

边莫英 主编

软件技术基础

```
NODE *p,*pa
pa=(NODE*)m
pa->row=i;
pa->col=j;
pa->value=v
p=matrix->d
while(p->ro
    p=p->down;
pn=p->right
while(pa->c
{
    p=pn;
    pn=pn->rig
```

天津大学出版社

内容提要

本书为计算机软件基础教程。主要内容有软件基础知识、操作系统、数据库结构、应用软件开发技术、算法基础和网络等七部分。全书各部相对独立,叙述简明扼要,结构严谨,理论联系实际,实用性强。

本书可为大专院校非计算机专业本科生、研究生教材和应用软件人员的培训教材,也可供广大科技人员参考。

(津)新登字 012 号

软件技术基础

边莫英主编

*

天津大学出版社出版

(天津大学内)

南大赛智公司照排

河北省邮电印刷厂印刷

新华书店天津发行所发行

*

开本:787×1092 毫米 1/16 印张:24.5 字数:612 千字

1993 年 11 月第一版 1993 年 11 月第一次印刷

印数:1—10000

ISBN 7-5618-0582-9

TP·55

定价:16.00 元

前 言

《软件技术基础》是为非计算机专业大学生和研究生学习计算机软件基础知识而编写的一本综合性教材,也可作为各有关计算机的技术人员和科技工作者的培训教材或参考书。

软件课程种类繁多。本书仅从实用角度出发选编如下内容:软件基础知识、操作系统、数据库、数据结构、应用软件开发技术、算法基础和计算机网络等七部分。在编写上力求简明、扼要、全面、结合实际、通俗易懂。书中各部分相对独立,自成体系,讲授时可根据专业需要在次序和内容上任意调整。

上机实践是学习本课程的重要一环。书中每章附有适量习题。附录 B 中提供了大量完整实用的 C 程序,所有程序均在 Bull 机上调试通过。考虑到有的读者未学过 C 语言,在附录 A 中全面而扼要地介绍了 C 语言的基本内容。

本书由边莫英主编。第一、五章由匙彦斌撰写,第二、七章由沙启亭撰写,第三章由赵国瑞撰写,第四、六章及附录由边莫英、王温君、刘捐献、李英慧和王津玲撰写。全书由边莫英负责组织编写、审阅、定稿。

天津大学教务处为本书编写大纲组织各系有关教师进行了讨论和修改。在本书的出版过程中,得到了天津大学教务处和天津大学出版社的大力支持,作者在此表示由衷的感谢。

编写软件综合性教材是一次尝试,错误和缺点在所难免,恳请读者提出宝贵意见。

编者 1993. 8.

目 录

第一章 计算机软件知识	(1)
§ 1.1 计算机与计算机系统	(1)
§ 1.2 计算机软件系统	(1)
1.2.1 系统软件	(2)
1.2.2 应用软件	(2)
§ 1.3 计算机各主要系统软件功能概述	(3)
1.3.1 操作系统	(3)
1.3.2 数据库管理系统	(4)
1.3.3 软件支持系统	(4)
§ 1.4 计算机语言与语言处理程序	(5)
1.4.1 计算机语言概述	(5)
1.4.2 语言处理程序简介	(6)
§ 1.5 当代微型计算机常用软件简介	(10)
1.5.1 IBM pc 微型计算机的系统软件	(11)
1.5.2 IBM pc 微型计算机的应用软件	(13)
§ 1.6 计算机软件的发展与未来	(15)
1.6.1 计算机软件的发展历史简述	(15)
1.6.2 计算机软件的发展趋势	(16)
§ 1.7 计算机病毒	(17)
1.7.1 计算机病毒的概念	(17)
1.7.2 常见计算机病毒	(19)
1.7.3 计算机病毒的预防	(20)
参考文献	(20)
 第二章 操作系统	 (21)
§ 2.1 操作系统	(21)
2.1.1 操作系统的发展	(21)
2.1.2 操作系统的功能	(22)
2.1.3 操作系统的分类	(25)
§ 2.2 计算机资源管理	(27)
2.2.1 进程与处理机管理	(27)
2.2.2 作业管理	(34)
2.2.3 存储管理	(38)
2.2.4 设备管理	(44)
2.2.5 文件管理	(47)
§ 2.3 微型计算机操作系统	(53)

2.3.1 概述	(54)
2.3.2 MS-DOS 的功能与构成	(54)
2.3.3 MS-DOS 命令和使用	(55)
2.3.4 中文 DOS 简介	(59)
§ 2.4 UNIX 操作系统	(62)
2.4.1 系统概述	(62)
2.4.2 UNIX 的命令与使用	(64)
参考文献	(67)
习题	(67)

第三章 数据库 (69)

§ 3.1 数据库基本概念	(69)
3.1.1 信息、数据和数据处理	(69)
3.1.2 计算机数据管理技术的发展	(69)
3.1.3 数据库系统的基本概念	(70)
3.1.4 信息结构的三个世界及其相互关系	(70)
3.1.5 数据模型的分类	(71)
3.1.6 用户访问数据的过程	(73)
3.1.7 数据库技术的发展方向	(74)
§ 3.2 关系数据库及其设计	(75)
3.2.1 关系模型的定义	(75)
3.2.2 关系代表	(76)
3.2.3 关系数据库 SQL 语言	(79)
3.2.4 关系数据库设计过程	(83)
§ 3.3 DBASE ■ 基本命令	(84)
3.3.1 DBASE ■ 主要性能指标和运行环境	(85)
3.3.2 数据类型和文件类型	(85)
3.3.3 DBASE ■ 命令一般格式	(87)
3.3.4 数据结构的建立和修改	(87)
3.3.5 数据库的数据输入、显示和修改	(89)
3.3.6 数据库的数据分类、索引与查找	(91)
3.3.7 数据库的数值统计	(93)
3.3.8 数据库之间的操作	(95)
§ 3.4 DBASE ■ 程序设计	(97)
3.4.1 DBASE ■ 程序的编辑与执行	(97)
3.4.2 赋值及数据输入输出命令	(98)
3.4.3 分支程序设计	(101)
3.4.4 循环程序设计	(103)
3.4.5 命令文件的嵌套调用	(105)
§ 3.5 DBASE ■ 与高级语言程序间的数据交换	(107)
3.5.1 DBASE ■ 调用高级语言程序	(107)
3.5.2 DBASE ■ 高级语言程序间的数据交换	(107)

3.5.3 DBASE Ⅱ 与 FORTRAN 程序的数据交换	(108)
3.5.4 DBASE Ⅱ 与 BASIC 程序的数据交换	(109)
§ 3.6 DBASE Ⅱ 命令、函数及 FOXBASE 简介	(110)
3.6.1 DBASE Ⅱ 函数清单	(110)
3.6.2 DBASE Ⅱ 命令清单	(111)
3.6.3 FOXBASE 简介	(114)
参考文献	(121)
习题	(121)

第四章 数据结构..... (124)

§ 4.1 引言	(124)
4.1.1 什么是数据结构	(124)
4.1.2 数据的逻辑结构	(125)
4.1.3 数据的存储结构	(125)
4.1.4 数据的运算	(126)
§ 4.2 线性结构	(126)
4.2.1 线性表	(126)
4.2.2 栈	(127)
4.2.3 队列	(128)
4.2.4 链表	(128)
§ 4.3 树形结构	(134)
4.3.1 树及其遍历	(134)
4.3.2 二叉树	(135)
4.3.3 遍历二叉树	(139)
4.3.4 线索二叉树	(140)
4.3.5 树的应用	(141)
§ 4.4 图形结构	(145)
4.4.1 图的概念	(145)
4.4.2 图的存储表示法	(146)
4.4.3 图的遍历和生成树	(147)
4.4.4 最短路径	(149)
4.4.5 拓扑排序	(152)
4.4.6 关键路径	(153)
§ 4.5 内部排序	(154)
4.5.1 插入排序	(154)
4.5.2 选择排序	(155)
4.5.3 交换排序	(157)
4.5.4 分配排序	(158)
4.5.5 归并排序	(159)
§ 4.6 检索	(159)
4.6.1 顺序检索	(159)
4.6.2 二分检索	(161)

4.6.3 分块检索.....	(161)
4.6.4 散列表检索.....	(162)
参考文献	(164)
习题	(164)

第五章 应用软件开发技术..... (168)

§ 5.1 软件工程概述.....	(168)
5.1.1 软件工程的定义.....	(168)
5.1.2 软件的生命期.....	(168)
§ 5.2 应用软件开发方法综述.....	(169)
5.2.1 应用软件开发的原则.....	(170)
5.2.2 传统的应用软件的开发方法.....	(170)
§ 5.3 结构化分析方法.....	(171)
5.3.1 SA 方法的特点	(171)
5.3.2 数据流程图.....	(172)
5.3.3 数据字典.....	(177)
5.3.4 小说明.....	(178)
§ 5.4 结构化设计方法.....	(179)
5.4.1 结构化设计方法.....	(179)
5.4.2 系统流程图法.....	(184)
5.4.3 原型化方法.....	(188)
5.4.4 结构化程序设计.....	(191)
§ 5.5 软件测试与软件开发工具简介.....	(195)
5.5.1 软件测试的概念.....	(195)
5.5.2 测试方法概述.....	(195)
5.5.3 软件开发工具简介.....	(199)
参考文献	(200)
习题	(200)

第六章 算法基础..... (202)

§ 6.1 算法设计与分析的基本概念.....	(202)
6.1.1 算法的定义.....	(202)
6.1.2 算法设计的步骤.....	(202)
6.1.3 算法的复杂性.....	(204)
§ 6.2 分治法.....	(207)
§ 6.3 贪心法	(209)
§ 6.4 动态规划.....	(213)
§ 6.5 回溯法.....	(216)
§ 6.6 穷举法.....	(221)
§ 6.7 倒推法.....	(221)
参考文献	(223)
习题	(223)

第七章 计算机网络	(225)
§ 7.1 计算机网络概述.....	(225)
7.1.1 计算机网络的发展.....	(225)
7.1.2 计算机网络的拓扑结构.....	(227)
7.1.3 计算机网络通信基础.....	(228)
§ 7.2 计算机网络体系结构.....	(231)
7.2.1 基本概念.....	(231)
7.2.2 ISO/OSI 参考模型.....	(233)
7.2.3 IEEE 802 网络协议	(235)
7.2.4 TCP/IP 网络协议	(237)
§ 7.3 微型计算机局部网络.....	(239)
7.3.1 微型计算机局部网络.....	(239)
7.3.2 微型计算机局部网络构成.....	(241)
7.3.3 NOVELL 网络	(243)
参考文献	(247)
习题	(247)

附录 A C 语言简介	(248)
§ A.1 C 语言的基本结构及程序设计	(248)
A.1.1 UNIX 和 C 语言	(248)
A.1.2 C 语言的成分	(248)
A.1.3 简单的 C 程序设计	(252)
A.1.4 流程控制	(255)
A.1.5 函数	(257)
§ A.2 数据类型	(261)
A.2.1 数组	(261)
A.2.2 指针类型	(261)
A.2.3 结构类型与联合类型	(264)
A.2.4 枚举类型	(268)
A.2.5 变量的初始化	(268)
A.2.6 预处理	(268)
§ A.3 文件	(270)
A.3.1 文件类型简介	(270)
A.3.2 文件的打开和关闭	(270)
A.3.3 文件和读写	(271)
A.3.4 文件的定位	(271)
A.3.5 出错的检测	(271)

附录 B C 程序实例	(273)
§ B.1 有关线性表运算的 C 程序	(297)
§ B.2 有关树运算的 C 程序	(298)
§ B.3 有磁图运算的 C 程序	(328)

§ B. 4 有关排序的 C 程序	(354)
§ B. 5 有磁检索的 C 程序	(376)

第一章 计算机软件知识

§ 1.1 计算机与计算机系统

电子计算机是一种帮助人们进行计算与处理数据的电子工具。它能够快速而准确地进行各种数值运算,又能自动完成大量的数据处理,也能够在生产现场实现复杂生产过程的自动控制。从目前情况来看,大至宇宙航行,小至个人的家庭琐事,从简单的加减运算到模拟人的大脑思维活动,都已经使用了电子计算机。

电子计算机的功能确实十分强大,其应用范围已波及到人类活动的各个领域。所以如此,最重要的原因就是计算机具有丰富的“功能资源”。功能资源有硬资源和软资源。硬资源包括计算机的全部硬设备和功能部件,如输入输出部件、信息的存储部件、计算机工作的控制部件及加工数据的运算部件等,习惯上称这些部件为硬件;软资源主要指为完成计算机的某种功能(如调度、控制等)用一些代码编写的工作步骤。这部分统称为计算机的软件,是计算机的灵魂。这两大部分组成一个统一的整体,使计算机协调一致地工作,形成了电子计算机的资源系统,简称计算机系统。

计算机系统工作的协调一致性,主要体现在以下几方面。

①计算机的工作方式是全自动化的。它可以在预先编排好的工作步骤指挥下,有条不紊的自动工作下去。这是计算机与迄今为止人类发明的一切其他运算工具的本质区别。

②计算机的动作是按一定的指令进行的。就计算机本身而言,它的能力极为有限,但它有一组由其本身的硬件所预先确定的指令系统。指令系统中的一条指令,可以完成一种对数据进行处理的操作,如存数、取数操作,加减操作,数据的传输操作等。虽然不同的计算机系统,指令的类别、数量、功能各不相同,但就整个计算机而言,这些指令为计算机一切复杂工作提供了最原始但功能足够充分的动作元。实际上,计算机之所以可以完成各种复杂的运算和数据处理,而且能有条不紊地工作,均是由这些基本指令的反复动作实现的。

③计算机的工作步骤是靠计算机程序指挥的。前面已经提到,计算机是当今自动化程度最高的运算工具,但计算机本身并不具备这种功能。它能在没有人的干预下连续、自动地进行各种运算和处理,完全是由计算机程序来指挥的。所谓计算机程序,是人们为解决某种问题用计算机可以识别的代码(称为计算机语言)编排的一系列加工步骤。计算机能严格按照这些步骤去做,包括计算机对其资源的调度、指挥等,都是人预先安排的。可以说,如果没有计算机程序指挥计算机工作,计算机将是一堆废铜烂铁。

计算机程序是计算机的灵魂,也是计算机功能的扩展。程序给人们提供了使用计算机解决问题的途径,也简化了使用计算机的过程。计算机中这些专用的或公用的程序系列,由于已经成了计算机系统不可分割的组成部分,因此称这些程序系列(包括各种程序的使用说明等文字资料)为计算机软件资源,或称计算机软件。

§ 1.2 计算机软件系统

计算机软件资源是计算机系统的重要组成部分。由于它们形成了一个决定着计算机功能的

整体,因此又称为计算机软件系统。它是计算机发展中最为活跃的因素。世界各国都投入了大量的人力和物力开发和研制新型软件。计算机软件的技术成就标志着计算机应用和发展水平。

计算机软件系统一般分成两大类:一类是系统软件;另一类是应用软件。

1.2.1 系统软件

一、系统软件及其功能

计算机系统软件是处于计算机硬件与用户之间的软件。它具有计算机各种应用所需的全部通用功能。但由于计算机技术发展很快,系统软件与应用软件的界限往往不十分清楚。

一般说来,计算机系统软件是指包含具有下列功能的程序:

- ① 高效使用硬件功能的程序;
- ② 提供各种应用的通用服务功能的程序;
- ③ 与其他计算机或设备进行通信时,控制通信处理的程序;
- ④ 保护数据和程序等信息,支持计算机系统正确、安全运行的程序;
- ⑤ 支持计算机系统预防故障、处理异常的程序;
- ⑥ 提供编制软件环境、提高软件编制效率的程序;
- ⑦ 使计算机系统操作高效、简便的支持程序等。

二、系统软件的种类

系统软件主要包括以下几种类别的程序。

1) 控制程序。这主要是指对作业、任务、资源、数据及故障的处理等进行管理和控制的程序。控制程序是操作系统的主体组成部分。

2) 通信程序。这主要是指对通信网络各层间的通信规程进行处理并进行通信故障检出和网络运行控制的程序。

3) 语言处理程序。这是与用户接触最多的一类软件。它的作用是将程序设计语言书写的源程序翻译成机器语言程序,主要包括汇编程序、编译程序和解释程序等。

4) 服务程序。它的作用是给程序的执行和大量数据操作提供通用服务功能,如连接编辑程序、输入编辑程序和分类合并程序等。

5) 数据库管理程序。它是将各种文件、数据统一管理,使其内容规范化、结构化,达到统一使用的目的,其功能包括控制和管理数据库的建立、检索、更新和故障的处理等。

6) 软件编程支持程序。这是为提高软件编制效率和质量而设计的程序。其中有多种支持程序与软件开发过程中各个阶段相对应,如系统分析工具、代码生成程序、软件测试程序等。

1.2.2 应用软件

应用软件是计算机所有应用程序的总称。应用软件可分为两类,一类是不分业务、行业而可以公共应用的软件;另一类是按业务、行业分类的应用软件。

一、公共应用软件

公共应用软件可以分成以下几组:

- ① 数据处理类软件,如进行数值分析、统计分析及模拟处理的程序等;
- ② 进行声音、图形、图像、文献等信息处理的程序;
- ③ 进行各种有用信息的检索处理的程序;
- ④ 关于自然语言处理、模式识别、专家系统等人工智能方面的应用程序;

⑤计算机辅助设计与制造、计算机辅助教学、计算机辅助分析以及决策支持系统等方面的通用程序。

二、按行业、业务分类的应用软件

这种软件专业性很强,有的仅可用于某一个行业或业务领域,如应用于医疗保健、教育、服务等方面的应用软件;应用于防灾、防范、宗教、法律方面的应用软件等。

当然,软件的分类并不是固定不变的,分类形式也多种多样。系统软件、应用软件的区分有时也相当灵活。尤其是系统软件与公共应用软件更难区别。

§ 1.3 计算机各主要系统软件功能概述

计算机系统软件为使用计算机提供了方便的操作环境。它的作用随着计算机技术的不断发展越来越重要,包括的内容也越来越广。本节将最主要的几个系统软件的作用做一介绍,目的是使读者对计算机的工作先有一个概括的了解。这些软件的详细内容在以后的章节中叙述。

1.3.1 操作系统

操作系统是一个庞大的程序。它是计算机软件的核心。在它的控制下,计算机的全部资源可以协调一致地工作。它可以有条不紊、高效率地管理和调度计算机的硬件设备和各种软件资源,使它们最大限度地发挥作用,在以下三个方面实现其管理目标。

1)充分发挥计算机的工作效率。计算机的处理速度是很高的。为了能有效地利用其高速、高效的优势,计算机的工作过程应当是全自动化的。它具有自动调度、自动协调、自动管理、自动运行的能力。就计算机本身而言,它仅是一个可以进行四则运算和逻辑运算的工具。如果没有一个统一指挥和调度的管理程序,上述功能难以实现。操作系统就是选择一个软件,它可以最大限度地减少人对机器工作的干预,并可在工作过程中,随时接受和处理用户提交的作业。

2)处理高速中央处理机与低速外部设备的速度匹配问题。计算机的中央处理机(称做CPU,即Central processing unit)的处理速度每秒钟几百万次。各种外部设备,如打印机、绘图仪,由于多为机械式的,处理速度每分钟几十次到几百次,二者相差几百万倍乃至上千万倍。如何解决二者在速度上的不匹配,即既不能让高速的中央处理机等待外部设备,又不能让被加工之后且需输出的数据堆积或丢失,也需要有统一协调和统一调配的软件。操作系统的一个重要功能称为“中断”,就是以某种优先顺序模式,按任务的轻、重、缓、急,在总的队列中,可以优先或靠后处理。这就为用户提供了一种简便的、统一调用外部设备的手段。

3)能使多个用户同时共用一台计算机。计算机的处理能力很强,只要配备足够的软件,可同时为多个不同的用户服务。此时从用户的角度来看计算机,似乎每个用户拥有一台自己的计算机一样。要做到这一点,也必须有一个统一管理和协调一致的软件,既可管理甲用户的要求,又可管理乙用户的要求。操作系统允许在一台计算机上装配多个用户操作设备(称为用户终端)。它可以随时接受和处理来自不同终端上的作业,并使这些不同用户作业在它的统一管理下同时运行。

为了实现上述管理目标,目前的操作系统一般由五个管理模块组成。它们是处理机管理模块、内存管理模块、文件系统管理模块、设备管理模块和作业控制管理模块。尽管这些模块有明确分工,但基本目的不外乎有两个,一是管理好计算机的各种资源,使它们协调一致地工作;二是给用户提供一个使用方便、操作简单的计算机系统。

根据管理目标的不同,操作系统有不同类型:有以自动地、顺序地执行作业流为目标的批处理

操作系统;有以多个作业分时共享资源和多个部件并行工作为目标的分时操作系统;有以要求及时接收来自现场数据并快速加以分析和做出必要反应为目标的实时操作系统。近 10 年来,由于计算机技术的飞速发展,操作系统也相应产生了新的类型,如虚拟机操作系统、网络操作系统和分布式操作系统等。

1.3.2 数据库管理系统

数据库是 60 年代末期开始发展起来的一门新技术,是用计算机管理信息资源的先进工具。这门技术目前已广泛应用于各个方面,如科技情报检索、图书管理、库存管理、会计核算、银行业务、人事管理、医院病历管理等等。当然,在这类问题的处理中,也会有一些数值计算任务,但大量的、占主导作用的是解决事务处理方面的非数值计算。这种非数值计算,由于涉及到大量的数据,因此不称做数据处理。现在,计算机在数据处理方面的应用已占很大比例。据不完全统计,当今世界总台数的百分之八十左右用于数据处理业务,而且这个比例还在继续上升。数据处理中应用计算机,促进了社会生产力的发展,同时也促进了计算机技术的发展。目前的微型计算机及其各种数据处理软件的出现和普及,有力地说明了数据处理技术越来越广泛地被人们所接受。

数据处理是一种计算过程简单的数据加工处理过程。它要涉及大量数据,数据间又存在着复杂的逻辑关系。数据处理业务的重要问题在于,如何对参与处理的数据进行有效的管理,以充分发挥计算机的作用,达到最佳的适应数据处理的要求。数据库技术的出现和广泛应用正是这一课题研究和实践的结晶。

数据库系统是一个复杂的计算机应用系统。它是当代计算机系统中最活跃、发展最快的一个分支。它包括四部分,即硬件、软件、数据库、数据库管理员。由于数据处理的特点,数据库系统对计算机硬件和软件均有特殊要求。比如,数据库系统要求计算机具有足够大的存储空间,以存储操作系统、数据库管理系统程序、用户应用程序及数据等。数据库系统的软件,最核心的是数据库管理系统(DBMS)。它的主要功能是允许用户逻辑地、抽象地处理和使用数据,而不必涉及这些数据在计算机中是怎样存放的。

DBMS 提供两种数据语言,一是数据描述语言(DDL),又称数据定义语言;另一是数据操作语言(DML),又称数据处理语言。前者的作用是描述数据的逻辑结构,即定义模式和子模式,并描述数据的物理特征,描述逻辑数据到物理数据的映射系统及访问数据的规则等。而后者的功能则因不同的 DBMS 模型而异,一般来说它具有以下功能:

- ①可以完成从数据库中检索数据;
- ②可以实现向数据库中添加数据;
- ③可删除数据库中过时的数据;
- ④可修改某些变化了的数据;
- ⑤可以进行并发访问(即两个以上用户同时访问某一数据)的控制等。

以上只是从计算机系统的角度简要地介绍了数据库系统的概貌,较详细地讲述参见第三章。

1.3.3 软件支持系统

软件支持系统是计算机软件系统中的重要组成部分。它主要指的是以下三方面的应用程序。

一、系统测试程序和系统诊断程序

这是为计算机硬件维护人员提供的一种维护工具。它可以在操作系统控制下运行,也可以在没有操作系统的情况下独立运行。该类软件主要在计算机系统出现较大的故障、而且一时无法查清

是硬件原因还是软件破坏造成的时候使用。它测试、诊断、记录并输出各主要部件的工作情况。对于大型计算机系统,对各主要部件及设备,如磁盘机、主机、内存存储器、通道、电源等均有相应的测试程序和诊断程序。严格说来,测试程序仅可测试有无问题,诊断程序可以查出问题的性质及发生的部位。根据程序功能的不同,测试程序或诊断程序有的是以软件形式提供,使用时由引导程序(一般为固化的程序)导入,并启动进行工作;也有的制成固件,一通电开机即可自行引导进行工作。如 IBM-PC 系列机的诊断程序就是如此。

二、程序编辑与调试的服务程序

这是为用户简便、高效地编辑和调试源程序服务的支持软件。常用的有编辑程序(如行编辑程序、全屏幕编辑程序等)和调试程序、链接程序、查错程序及装配程序等。在广泛使用的 IBM PC 系列机上,广为应用的 EDLIN、DEBUG、PE、CCED 等均属此类软件。

三、软件开发支持工具

这是近些年来发展起来且正在迅速推广的一种软件支持程序。它们对应于软件开发各个阶段,都有软件开发工具与其适应。如应用于系统分析的系统分析工具,应用于系统设计阶段的系统设计自动生成工具及自动编程工具、软件测试工具等。

§ 1.4 计算机语言与语言处理程序

1.4.1 计算机语言概述

目前世界上经常使用的计算机语言虽然只有十几种,但已经设计和实现的却有上千种之多。这些计算机语言可以分成两大类:一类是因不同的计算机主机而异的汇编语言;另一类是通用的程序设计语言。前者称为较低级语言;后者称为高级语言。

高级语言还可以从不同角度再做分类。

从应用范围来分,有通用语言和专用语言之分。目标非单一的语言称为通用语言,如用于数值计算方面的 FORTRAN 语言、ALGOL 语言、BASIC 语言;用于进行数据处理及商业、企业管理的通用 COBOL 语言等。目标单一的语言称为专用语言,如数控语言 APT、系统模拟语言 MIMIC 等。

从使用方式来看,有交互式和非交互式之分。具有反映人一机交互作用的语言,有 BASIC 语言、LISP 语言等;而 FORTRAN、COBOL、PASCAL 等语言则属非交互式语言。

从编程方法来看,有过程式语言和非过程式语言之分。如 FORTRAN 语言、PASCAL 语言、C 语言等均属过程式语言。因为在程序设计时,必须用语言的语句、函数、命令一步一步描述解决问题的处理过程。而有的语言,如数据库管理语言 FOXBASE、人工智能语言 PROLOG 等,编程时,只要说明和定义输入、输出及要求做的工作即可,因此,它们称为非过程式语言。

下面列出几种常用的语言及其主要应用场合:

- ①FORTRAN (Formula Translation) 公式翻译语言,主要用于科学计算。
- ②COBOL (Common Business Oriented Language) 用于商业、金融业务及企业管理方面。
- ③BASIC (Beginner's All-purpose Symbolic Instruction Code) 是用于计算、管理等多面的通用高级语言。近几年该语言功能扩充很多,它几乎包括了其他高级语言的全部优点,如绘图、动画和对系统硬件的操作、控制等。

④C 是较靠近机器而又不被机器所束缚的一种程序设计语言,多用于系统软件的开发及公用程序设计。

⑤PASCAL 是一种结构化程序设计语言,主要用于构造数据结构的场合,也是在系统软件开发时常用的一种语言。

⑥APT (Automatically Programming Tools) 自动数控程序语言,主要用于数控机床,具有自动编程功能。

⑦LISP (LIST Processing) 表处理语言,用于智能软件的开发,是迄今在人工智能领域内应用最广泛的一种语言。

⑧PROLOG (PROgramming in logic) 是一种处理逻辑问题的语言,也是一种人工智能语言,用于专家知识的描述和模拟,是第五代计算机的主要语言。

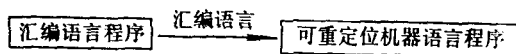
⑨PL/I (Programming Language) 是一种大型汇集型语言。它汇集了当时通用语言的各种功能,如同公共汽车一样可供广泛使用,故有“公共汽车”语言之称。

1.4.2 语言处理程序简介

用上述各种语言(包括汇编语言在内)书写的源程序,都不能直接在计算机上运行,而必须经过语言处理程序翻译加工之后方可执行。一般将源程序变换成语义等价的较低级语言程序的过程称为语言处理过程。语言处理过程根据源程序处理方法与处理结果的不同,可分为汇编处理、编译处理 and 解释处理。现分述如下。

一、汇编处理程序

汇编处理程序,又叫汇编系统。它的功能是把用汇编语言书写的源程序翻译成机器语言程序,如下图所示:



汇编语言是一种较低级语言,它接近于机器语言。由于它每一条指令与机器语言的指令保持了一一对应关系,所以汇编处理过程就是对汇编指令逐行进行翻译的过程。基本处理步骤是:

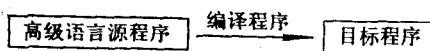
- ①将指令的助记符操作码转换成相应的机器操作码;
- ②将指令中的符号操作数转换成相应的地址码;
- ③将操作码和操作数构造成机器指令。

在具体的处理过程中,汇编程序通过对源程序进行逐行扫描的办法,得到一系列的表,然后将这些表作为产生目标程序的基础,形成目标代码程序。在两遍扫描汇编系统中,第一遍扫描是把源程序中出现的名字(包括标号、变量和常数名字)一个个地按统一格式放入一个表格之中,并确定每个名字将占用的内存位置。第二遍扫描时,按造出的表,把每条符号化指令代以真正的代码形式的机器指令,并将每个符号名字分配一个真正的地址,这叫代真。最后完成汇编处理任务。

汇编语言是为特定的计算机系统设计的面向机器的语言。但用汇编语言书写的源程序可以在不同的机器上进行汇编和执行时,称为定义汇编。

二、编译处理程序

编译程序,又叫编译系统。它的功能是把高级语言编写的面向过程的源程序翻译成目标程序的一种语言处理程序。其功能可用下图表示:



一般来说,一个编译程序把一个源程序翻译成目标程序的工作可分为前后衔接的五个阶段,即词法分析、语法分析、中间代码生成、优化和目标代码生成。相应于这五个阶段的工作任务,一

个编译程序也由五个相应的部分组成,它们分别是读词程序(或称词法分析程序)、语法分析程序、语句的处理程序、优化程序及装配程序。

1、编译程序的工作流程

编译程序的工作流程因扫描遍数的多少而异。为完成上述五个阶段的工作,在采用多遍扫描的编译程序中,一般每经过一遍扫描产生一种中间语言代码,前面的加工结果,又当作后面加工的对象。图 1-1 是一个三遍扫描的编译程序的工作流程图。

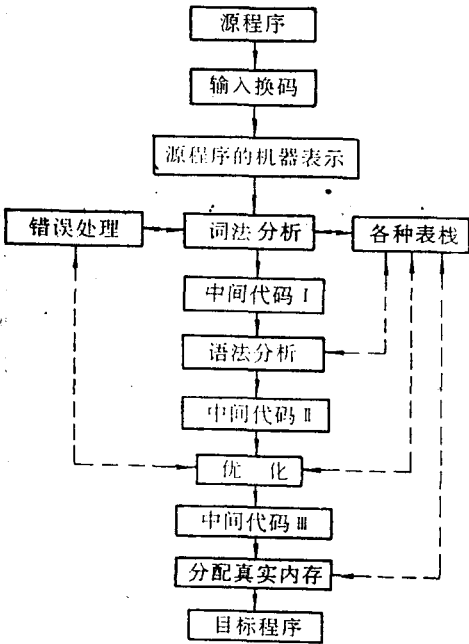


图 1-1 三遍扫描的编译程序工作流程图

2、编译过程各阶段功能简述

(1)词法分析 词法分析是编译程序的基础,也是下一步进行语法分析的必要准备。在多遍扫描的编译程序中,词法分析是用一次扫描完成的。词法分析的任务是从左至右逐个字符扫描源程序,识别出具有独立意义的语法符号,同时指出这些语法单位的属性。

众所周知,一个高级语言中,有独立意义的语法单位可分成五类,即保留字(如 FORTRAN 语言中的各种语句、函数等),语言允许的常数,用来表示变量名、数组名和过程名的标识符,语言中允许使用的运算符、分界符,如分号、逗号、回车换行符号等。词法分析程序工作时,就是把源程序中的上述基本符号分离出来。

目前,词法分析的方法有多种,但常用的有两个,一是直接分析法;另一是状态矩阵法。

直接分析法是根据读入的语法符号第一个字符的类别,转到各子程序去处理。这些子程序的功能就是识别以相应字符开头的各种语法符号。各种语言中,各种语法符号的开头字符均有规律。比如 FORTRAN 语言中,由基本字符组成的语法符号有以下几类:

- 以字母开头的有符号名、格式语句的字域说明符及专用定义符(如语句、函数等);
- 以小数点开头的有实型数、逻辑运算符等;
- 以数字开头的有常数、格式语句中的重复次数等;

以 * 开头的有乘法运算符、乘方运算符。

除上述外,其他均为一个基本符号表示一个语法符号。因此,只要编制出处理以上几种情况的子程序,即可达到识别各种语法符号的目的。

至于状态矩阵法,是根据特定语言的各种符号与处理方法,用一个矩阵表描述出来,词法分析程序从表开始进行识别与分析,达到分离出各语法单位的目的。

(2)语法分析 语法分析是编译程序的核心。它是以单词符号为输入,分析单词符号串是否形成符合语法规则的语法单位(即语句),最后是否构成一个符合要求的最终语法单位——程序。

语法分析的任务具体有以下三点。首先,识别出各种类型的语句,并做语法检查。然后,在识别各种语句的过程中,根据当时被处理的语法符号的语义,分别转到各种加工子程序去进行处理。这些加工包括对各类符号名分配单元地址;确定标号和地址的对应关系;记录各过程段的有关信息及分析循环的嵌套结构等。上述这些加工的结果,要分别记录在名表、标号表、过程表、循环层次表等一系列表格中。最后,产生目标程序或半目标程序。

语法分析的方法也有多种,一般有自上而下的分析方法和自下而上的分析方法。其中用得比较多的是递归下降法(属自上而下的分析方法)和算符优先法(属自下而上的分析方法)。下面以算符优先法为例,说明语法分析的过程。

算符优先法的基本思想是,把算法语言的语法单位分成两类,一类叫广义运算对象(简称运算对象),包括符号名、常数和标号等;另一类叫广义运算符,包括各种运算符、分隔符、拼写定义符及定界符。在处理和析过程中,采用一种类似算法语言的语言规则,使各广义运算符之间按某种优先关系进行广义运算。这种语言规则称为算符优先文法语言,并采用 $<\cdot$ 、 $\cdot$ 、 $>$ 三个记号表示两个广义运算符之间的优先关系:

$Q_1 < \cdot Q_2$ 表示运算符 Q_1 优先性低于 Q_2 ;

$Q_1 \cdot > Q_2$ 表示运算符 Q_1 优先于 Q_2 ;

$Q_1 \cdot = Q_2$ 表示运算符 Q_1 等于 Q_2 。

这样,可将被分析对象包含的算符之间的优先关系用一张二维表(称为优先矩阵)表示出来。若分析的是一数学表达式,则优先矩阵表如表 1-1 所示。

表 1-1 某语言算术表达式的优先矩阵

WRD SK 顶	┐	)	+	-	*	/	**	(	=
┐	$\cdot >$								$< \cdot$
=	$\cdot >$		$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	
(		$\cdot >$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	
+	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	
-	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$< \cdot$	$< \cdot$	$< \cdot$	$< \cdot$	
*	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$< \cdot$	$< \cdot$	
/	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$< \cdot$	$< \cdot$	
**	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$\cdot >$	$< \cdot$	

表中 WRD 存放刚读入的一个语法符号,SK 顶为栈 SK 顶当前的语法符号。