提高分析问题、解决问题的能力。

全书内容涵盖了编译原理涉及的所有内容。每章精选了该章的有关概念、方法和原理,各章的习题分为基本题和习题解析,基本题分填空题与单项选择题,习题解析部分既收录了一些较容易的内容,也收集了有一定难度的研究生入学试题,针对每道试题给出问题的求解思路和解答过程,以适应不同读者的需要。

本书不仅可作为计算机专业本科生学习指导书,也适合报考相关专业的研究生作为考试辅助参考书。

版权所有·侵权必究

本书封面贴有清华大学出版社激光防伪标签,无标签者不得进入各书店。

书 名: 编译原理——习题与解析

作 者: 伍春香

出版者: 清华大学出版社(北京清华大学校内,邮编 100084)

印刷者: 北京朝阳科普印刷厂

发行者: 新华书店总店北京科技发行所

开 本: 16 印张: 20.25 字数: 492 千字

版 次: 2001 年 6 月第 1 版 2001 年 6 月第 1 次印刷

印 数: 0001~5000

书 号: ISBN 7-302-04512-7/TP · 2668

定 价: 25.00 元

前　言

编译原理是计算机专业一门重要的专业课，它介绍了编译过程中的基本原理、基本技术及其实现方法。学习这方面的知识，对计算机专业许多问题的理解能够起到融会贯通的作用。

编译原理这门课涉及内容很广，具有较强的理论性，学生掌握起来有一定的难度，解答习题也往往需要花大量的时间。编者通过多年教学实践，收集、整理资料，编写了这本《编译原理习题与解析》，目的不仅仅是对习题给出答案，更多的是想通过对习题的分析与解答帮助学生充分理解编译的基本原理，了解和掌握基本的编译方法，提高分析问题、解决问题的能力。

本书共分 14 章。第 1 章概括了编译课程的前期知识。第 2 章涉及程序的翻译和编译程序的组成。第 3 章主要是对文法、语言的形式定义及编译过程中一些基本概念的综合。第 4 章包括词法分析的基本方法和有穷自动机。第 5 章包括自顶向下分析的一些基本方法。第 6 章为自底向上分析方法。第 7 章总结了符号表及符号表在编译过程中的作用。第 8 章综合了编译程序为源程序在运行阶段的存储所做的工作。第 9 章介绍了编译过程中常用的一些中间语言。第 10 章涉及语法制导翻译。第 11 章为代码生成的有关内容。第 12 章综合了代码优化的方法。第 13 章为编译程序对源程序中错误的检测和处理方法。第 14 章给出了一些综合题。根据教学大纲要求，本书内容重点在第 3、4、5、6 章。

每章精选了该章的有关概念、方法和原理，其后的习题分为基本题和习题解析。基本题有填空题与单项选择题。习题解析部分既收录了一些较容易的内容，也有一部分是取自 1984~2000 年部分高校、科研院所的研究生入学试题，（每题题号后有标注，并根据本书内容需要，对某些题目的题型作了改动。）有一定难度。习题解析部分每一道题基本上给出了求解思路和解答过程，以适应不同层次学生的需要。

本书习题难易适当，不仅可以作为计算机专业本科生编译原理课程的参考书，也适合作为报考计算机专业研究生的学生的复习指导书，对于参加计算机软件专业技术水平考试的人来说，也有一定的参考价值。

编写本书时，编者力求将自己多年教学积累下来的一些想法以及教学中发现的学生难以理解的问题融于各章习题解答中，希望能有助于读者对编译课程所涉及知识的理解。由于习题较多，时间仓促，加上编者水平有限，书中难免有错漏之处，内容编排上也可能存在不够合理的地方，希望读者发现后能与编著者联系。诚请广大读者批评指正！

作　者

2001 年 1 月

目 录

第 1 章 预备知识	1
1.1 基本内容	1
1.1.1 程序设计语言与程序的翻译	1
1.1.2 程序设计语言的语法描述	1
1.1.3 程序的执行	2
1.2 基本题	2
1.2.1 填空题	2
1.2.2 单项选择题	2
1.3 习题解析	3
第 2 章 编译程序概述	4
2.1 基本内容	4
2.1.1 程序的翻译	4
2.1.2 编译程序的组成	5
2.2 基本题	6
2.2.1 填空题	6
2.2.2 单项选择题	7
2.3 习题解析	11
第 3 章 文法和语言的形式定义	13
3.1 基本内容	13
3.1.1 符号串	13
3.1.2 文法和语言的形式定义	14
3.1.3 与文法有关的一些关系和集合	15
3.1.4 文法的分类	15
3.1.5 语法树和二义性	16
3.1.6 文法的实用限制和文法变换	17
3.2 基本题	18
3.2.1 填空题	18
3.2.2 单项选择题	19
3.3 习题解析	25
3.3.1 文法、语言的形式定义	25
3.3.2 短语、简单短语、句柄	28
3.3.3 文法的递归	31

3.3.4 文法描述的语言和构造语言的文法.....	33
3.3.5 文法的分类	39
3.3.6 规范推导与规范归约	42
3.3.7 语法树与二义性	44
3.3.8 文法变换	47
第 4 章 词法分析与有穷自动机.....	49
4.1 基本内容	49
4.1.1 词法分析程序的任务	49
4.1.2 状态(转换)图	49
4.1.3 正则表达式和有穷自动机	50
4.1.4 词法分析程序的设计	54
4.2 基本题	55
4.2.1 填空题	55
4.2.2 单项选择题	55
4.3 习题解析	56
4.3.1 词法分析	56
4.3.2 正则式与正则语言	57
4.3.3 有穷自动机	58
第 5 章 自顶向下语法分析	88
5.1 基本内容	88
5.1.1 自顶向下语法分析	88
5.1.2 递归子程序法	89
5.1.3 LL(1) 方法(预测分析技术)	91
5.2 基本题	92
5.2.1 填空题	92
5.2.2 单项选择题	92
5.3 习题解析	96
5.3.1 自顶向下语法分析	96
5.3.2 递归子程序法	98
5.3.3 LL(1) 分析法	102
第 6 章 自底向上语法分析	110
6.1 基本内容	110
6.1.1 自底向上语法分析概述	110
6.1.2 简单优先分析算法	111
6.1.3 算符优先分析算法	113
6.1.4 LR(k) 分析程序	115

6.1.5 LR (0) 分析表的构造	116
6.1.6 SLR (1) 分析表的构造	119
6.1.7 LR (1) 分析表的构造	119
6.1.8 LALR (1) 分析表	121
6.2 基本题	121
6.2.1 填空题	121
6.2.2 单项选择题	122
6.3 习题解析	127
6.3.1 自底向上语法分析	127
6.3.2 优先分析法	127
6.3.3 LR (k) 分析法	139
第 7 章 符号表	174
7.1 基本内容	174
7.1.1 符号表的作用与生存期	174
7.1.2 符号表的内容	174
7.1.3 符号表的数据结构与符号表的组织	175
7.1.4 符号表的结构与符号表的操作	175
7.1.5 FORTRAN 的符号表	176
7.1.6 PASCAL 的符号表	176
7.2 基本题	177
7.2.1 填空题	177
7.2.2 单项选择题	177
7.3 习题解析	179
第 8 章 运行阶段的存储组织与分配	187
8.1 基本内容	187
8.1.1 概述	187
8.1.2 静态存储分配	188
8.1.3 动态存储分配	188
8.1.4 简单的栈式存储分配	189
8.1.5 以过程为单位的栈式动态存储分配方案	190
8.1.6 堆存储分配	192
8.1.7 形实参数结合	193
8.1.8 关于本章的学习	193
8.2 基本题	194
8.2.1 填空题	194
8.2.2 单项选择题	194
8.3 习题解析	197

8.3.1 存储分配	197
8.3.2 参数传递	207
第 9 章 中间语言	221
9.1 基本内容	221
9.1.1 逆波兰表示	221
9.1.2 四元式表示	221
9.1.3 三元式表示	221
9.1.4 伪(抽象机器)代码	221
9.1.5 PL/0 抽象机器的伪代码指令	222
9.1.6 形成中间形式的注意事项	223
9.2 基本题	223
9.2.1 填空题	223
9.2.2 单项选择题	223
9.3 习题解析	228
第 10 章 语法制导翻译	236
10.1 基本内容	236
10.1.1 语义分析概述	236
10.1.2 属性文法	236
10.1.3 语义分析方法	237
10.1.4 语义分析的翻译过程	238
10.2 基本题	238
10.2.1 填空题	238
10.2.2 单项选择题	239
10.3 习题解析	239
第 11 章 代码生成	251
11.1 基本内容	251
11.1.1 概述	251
11.1.2 目标代码结构	251
11.2 基本题	252
11.2.1 填空题	252
11.2.2 单项选择题	252
11.3 习题解析	253
第 12 章 代码优化	260
12.1 基本内容	260
12.1.1 优化的原则	260
12.1.2 代码优化的基本方法	260

12.1.3 基本块内的局部优化	260
12.1.4 基于循环的优化	264
12.1.5 窥孔优化	266
12.2 基本题	267
12.2.1 填空题	267
12.2.2 单项选择题	267
12.3 习题解析	269
第 13 章 错误的检测与处理	292
13.1 基本内容	292
13.1.1 错误的种类	292
13.1.2 对错误的处理方法	292
13.1.3 词法分析阶段的错误检测与处理	292
13.1.4 语法分析阶段的错误检测与处理	293
13.1.5 语义错误的检测与处理	293
13.2 基本题	294
13.2.1 填空题	294
13.2.2 单项选择题	294
第 14 章 综合题	296
14.1 基本题	296
14.1.1 选择题	296
14.2 习题解析	301
参考文献	313

第1章 预备知识

1.1 基本内容

1.1.1 程序设计语言与程序的翻译

一般的程序设计语言的定义都涉及语法、语义和语用三个方面。

1. 语法

是指由程序语言基本符号组成程序中各个语法成分（包括程序）的一组规则。其中由符号（单词）构成语法成分的规则称为一般语法规则，而由基本符号构成的符号（单词）书写规则特称为词法规则。

2. 语义

是程序设计语言中按语法规则构成的各个语法成分的意义，又分静态语义（编译时刻可确定的语法成分含义）和动态语义（运行时刻才能理解与确定的语法成分含义）。

3. 语用

表示语言符号及其使用者之间的关系，涉及符号的来源、使用和影响。

1.1.2 程序设计语言的语法描述

1. 语法图

用图解形式来描述程序设计语言语法规则的工具。这种表示简洁、直观。

2. BNF 范式

用尖括号括住待描述的语法成分，用“::=”表示语法成分的定义，“|”表示“或者”。采用这种方式描述语法具有简洁、严谨、精确、无歧义等优点。

3. 扩充的 BNF 范式

增加了三个符号：

{x} 表示 x 可以出现 0 到多次。

[x] 表示 x 可能出现，也可能不出现。

(x|y) 表示 x 和 y 二者取一。

4. 口语

是指如何用自然语言来描述程序设计语言的语法成分。这种表示方式显得不简洁甚至

累赘，也不直观，且易导致定义不确切。

1.1.3 程序的执行

正确的程序才能被执行。“一个程序是正确的”包括两层含义：一是书写正确，即合乎语法规则；二是含义正确，能够正确理解与应用程序中各种语法成分的语义定义，并在逻辑上体现了程序书写者的意图，因而在正确地输入之后，就能获得预期的运行效果。

1.2 基 本 题

1.2.1 填空题

- 一个程序是正确的，包括两层含义：一是(1)；二是(2)。

答案：(1) 书写正确（或合乎语法规则） (2) 含义正确（或合乎语义规则）

- 描述高级语言语法的常用方法有(1)、(2)、(3)等。

答案：(1) 语法图 (2) BNF 范式 (3) 扩充的 BNF 范式

- 描述如何用语言基本符号组成程序中各个语法成分的一组规则称为(1)；对程序中各个语法成分含义的描述称为(2)；而涉及语言符号及其使用者之间的关系的内容称为(3)。

答案：(1) 语法 (2) 语义 (3) 语用

1.2.2 单项选择题

- 一般程序设计语言的定义都涉及_____三个方面。

① 语法 ② 语义 ③ 语用 ④ 程序基本符号的确定

可选项有：

- a. ①②③ b. ①②④ c. ①③④ d. ②③④

答案：a

- 程序语言一般分为(1)和(2)两大类，其中(1)通常又称为面向机器的语言。面向机器语言指的是(3)，其特点是(4)，在此基础上产生了与人类自然语言比较接近的(2)。

可选项有：

- (1)、(2) a. 高级语言 b. 专用程序语言 c. 低级语言 d. 通用程序语言

- (3) a. 用于解决机器硬件设计问题的语言
b. 特定计算机系统所固有的语言

- c. 各种计算机系统都通用的语言
 - d. 只能在一台计算机上使用的语言
- (4) a. 程序的执行效率低, 编制效率低, 可读性差
 b. 程序的执行效率高, 编制效率高, 可读性强
 c. 程序的执行效率低, 编制效率高, 可读性强
 d. 程序的执行效率高, 编制效率低, 可读性差

答案: (1) c (2) a (3) b (4) d

1.3 习题解析

1. 请用语法图、BNF 范式、扩充的 BNF 范式、自然语言四种方式给出<标识符>的语法描述。

解答: (1) 标识符的构成规则可以用图 1.1 所示的语法图描述。

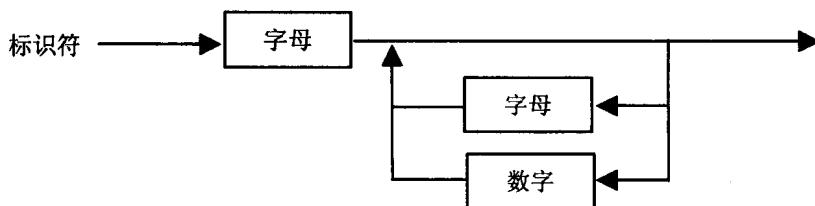


图 1.1 标识符的构成规则

- (2) 标识符的构成规则用 BNF 范式描述为:

<标识符> ::= <字母> | <标识符> <字母> | <标识符> <数字>

- (3) 标识符的构成规则用扩充的 BNF 范式可表示为

<标识符> ::= <字母> { (<字母> | <数字>) }

- (4) 可以用如下的自然语言来描述标识符的构成规则:

标识符是由字母后跟若干个(包括 0 个)字母或数字的符号串组成的。

2. 试以 PASCAL 语言的赋值语句为例, 说明什么是一个语法成分的语法、语义、语用。

解答: 赋值语句语法的非形式化描述为: 赋值语句由一个变量, 后随一个符号“:=”, 再在后面跟一个表达式所构成。

赋值语句的语义为: 先对该语句的右部表达式求值, 然后把所得结果与语句左部的变量相结合, 并取代该变量原有的值。

赋值语句的语用为: 赋值语句可用来计算和保存表达式的值。

第 2 章 编译程序概述

2.1 基本内容

2.1.1 程序的翻译

1. 程序设计语言与程序的翻译

程序设计语言有多种：不需翻译就可直接执行的机器语言；机器指令助记符形式的汇编语言，以及接近自然语言的高级程序设计语言（如 FORTRAN, PASCAL, C 等）。

除机器语言程序外，用其他语言书写的程序都必须经过翻译才能被计算机识别，这一过程由翻译程序来完成。所谓翻译程序是指这样一种程序，它能将用甲语言编写的程序翻译成与之等价的乙语言程序。甲语言称为该翻译程序的源语言，乙语言称为该翻译程序的目标语言。用源语言书写的程序称为源程序，与源程序等价的目标语言程序称为目标程序。

程序的翻译通常有两种方式：一种是“编译”方式，另一种是“解释”方式。

2. 编译方式与编译程序

编译方式是一种分阶段进行的方式。一般说来，首先进行“翻译”，把用高级语言或汇编语言编写的程序翻译成与之等价的机器语言程序，然后对翻译出来的程序进行运行计算。前一阶段的翻译工作由翻译程序完成，后一阶段的运行计算需要有运行程序来配合完成。

如果翻译程序是将用高级语言书写的源程序翻译成与之等价的某计算机的汇编语言程序或机器语言程序，则这种翻译程序特称为编译程序。

如果翻译程序是将汇编语言的程序翻译成某台计算机的机器语言程序，则这种翻译程序特称为汇编程序。

用机器语言构成的目标程序又称为目标代码程序或简称为代码程序，有时又称为目标代码或结果代码。

所谓运行程序是指运行目标代码程序时必须配置的各种子程序全体，通常以库子程序的形式存在，如一些连接装配程序及一些连接库等。

3. 解释方式与解释程序

在解释方式下，源程序的执行只有一个阶段——解释执行阶段。具体地说，完成解释工作的解释程序将按源程序中语句的动态顺序逐句地进行分析解释，并立即予以执行。

在解释方式下，并不生成目标代码，而是直接执行源程序本身。这是编译方式与解释方式的根本区别。

2.1.2 编译程序的组成

1. 编译程序的组成

每一个编译程序都要完成五个方面的任务，完成这五个方面任务的程序分别为：词法分析程序，语法分析程序，语义分析程序，代码生成程序，代码优化程序。

(1) 词法分析程序（又称扫描器）

依据语言词法规则，分析由字符组成的源程序，把它识别为一个一个具有独立意义的最小语法单位，即“单词”，并识别出与其相关的属性（如标识符，或界限符，或数，等等），再转换成长度统一的标准形式（这种统一的标准形式既刻画了单词本身，又刻画了它所具有的属性，称为属性字），以供其他部分使用。

词法分析程序往往还完成那些在语法分析之前需做的工作：如删除注解之类非必要信息，处理编译程序的控制指示，把标识符登录入符号表以及加工宏功能等各项预处理。

(2) 语法分析程序

依据语言的语法规则，逐一地分析词法分析时得到的单词（属性字），以确定它们是怎样构成说明和语句的，以及说明和语句是怎样组成程序的。分析时如发现有不合语法规则的地方，便把这些出错的地方及出错性质打印输出报告给程序员；如无语法错误，则以另一种内部表示（如语法分析树或其他中间表示）给出正确的语法结构，供下一阶段分析使用。

(3) 语义分析程序

依据语言的语义规则对语法分析得到的语法结构进行静态语义检查（确定类型、类型和运算合法性检查、识别含义与相应的语义处理及进行其他一些静态语义检查），并用另一种内部形式表示出来，或者直接用目标语言表示出来。

(4) 代码生成程序

如果语义分析时把源程序表示成内部形式而不是表示成目标指令，则由本部分完成从内部形式到目标指令的生成工作。如果语义分析时已直接生成目标指令，则无需做代码生成工作。

(5) 代码优化程序

依据程序的等价变换规则，尽量压缩目标程序运行所需的时间和所占的存储空间，以提高目标程序的质量。

2. 编译程序的逻辑结构

(1) 遍（趟，趟程）

所谓一趟或一遍是指一个编译程序在编译时刻把源程序或源程序的等价物（中间程序）从头到尾扫描一遍并转换成另一紧邻的等价物的全过程。

(2) 单遍扫描与多遍扫描

根据编译程序在完成翻译任务过程中需要对源程序或其中间等价物扫描的遍数，可把编译程序分为单遍扫描的编译程序（只需扫描一遍）和多遍扫描的编译程序（需扫描多遍）。

在单遍扫描的编译程序中，词法分析程序往往作为语法分析的一个子程序，当语法分析需要一个新单词时，便调用词法分析程序。

(3) 典型的编译程序结构

典型的编译程序结构图如图 2.1 所示。

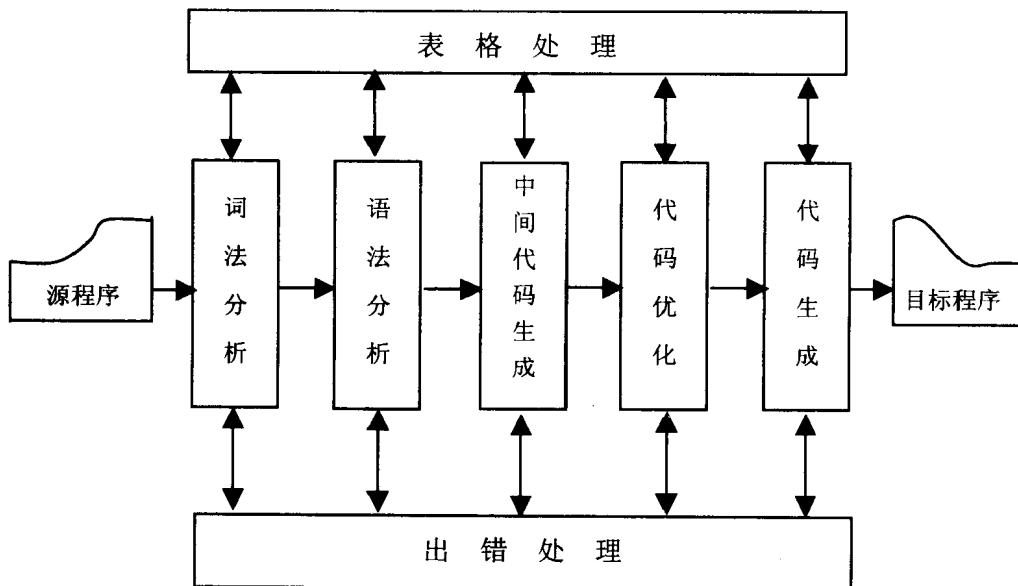


图 2.1 典型的编译程序结构

不同的编译程序其结构不一定相同，图 2.1 中给出的只是编译程序各部分的逻辑结构图，它并不代表时间上的执行顺序。有些编译程序可能恰好按图中的顺序执行这些逻辑过程，另一些编译程序可能按平行、互锁方式执行这些逻辑过程。

采用哪种执行方式，视语种、机型等因素不同而异。

2.2 基 本 题

2.2.1 填空题

1. (浙江大学 1984 年) 编译程序的工作过程一般可以划分为 (1) 等几个基本阶段，同时还会伴有 (2) 和 (3)。

答案：(1) 词法分析、语法分析、语义分析、中间代码生成、代码优化

(2) 表格处理 (3) 出错处理

2. (湖北省高等教育自学考试) 若源程序是用高级语言编写的，目标程序是 _____，则其翻译程序称为编译程序。

答案：机器语言程序或汇编语言

3. 编译方式与解释方式的根本区别在于_____。

答案：是否生成目标代码

4. (华中科技大学1985年) 翻译程序是这样一种程序，它能够将(1)转换成与其等价的(2)。

答案：(1)用甲语言书写的程序 (2)用乙语言书写的程序

5. 对编译程序而言，输入数据是(1)，输出结果是(2)。

答案：(1)源程序 (2)目标程序

6. 如果编译程序生成的目标程序是机器代码程序，则源程序的执行分为两大阶段：(1)和(2)。如果编译程序生成的目标程序是汇编语言程序，则源程序的执行分成三个阶段：(3)、(4)和(5)。

答案：(1)编译阶段 (2)运行阶段 (3)编译阶段
(4)汇编阶段 (5)运行阶段

2.2.2 单项选择题

1. (湖北省高等教育自学考试) 在使用高级语言编程时，首先可通过编译程序发现源程序的全部(1)错误和部分(2)错误。

可选项有：

- a. 语法 b. 语义 c. 语用 d. 运行

答案：(1) a (2) b

2. 一个编译程序中，不仅包含词法分析、语法分析、中间代码生成、代码优化、目标代码生成等五个部分，还应包括(1)。其中，(2)和代码优化部分不是每个编译程序都必需的。词法分析器用于识别(3)，语法分析器则可以发现源程序中的(4)。

可选项有：

- | | | | |
|--------------|------------|--------------|-----------|
| (1) a. 模拟执行器 | b. 解释器 | c. 表格处理和出错处理 | d. 符号执行器 |
| (2) a. 语法分析 | b. 中间代码生成 | c. 词法分析 | d. 目标代码生成 |
| (3) a. 字符串 | b. 语句 | c. 单词 | d. 标识符 |
| (4) a. 语义错误 | b. 语法和语义错误 | c. 错误并校正 | d. 语法错误 |

答案：(1) c (2) b (3) c (4) d

3. 程序语言的语言处理程序是一种(1)。(2)是两类程序语言处理程序，它们的主要区别在于(3)。

可选项有：

- (1) a. 系统软件 b. 应用软件 c. 实时系统 d. 分布式系统

- (2) a. 高级语言程序和低级语言程序 b. 解释程序和编译程序
 c. 编译程序和操作系统 d. 系统程序和应用程序
 (3) a. 单用户与多用户的差别 b. 对用户程序的查错能力
 c. 机器执行效率 d. 是否生成目标代码

答案: (1) a (2) b (3) d

4. (华东计算所 1985) 汇编程序是将 (1) 翻译成 (2); 编译程序是将 (3) 翻译成 (4)。

可选项有:

- | | |
|------------------|------------------|
| a. 汇编语言程序 | b. 机器语言程序 |
| c. 高级语言程序 | d. 汇编语言程序或机器语言程序 |
| e. 汇编语言程序或高级语言程序 | f. 机器语言程序或高级语言程序 |

答案: (1) a (2) b (3) c (4) d

5. 下面关于解释程序的描述正确的是_____。

- ① 解释程序的特点是处理程序时不产生目标代码
- ② 解释程序适用于 COBOL 和 FORTRAN 语言
- ③ 解释程序是为打开编译程序技术的僵局而开发的

可选项有:

- a. ①② b. ① c. ①②③ d. ②③

答案: b

6. 高级语言的语言处理程序分为解释程序和编译程序两种。编译程序有五个阶段, 而解释程序通常缺少 (1) 和 (2)。其中, (1) 的目的是使最后阶段产生的目标代码更为高效。与编译系统相比, 解释系统 (3)。解释程序处理语言时, 大多数采用的是 (4) 方法。(5) 就是一种典型的解释型语言。

可选项有:

- (1)、(2) a. 中间代码生成 b. 目标代码生成 c. 词法分析
 d. 语法分析 e. 代码优化
- (3) a. 比较简单, 可移植性好, 执行速度快
 b. 比较复杂, 可移植性好, 执行速度快
 c. 比较简单, 可移植性差, 执行速度慢
 d. 比较简单, 可移植性好, 执行速度慢
- (4) a. 源程序命令被逐个直接解释执行
 b. 先将源程序转化为中间代码, 再解释执行
 c. 先将源程序解释转化为目标程序, 再执行
 d. 以上方法都可以