

依据教育部考试中心2002年最新大纲编写
National Computer Rank Examination

全国计算机等级考试

三合一精典版本

上机指导·应试指导·模拟试题

三级数据库技术

全国计算机等级考试命题研究组 编



海潮出版社

TP312
J55v



全国计算机等级考试

49/50

三合一精典版本

三级数据库技术上机指导·应试指导·模拟试题

全国计算机等级考试命题研究组 编

本书是为了配合全国计算机等级考试三级数据库技术而编写的应试辅导用书，全面覆盖了三级数据库技术考试的要求及范围。

全书共有三部分，第一部分是应试指导，由考试大纲串讲、精典例题分析、实战模拟练习组成；第二部分是专门针对上机考试编写的，内容主要包括考试要求、考试环境及大量的上机实践练习题；通过本章的学习，考生可以对上机考试的内容事先做到心中有数，更好地通过上机考试；第三部分是依据最新大纲设计的全套模拟试题及标准答案。

本书的作者均是从事多年计算机等级考试命题及辅导工作的专家，其中有的同志在计算机等级考试中担任过考官，具有丰富的经验。

本书由张翼、赵伟群主编。在编写本书系列丛书的过程中，得到了许多同志的帮助，在此表示衷心的感谢。

1. 与大纲同步，与教材吻合，突出重点难点，针对性强，便于考生学习。

2. 预测考试命题，精心设计练习，提高应试能力。

由于时间仓促，不足之处在所难免，敬请读者批评指正。

海潮出版社

图书在版编目(CIP)数据

三级数据库技术上机指导·应拭指导·模拟试题三合一精典版本/全国计算机等级考试命题研究组编.北京:海潮出版社,2002
(全国计算机等级考试丛书)

ISBN 7-80151-480-7

I.三… II.全… III.数据库 - 工程技术人员 - 水平考试 - 自学参考资料 IV.
TP312

中国版本图书馆 CIP 数据核字(2002)第 045873 号

丛 书 名: 全国计算机等级考试丛书

书 名: 三级数据库技术上机指导·应拭指导·模拟试题三合一精典版本

责任编辑: 孟庆华 宋树根

经 销: 全国各地新华书店

印 刷: 北京时事印刷厂

版 次: 2002 年 8 月第 1 版

印 次: 2002 年 8 月第 1 次印刷

开 本: 850×1168 1/16 字数:2000 千字

印 张: 200

书 号: ISBN7-80151-480-7/G.178

总 定 价: 350.00 元(共 12 册)

前 言

计算机与计算机科学正以无比的优越性和强劲的势头迅猛地进入人类社会的各个领域,急剧地改变着人们的生产方式和生活方式,而信息化社会必然对人才的素质及其知识结构提出新的要求。各行各业的人员不论年龄、专业 and 知识背景如何,都应掌握和应用计算机,以便提高工作效率和管理水平。既掌握一定的专业技术,又具备计算机应用能力的人员越来越受到用人单位的重视和欢迎。21 世纪将是信息时代,计算机技能是当今世界的“第二文化”。

国家教育部考试中心顺应社会发展的需要,于 1994 年推出“全国计算机等级考试”,其目的是以考促学,向社会推广普及计算机知识,为选拔人才提供统一、公正、客观和科学的标准。现在全国每年都有百万人参加这种考试。根据我国计算机应用水平的实际情况,教育部

参加全国等级考试的许多人都普遍感到,这种考试与传统考试不同,除指定的教材外,缺少关于应试指导以及模拟试题方面的资料,为此,为配合社会各类人员参加考试,并能顺利通过“全国计算机等级考试”,我们组织多年从事辅导计算机等级考试的专家在对近几年的考题深刻分析、研究基础上,编写出这套指导应考者备考和参加考试的辅导资料——计算机等级考试上机指导·应试指导·模拟试题三合一精典版本丛书。

本书是为了配合全国计算机等级考试三级数据库技术而编写的应试辅导用书,全面覆盖了三级数据库技术考试的要求及范围。

全书共有三部分,第一部分是应试指导,由考试大纲串讲、精典例题分析、实战模拟练习组成,第二部分是专门针对上机考试编写的,内容主要包括考试要求、考试环境及大量的上机实战练习题,通过本章的学习,考生可以对上机考试的内容事先做到心中有数,更好地通过上机考试;第三部分是依据最新大纲设计的全真模拟试题及标准答案。

本丛书的作者均是在各高等学校或研究单位工作、具有丰富教学和研究经验的专家、教授,其中有的同志在计算机教育界中享有盛名,颇有建树,并且编写过多种计算机书籍。

本书由张翼、赵晓环主编。作者提示本系列丛书的特点如下:

1、与大纲同步,与教材吻合,突出重点难点,针对考生学习规律有的放矢。让考生得到学习质量和效率双收益。以应试为目标,既强调知识体系,又着重基本功训练,从理论和实践的相结合上,让学生准确高效进入应试状态。

2、预测考试命题,精心设计模拟试卷,掌握学习要点,提高作题速度,巩固所学知识,熟练答题技巧,以期事半功倍。在本丛书的帮助下,您将会顺利通过考试。

由于时间仓促,不足之处在所难免,恳请广大读者批评指正。

全国计算机等级考试命题研究组

2002 年 8 月

目 录

三级教程数据库技术等级考试大纲	(1)
第一部分 应试指导	(3)
1.1 考试大纲串讲	(3)
1.1.1 基础知识	(3)
1.1.2 数据结构与算法	(7)
1.1.3 操作系统	(20)
1.1.4 数据库技术基础	(30)
1.1.5 关系数据库系统	(32)
1.1.6 关系数据库标准语言 SQL	(34)
1.1.7 关系数据库的规范化理论与数据库设计	(36)
1.1.8 数据库管理系统	(40)
1.1.9 事务管理与数据库安全性	(46)
1.1.10 新一代数据库应用开发工具	(47)
1.1.11 数据库技术的发展	(52)
1.2 精典例题分析	(55)
1.3 实战模拟练习	(66)
1.3.1 实战模拟练习(一)	(66)
实战模拟练习(一)参考答案	(69)
1.3.2 实战模拟练习(二)	(70)
实战模拟练习(二)参考答案	(87)
1.3.3 实战模拟练习(三)	(89)
实战模拟练习(三)参考答案	(104)
1.3.4 实战模拟练习(四)	(109)
实战模拟练习(四)参考答案	(119)
1.3.5 实战模拟练习(五)	(122)
实战模拟练习(五)参考答案	(127)
1.3.6 实战模拟练习(六)	(128)
实战模拟练习(六)参考答案	(133)
1.3.7 实战模拟练习(七)	(134)
实战模拟练习(七)参考答案	(144)
1.3.8 实战模拟练习(八)	(148)
实战模拟练习(八)参考答案	(148)
1.3.9 实战模拟练习(九)	(149)
实战模拟练习(九)参考答案	(156)

1.3.10 实战模拟练习(十)	(157)
实战模拟练习(十)参考答案	(160)
1.3.11 实战模拟练习(十一)	(161)
实战模拟练习(十一)参考答案	(164)
第二部分 上机指导	(165)
2.1 考试环境	(165)
2.2 Turbo C系统的上机操作	(166)
2.3 Turbo C2.0 标准函数	(205)
2.4 实战模拟练习	(294)
2.5 实战模拟练习参考答案	(412)
第三部分 全真模拟试题	(449)
模拟试题(一)	(449)
模拟试题(一)参考答案	(456)
模拟试题(二)	(457)
模拟试题(二)参考答案	(463)
模拟试题(三)	(464)
模拟试题(三)参考答案	(471)
模拟试题(四)	(472)
模拟试题(四)参考答案	(479)
模拟试题(五)	(480)
模拟试题(五)参考答案	(486)
(68)	参考答案(一)区测试题答案
(69)	参考答案(二)区测试题答案 2.8.1
(70)	参考答案(二)区测试题答案
(71)	参考答案(二)区测试题答案
(72)	参考答案(三)区测试题答案 6.8.1
(73)	参考答案(三)区测试题答案
(74)	参考答案(四)区测试题答案 4.8.1
(75)	参考答案(四)区测试题答案
(76)	参考答案(五)区测试题答案 2.8.1
(77)	参考答案(五)区测试题答案
(78)	参考答案(六)区测试题答案 0.8.1
(79)	参考答案(六)区测试题答案
(80)	参考答案(七)区测试题答案 7.8.1
(81)	参考答案(七)区测试题答案
(82)	参考答案(八)区测试题答案 8.8.1
(83)	参考答案(八)区测试题答案
(84)	参考答案(九)区测试题答案 9.8.1
(85)	参考答案(九)区测试题答案

三级教程数据库技术等级考试大纲

基本要求

1. 掌握计算机系统和计算机软件的基本概念、计算机网络的基本知识和应用知识、信息安全的基本概念。
2. 掌握数据结构与算法的基本知识并能熟练应用。
3. 掌握并能熟练运用操作系统的基本知识。
4. 掌握数据库的基本概念,深入理解关系数据模型、关系数据理论和关系数据库系统,掌握关系数据语言。
5. 掌握数据库设计方法,具有数据库设计能力。了解数据库技术发展。
6. 掌握计算机操作,并具有用 C 语言编程,开发数据库应用(含上机调试)的能力。

考试内容

(一)基础知识

1. 计算机系统的组成和应用领域。
2. 计算机软件的基础知识。
3. 计算机网络的基础知识和应用知识。
4. 信息安全的基本概念。

(二)数据结构与算法

1. 数据结构、算法的基本概念。
2. 线性表的定义、存储和运算。
3. 树形结构的定义、存储和运算。
4. 排序的基本概念和排序算法。
5. 检索的基本概念和检索算法。

(三)操作系统

1. 操作系统的基本概念、主要功能和分类。
2. 进程、线程、进程间通信的基本概念。
3. 存储管理、文件管理、设备管理的主要技术。
4. 典型操作系统的使用。

(四)数据库系统基本原理

1. 数据库的基本概念,数据库系统的构成。
2. 数据模型概念和主要的数据模型。
3. 关系数据模型的基本概念,关系操作和关系代数。
4. 结构化查询语言 SQL。
5. 事务管理、并发控制、故障恢复的基本概念。

(五)数据库设计和数据库应用

1. 关系数据库的规范化理论。
2. 数据库设计的目标、内容和方法。
3. 数据库应用开发工具。
4. 数据库技术发展。

(六)上机操作

1. 掌握计算机基本操作。
2. 掌握 C 语言程序设计基本技术、编程和调试。
3. 掌握与考试内容相关的知识的上机应用。

考试方式

(一)笔试:120 分钟

(二)上机考试:60 分钟

第一部分 应试指导

1.1 考试大纲串讲

1.1.1 基础知识

本章考试要求是:1.计算机系统的组成和应用领域。2.计算机软件的基础知识。3.计算机网络的基础知识和应用知识。4.信息安全的基本概念。

(一) 计算机发展阶段、应用领域、分类、主要技术指标

1. 第一台计算机

1946年,美国宾夕法尼亚大学研制成功全世界第一台电子数字计算机ENIAC,用电子管和继电器等元器件制成,面积 170m^2 ,重约30吨,耗电140KW。

2. 计算机时代

人们通常按计算机所使用的元器件来划分计算机发展的几个时代:

第一代是电子管计算机(1946~1957年),第二代是晶体管计算机(1958~1964年),第三代是中、小规模集成电路计算机(1965~1970年),第四代是大规模集成电路计算机(1971年至今)。

曾经有第五代计算机的说法,即基于处理知识的计算机,但并未获得广泛的认同。尽管近年来计算机的新技术层出不穷,但迄今为止,尚没有哪一种技术足以成为新一代计算机的标志,所以有人认为现在是无代计算机时代。

3. 微处理器、微计算机、单片机

微处理器和单片机是1971年问世的,这对计算机的发展和应用具有极其重大的意义。

微处理器(Microprocessor,MP)是以单片大规模集成电路制成的具有运算和控制功能的处理器。

微计算机(Microcomputer)是以微处理器作为中央处理器(CPU)的计算机。

单片机是在单个芯片上集成了微计算机的CPU、存储器、输入/输出接口电路等各部件的可嵌入各种工业或民用设备的极小的计算机。

微处理器的代表产品有4位的4004,8位的8088、280。从16位开始Intel公司的系列产品最具代表性。

4. 计算机的应用领域

计算机的应用可归纳为如下5个领域:

①科学计算,包括计算在科学研究和工程设计中遇到的大量复杂、难度较大的数学计算问题,要求快速和准确的计算结果。

②数据通信与数据处理,包括企、事业的管理营运中存在的大量数据搜集及统计工作,其特点是计算比较简单,但数据量特别大,是目前计算机应用最多的领域。

③自动控制,用于工业和民用设备的计算机自动控制。

④计算机辅助设计(CAD)与计算机辅助制造(CAM),可大大提高生产率,并使整个生产过程可以达到最优化。

⑤计算机人工智能,包括专家系统、模式(声、图、文)识别、机器翻译等。

5. 计算机的硬件组成

计算机硬件可分为5大部分:CPU(中央处理器)、主存储器、总线、输入/输出设备和辅助存储器。

6. 计算机分类

计算机可分为超级计算机、大型计算机、小型计算机和微型计算机。

超级计算机有许多CPU同时并行处理,运算速度可达每秒万亿次。小型计算机现在则作为网络的高性能服务器。微型计算机也称为个人计算机(PC),得到了最广泛的应用。个人计算机又可分为台式机(桌面机)和便携机(笔记本计算机)。随着因特网的普及应用,有一些简易的上网设备出现,这类设备都属于网络计算机(NC)的范畴。

7. CISC 计算机和 RISC 计算机

CISC(Complex Instruction Set Computer)即复杂指令集计算机,其指令种类与数量较多,以提供更完善的指令系统功能。RISC(Reduced Instruction Set Computer)即精简指令集计算机,其指令集中的指令数量较少,但使用频率高、速度快。在相应的硬件和软件的配合下,可以获得较高的性能/价格比。可以认为,CISC和RISC是计算机指令系统设计的两种风格,各适用于不同的情况,而Pentium II及其以上的CPU具有二者的优点。

8. 计算机主要技术指标

字长——进行运算的二进位数目,又称为位宽,如8位、16位、32位、64位等。字长越大,运算精度越高。

运算速度——一般用每秒钟执行的指令条数来表示。例如,每秒执行定点指令的平均数目,单位是MIPS(Million Instruction Per Second),即每秒百万条指令。也有用每秒执行浮点指令的平均数目来表示的,单位是MFIPS(Million Floating Instruction Per Second),即每秒百万条浮点指令。

主存容量——以字节为基本单位,如KB($1\text{KB}=1024\text{B}$),MB($1\text{MB}=1024\text{KB}$),GB($1\text{GB}=1024\text{MB}$)等。目前主存储器采用MOS集成电路制成,其存取时间(从给定地址到读出或写入数据的时间)约为几十纳秒(ns)。

综合性能——计算机的综合性能不仅与CPU、内存与外存的配置等硬件有关,还与系统软件和应用软件的配置情况有关。为了使测试结果能更接近于实际情况,常采用基准程序测试法(Benchmark),即通过模拟用户的实际负载,编制一组基准测试程序来测试计算机系统的性能。Intel公司对PC的性能测试就包含了四个方面:办公效率性能,多媒体运算性能,3D/浮点性能和Internet性能,并推出ICMP(Intel Comparable Microprocessor Performance)指数,作为综合反映微处理器的性能指标。例如,P II/350和P III/500的Icomp指数分别为1000和1650。

(二) 计算机系统、存储系统、输入和输出

1. 计算机系统的组成,16位及32位微机的结构及工作原理;

2. 存储系统的组成,随机存储器、磁盘、磁带和光盘;

3. 总线结构,中断方式与查询方式,A/D、D/A 转换,输入输出设备。

(三) 软件基础知识

1. 软件的基本概念,程序与文档,汇编与反汇编,解释与编译,实用程序及集成软件;

2. 软件保护与标准化。

(四) 计算机网络基础

1. 计算机网络基本概念

(1) 计算机网络的形成与发展

(2) 计算机网路的主要特征

资源共享观点将计算机网络定义为“以能够相互共享资源的方式互联起来的自治计算机系统的集合”。

资源共享观点的定义符合目前计算机网络的基本特征。

2. 计算机网络的分类

(1) 网络分类方法

计算机网络的分类方法可以是多样的,其中最主要的两种方法是:

- 根据网络所使用的传输技术(transmission technology)分类。
- 根据网络的覆盖范围与规模(scale)分类。

(2) 广域网

广域网(Wide Area Network, WAN)也称为远程网。

目前的广域网应具有以下特点:

- 适应大容量与突发性通信的要求;
- 适应综合业务服务的要求;
- 开放的设备接口与规范化的协议;
- 完善的通信服务与网络管理。

(3) 局域网

局域网(Local Area Network, LAN)是继广域网之后又一个网络研究与应用的热点,也是目前技术发展最快的领域之一。

局域网的技术特点主要表现在以下几个方面:

① 局域网覆盖有限的地理范围,它适用于公司、机关、校园、工厂等有限范围内的计算机、终端与各类信息处理设备联网的需求。

② 局域网提供高数据传输速率(10 Mb/s~1 000 Mb/s)、低误码率的高质量数据传输环境。

③ 局域网一般属于一个单位所有,易于建立、维护与扩展。

④ 决定局域网特性的主要技术要素为网络拓扑、传输介质与介质访问控制方法。

⑤ 从介质访问控制方法的角度看,局域网可分为共享式局域网与交换式局域网两类。

(4) 城域网

城域网(MAN, Metropolitan Area Network)是介于广域网与局域网之间的一种高速网络。城域网设计的目标是要满足几十公里范围内的大量企业、机关、公司的多个局域网互联的需求,以实现大量用户之间的数据、语音、图形与视频等多种信息的传输功能。早期的城域网产品主要是光纤分布式数据接口(Fiber Distributed Data Interface, FDDI)。

3. Internet 基础

(1) Internet 的形成与发展

(2) Internet 的结构与组成

(3) TCP/IP、域名与 IP 地址

① TCP/IP 的基本概念

TCP/IP 具有以下几个特点。

- 开放的协议标准,独立于特定的计算机硬件与操作系统。
- 独立于特定的网络硬件,可以运行在局域网、广域网,更适用于互联网中。
- 统一的网络地址分配方案,使得整个 TCP/IP 设备在网中都具有惟一的 IP 地址。
- 标准化的高层协议,可以提供多种可靠的用户服务。

② 域名与 IP 地址

4. Internet 提供的主要服务

(1) WWW 服务

(2) 电子邮件服务

5. Internet 的基本接入方式

(1) ISP 的作用

Internet 服务提供者(ISP)是用户接入 Internet 的入口点。一方面,它为用户提供 Internet 接入服务;另一方面,它也为用户提供各类信息服务。

一般来说,用户计算机接入 Internet 的方式主要有两种:通过局域网接入 Internet;通过电话网接入 Internet。

(2) 通过局域网接入 Internet

(3) 通过电话网接入 Internet

(五) 信息安全基础

1. 信息安全

信息安全从简单的意义来理解,就是要防止非法的攻击和病毒的传播,以保证计算机系统和通信系统的正常运作。而从更全面的意义来理解,就是要保证信息的保密性(confidentiality)、完整性(integrity)、可用性(availability)和可控性(controllability)。综合起来,就是要保障电子信息的有效性。

2. 信息保密

信息的保密是信息安全的重要方面,为保密而进行加密是防止破译信息系统中机密信息的技术手段。加密的办法就是使用数学方法来重新组织数据域信息,使除合法接收者外,其他任何人要想看懂变化后的数据或信息是非常困难的。一般人们将加密前的称为明文,而将加密后的称为密文,因此加密的目的就是将明文变为密文。而反过来将密文变为明文的过程则称为解密。加密技术可以使某些重要的数据或信息存放在一般的不安全的计算机上或在一条一般的不安全的信道上传送。只有持有合法解密办法的人才能获取明文。

3. 信息认证

信息认证是安全的另一重要方面。信息认证,首先是验证信息的发送者的真实性,即不是假冒的;其次是验证信息的完整性,即验证信息在传送或存储过程中未被篡改、重放或延迟

等。认证是防止对系统进行主动攻击,如伪造、篡改的重要手段。在有关认证的实用技术中,主要的有数字签名技术、身份识别技术和信息的完整性校验技术等。

(1) 数字签名

(2) 身份识别

(3) 消息认证

4. 密钥管理

密钥管理影响到密码系统的安全,而且还会涉及到系统的可靠性、有效性和经济性。

密钥管理包括密钥的产生、存储、装入、分配、保护、丢失、销毁以及保密等内容。其中解决密钥的分配和存储是最关键和有技术难点的问题。

5. 计算机病毒的基本概念

计算机病毒是一种特殊的具有破坏性的计算机程序,它具有自我复制能力,可通过非授权入侵而隐藏在可执行程序或数据文件中。当计算机运行时,源病毒能把自身精确拷贝或者有修改地拷贝到其他程序体内,影响和破坏正常程序的执行和数据的正确性。

(1) 计算机病毒的特征

(2) 病毒的破坏作用

(3) 病毒的来源

(4) 病毒的防治

6. 网络安全

(1) 构成对网络安全威胁的主要因素及相关技术

(2) 网络安全服务的主要内容

7. 操作系统安全

(1) 操作系统安全方法

(2) 操作系统安全措施

(3) 文件保护与保密

8. 数据库安全

(1) 安全性措施的层次

(2) 权限和授权

(3) 在 SQL 中进行安全性说明

1.1.2 数据结构与算法

本章考试要求是:1. 数据结构、算法的基本概念。2. 线性表的定义、存储和运算。3. 树形结构的定义、存储和运算。4. 排序的基本概念和排序算法。5. 检索的基本概念和检索算法。

本章内容主要是:数据结构、算法的基本概念;线性表逻辑结构,链表、数组的存储和运算;队列与栈的定义,存储及应用;树和二叉树的定义,互相转换,二叉树的存储,二叉树的周游;排序的基本概念与排序算法(选择排序,插入排序,交换排序,归并排序);检索的基本概念与检索算法(顺序检索,二分检索,散列技术,二叉排序树)。

以下介绍一些常用的数据结构,阐明各种数据结构内在的逻辑关系,讨论它们在计算机中的存储表示,以及在哪些数据结构上进行的各种运算和实际的执行算法,并对算法的效率进行

简单的分析。

(一) 什么是数据结构

数据是描述客观事物的数字、字符以及所有能直接输入到计算机中并被计算机程序处理的符号的集合。

数据对象是具有相同性质的数据元素的集合。通常,一个数据对象中的数据元素不是孤立的,而是彼此之间存在着一定的联系,这种联系就是数据结构。数据对象中数据元素之间的联系需要在对数据进行存储和加工中反映出来,因此,数据结构概念一般包括三方面的内容:数据之间的逻辑关系、数据在计算机中的存储方式、以及在这些数据上定义的运算的集合。

(1) 数据的逻辑结构

数据的逻辑结构只抽象地反映数据元素之间的逻辑关系,它与数据的存储无关,是独立于计算机的。

数据的逻辑结构分为线性结构和非线性结构两大类,线性结构的逻辑特征是:有且仅有一个开始结点和一个终端结点,并且所有的结点都最多有一个直接前驱和一个直接后继。线性表就是一个典型的线性结构。非线性结构的逻辑特征是:一个结点可能有多个直接前驱和直接后继。树、图等都是非线性结构。

(2) 数据的存储结构

数据的存储结构是数据的逻辑结构在计算机存储器里的实现(亦称为映象)。它是依赖于计算机的,并有四种基本的存储映象方法。它们是:

①顺序存储方法 该方法是把逻辑上相邻的结点存储在物理位置上相邻的存储单元内,结点间的逻辑关系由存储单元的邻接关系来体现。顺序存储方法主要用于线性的数据结构,非线性的数据结构也可以通过某种线性化方法来实现顺序存储。

②链接存储方法 在链接存储方法中,逻辑上相邻的结点在物理位置上未必相邻,结点间的逻辑关系是由附加的指针字段表示的。

③索引存储方法 该方法通常是在存储结点信息的同时,还建立一个附加的索引表,索引表中的每一项称为索引项,索引项的一般形式是:(关键字,地址)。关键字是能唯一标识一个结点的那些数据项。

④散列存储方法 在散列存储方法中,结点的存储地址是根据结点的关键字值直接计算出来的。

上述四种基本的存储方法也可以组合起来对数据结构进行存储映象。

(3) 数据的运算

数据的运算定义在数据的逻辑结构之上,每种逻辑结构都有一个运算的集合。常用的运算有:查找、插入、删除、更新、排序等。显然,对数据运算的具体实现方法只有在确定了存储结构之后才能加以考虑。

(二) 算法

(1) 算法及其特征

简单地说,一个算法就是一种解题方法,更严格地说,算法是由若干条指令组成的有穷序列,它必须具有以下特征:

①有穷性 一个算法必须在执行有穷步后结束。

②确定性 算法的每一步必须是确切地定义的,无二义性。

③可行性 算法中的所有待实现的运算必须在原则上能够由人使用笔和纸在做有穷次运算后完成。

④输入 一个算法具有 0 个或多个输入的外界量,它们是算法开始前对算法最初给出的量。

⑤输出 一个算法至少产生一个输出,它们是与输入有某种关系的量。

算法的含义与程序十分相似,但二者又有区别。一个程序不一定满足有穷性,操作系统就是如此,只要整个系统不被破坏,操作系统就永远不会停止,所以操作系统程序不是一个算法。另外,程序中的指令必须是机器可以执行的,而算法中的指令则无此限制。但是,一个算法如果用机器可执行的语言书写,则它就是一个程序。

对一个算法的描述可以采用自然语言、数学语言、约定的符号语言、以及图解等方式。

(2)算法的分析

求解同一个问题可以有多种不同的算法,评价一个算法的优劣除了正确性和简明性外,主要考虑两点:一是执行算法所耗费的时间,二是执行算法所耗费的存储空间,特别是辅助存储空间的耗费。就这两者而言,前者显得比后者更为重要,在数据结构中往往更注重对算法执行时间的分析。

一个算法所耗费的时间是该算法中每条语句的执行时间之和,而每条语句的执行时间是该语句执行次数(频度)与该语句一次执行所需时间的乘积。如果假定每条语句一次执行所需的时间均为单位时间,则一个算法的时间耗费就是该算法中所有语句的频度之和。

(三) 线性表

(1)线性表及其基本操作

线性表是 $n \geq 0$ 个元素的一个有限序列: $(a_1, a_2, a_3, \dots, a_{n-1}, a_n)$

表中元素的个数 n 称为表的长度,长度 $n=0$ 的表称为空表。

表元素又称为结点,线性表的一个重要特性是可以按照诸元素在表中的位置确定它们在表中的先后次序。若 $n \geq 1$,则 a_1 为第一个元素, a_n 为最后一个元素。元素 a_{i-1} 先于 a_i ,我们称 a_{i-1} 为 a_i 的前驱; a_i 在 a_{i-1} 之后, a_i 为 a_{i-1} 的后继。除第一个元素外,每个元素都有一个且仅有一个直接前驱;除最后一个元素外,每个元素都有一个且仅有一个直接后继,下面所列的是其中一些常用的运算。

①查找运算

- 查找线性表的第 i ($0 \leq i \leq n-1$) 个表元;
- 在线性表中查找具有给定键值的表元;

②插入运算

- 把新表元插在线性表的第 i ($0 \leq i \leq n$) 个位置上;
- 把新表元插在具有给定键值的表元的前面或后面;

③删除运算

- 删除线性表的第 i ($0 \leq i \leq n-1$) 个表元;
- 删除线性表中具有给定键值的表元;

④其他运算

- 统计线性表元的个数;
- 输出线性表各表元的值;
- 复制线性表;
- 线性表分析;
- 线性表合并;
- 线性表排序;
- 按某种规则整理线性表。

(2) 线性表的存储

有多种存储方式能将线性表存储在计算机内,其中最常用的是顺序存储和链接存储。

① 线性表的顺序存储

线性表的顺序存储是最简单的存储方式。程序通常用一个足够大的数组,从数组的第一个元素开始,将线性表的结点依次存储在数组中。即线性表的第 i 个结点存储在数组的第 i ($0 \leq i \leq n-1$) 个元素中,用数组元素的顺序存储来体现线性表中结点的先后次序关系。用数组存储线性表的优点是能直接访问线性表中的任一结点。

用数组存储线性表的缺点主要有两个:一是程序中的数组通常大小是固定的,可能会与线性表的结点可以任意增加和减少的要求相矛盾;二是执行线性表的结点插、删操作时要移动存于数组中的其他元素,使插和删操作不够简便。

② 线性表的链接存储

线性表链接存储是用链表存储线性表,最简单的用单链表。如从链表的第一个表元开始,将线性表的结点依次存储在链表的各表元中。即线性表的第 i 个结点存储在链表的第 i ($0 \leq i \leq n-1$) 个表元中。链表的每个表元除要存储线性结点的信息外,还要有一个成分用来存储其后继结点的指针。单链表就是通过链接指针来体现线性表中结点的先后次序关系。每个链表还要有一个指向链表的第一个表元,链表的最末一个表元的后继指针值为空。用链表存储线性表的优点是线性表的每个表元的后继指针就能完成插或删的操作,不需移动任何表元。

其缺点也主要有两条:一是每个表元增加了一个后继指针成分,要花费更多的存储空间;二是不便随机地直接访问线性表的任一结点。

(3) 线性表上的查找

线性表上的查找运算是指在线性表中找某个链值的结点。根据线性表的存储形式和线性表本身的性质差异,有多种查找算法,如:顺序查找、二分法查找、分块查找、散列查找等。

(4) 线性表的新结点插入

顺序存储线性表的插入:

设线性表结点的类型为整型,插入之前有 n 个结点,把值为 x 的新结点插在线性表的第 i ($0 \leq i \leq n$) 个位置上。完成插入主要有以下几个步骤:

- 检查插入要求的有关参数的合理性;
- 把原来第 $n-1$ 个结点至第 i 个结点依次往后移一个数组元素位置;
- 把新结点放在第 i 个位置上;
- 修正线性表的结点个数。

(5) 栈

堆栈的工作原理是采用后进先出(LIFO)技术,栈顶由中央处理器中的堆栈指示器(SP)指出。在执行 PUSH 操作中 SP 减量,而在 POP 操作中 SP 增量。

下面从数据结构的角度,进一步说明堆栈的基本概念与操作。需要说明的是,其工作原理与前面所介绍的是一致的,不同的是脱离了硬件背景,例如,栈顶指针不是中央处理器的某个寄存器的内容,而是一个抽象的数据结构。

栈是一种特殊的线性表,这种线性表只能在固定的一端进行插入和删除操作。允许插入和删除的一端称为栈顶,另一端称为栈底。一个新元素只能从栈顶一端进入,删除时,只能删除栈顶的元素,即刚刚被插入的元素。由于元素是按后进先出的次序入栈和出栈的,所以栈又称后进先出表(Last In First Out),简称 LIFO 表。

栈的基本操作有:

- ①create(s) 建立一个空栈 s。
- ②empty(s) 测试栈是否为空栈。
- ③full(s) 测试栈是否满。
- ④push(x,s) 将元素 x 插入栈 s 的栈顶。
- ⑤top(s) 取栈顶元素。
- ⑥pop(s) 删除栈顶元素。

由于栈是一种特殊的线性表,栈的各种操作实际上是线性表的操作的特殊情形,所以表示线性表的方法同样可以用来表示栈。

(6) 队列

队列可看作是插入在一端进行,删除在另一端进行的线性表,允许插入的一端称为队尾,允许删除的一端称为队头。在队列中,只能删除队头元素。队列的最后一个元素一定是最新入队的元素。因此队列又称先进先出表(First-In-First-Out)。

日常生活中排队购物就是队列应用的例子:新来的顾客排在队尾等待,排在队头的顾客购物后离开队伍。

队列的基本操作有:

- ①create(Q)建立一个空队列。
- ②empty(Q)测试队列是否为空队列。
- ③full(Q)测试队列是否为满。
- ④front(Q)取队头元素。
- ⑤enq(X,Q)向队列中插入一个元素 X。
- ⑥deq(Q)删除队头元素。

(四) 数组

线性表(包括栈和队列)都是线性结构,结构中的每个元素只是无结构的数据元素。我们对线性表作进一步的推广,使结构中的元素本身也可以是具有某种结构(如向量)的数据,从而引出了数组这一种新的数据结构。

(1) 数组的定义和运算

类似于线性表,一个二维数组(或称矩阵)可以看成是由 m 个行向量所组成的向量,也可以看成是由 n 个列向量所组成的向量。