



普通高等教育“十一五”国家级规划教材
普通高等教育精品教材

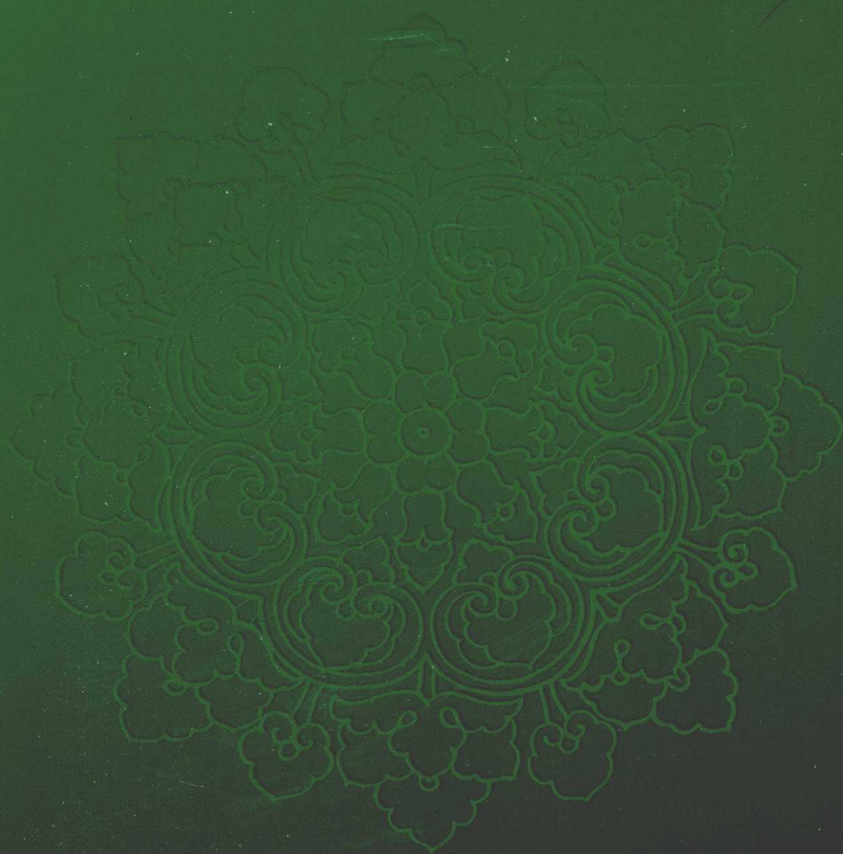
新世纪计算机基础教育丛书

丛书主编 谭浩强

计算机软件技术基础

(第三版)

徐士良 葛兵 编著



清华大学出版社



新世纪计算机基础教育丛书

丛书主编 谭浩强

计算机软件技术基础

(第三版)

徐士良 葛兵 编著

内 容 简 介

本书针对非计算机专业的大学生、研究生以及科技工作者与研究人员对计算机软件应用技术的需要,介绍了计算机软件设计的基础知识、方法与实用技术。主要内容包括:集合与算法的基本概念、基本数据结构及其运算、查找与排序技术、资源管理技术、数据库设计技术、编译技术概述、应用软件的设计与开发技术。每章都配有一定数量的习题。

本书内容丰富,通俗易懂,实用性强,可作为非计算机专业软件基础课程的教材,也可作为广大从事计算机应用工作的科技人员的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

计算机软件技术基础 / 徐士良,葛兵编著. —3版. —北京:清华大学出版社,2010.10
(新世纪计算机基础教育丛书)

ISBN 978-7-302-23407-4

I. ①计… II. ①徐… ②葛… III. ①软件—技术 IV. ①TP31

中国版本图书馆 CIP 数据核字(2010)第 153361 号

责任编辑:焦虹 李晔

封面设计:焦丽丽

责任印制:何芊

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62795954,jsjic@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015,zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185×260

印 张:22.5

字 数:540千字

版 次:2010年9月第3版

印 次:2010年9月第1次印刷

印 数:1~3000

定 价:29.50元

产品编号:038108-01

现

代科学技术的飞速发展,改变了世界,也改变了人类的生活。作为新世纪的大学生,应当站在时代发展的前列,掌握现代科学技术知识,调整自己的知识结构和能力结构,以适应社会发展的要求。新世纪需要具有丰富的现代科学知识、能够独立完成面临的任务、充满活力、有创新意识新型人才。

掌握计算机知识和应用,无疑是培养新型人才的一个重要环节。现在计算机技术已深入到人类生活的各个角落,与其他学科紧密结合,成为推动各学科飞速发展的有力的催化剂。无论什么专业的学生,都必须具备计算机的基础知识和应用能力。计算机既是现代科学技术的结晶,又是大众化的工具。学习计算机知识,不仅能够掌握有关知识,而且能培养人们的信息素养。这是高等学校全面素质教育中极为重要的一部分。

高校计算机基础教育应当遵循的理念是:面向应用需要;采用多种模式;启发自主学习;重视实践训练;加强创新意识;树立团队精神,培养信息素养。

计算机应用人才队伍由两部分人组成:一部分是计算机专业出身的计算机专业人才,他们是计算机应用人才队伍中的骨干力量;另一部分是各行各业中应用计算机的人员。这后一部分人一般并非计算机专业毕业,他们人数众多,既熟悉自己所从事的专业,又掌握计算机的应用知识,善于用计算机作为工具解决本领域中的任务。他们是计算机应用人才队伍中的基本力量。事实上,大部分应用软件都是由非计算机专业出身的计算机应用人员研制的。他们具有的这个优势是其他人难以代替的。从这个事实可以看到在非计算机专业中深入进行计算机教育的必要性。

非计算机专业中的计算机教育,无论目的、内容、教学体系、教材、教学方法等各方面都与计算机专业有很大的不同,绝不能照搬计算机专业的模式和做法。全国高等院校计算机基础教育研究会自1984年成立以来,始终不渝地探索高校计算机基础教育的特点和规律。2004年,全国高等院校计算机基础教育研究会与清华大学出版社共同推出了《中国高等院校计算机基础教育课程体系2004》(简称CFC2004);2006年、2008年又共同推出了《中国高等院校计算机基础教育课程体系2006》(简称CFC2006)及《中国高等院校计算机基础教育课程体系2008》(简称CFC2008),由清华大学出版社正式出版发行。

1988年起,我们根据教学实际的需要,组织编写了《计算机基础教育丛

书》，邀请有丰富教学经验的专家、学者先后编写了多种教材，由清华大学出版社出版。丛书出版后，迅速受到广大高校师生的欢迎，对高等学校的计算机基础教育起了积极的推动作用。广大读者反映这套教材定位准确，内容丰富，通俗易懂，符合大学生的特点。

1999年，根据新世纪的需要，在原有基础上组织出版了《新世纪计算机基础教育丛书》。由于内容符合需要，质量较高，被许多高校选为教材。丛书总发行量1000多万册，这在国内是罕见的。最近，我们又对丛书作了进一步的修订，根据发展的需要，增加了新的书目和内容。本丛书有以下特点：

(1) 内容新颖。根据21世纪的需要，重新确定丛书的内容，以符合计算机科学技术的发展和教学改革的要求。本丛书除保留了原丛书中经过实践考验且深受群众欢迎的优秀教材外，还编写了许多新的教材。在这些教材中反映了近年来迅速得到推广应用的一些计算机新技术，以后还将根据发展不断补充新的内容。

(2) 适合不同学校组织教学的需要。本丛书采用模块形式，提供了各种课程的教材，内容覆盖了高校计算机基础教育的各个方面。丛书中既有理工类专业的教材，也有文科和经济类专业的教材；既有必修课的教材，也包括一些选修课的教材。各类学校都可以从中选择到合适的教材。

(3) 符合初学者的特点。本丛书针对初学者的特点，以应用为目的，以应用为出发点，强调实用性。本丛书的作者都是长期在第一线从事高校计算机基础教育的教师，对学生的基础、特点和认识规律有深入的研究，在教学实践中积累了丰富的经验。可以说，每一本教材都是他们长期教学经验的总结。在教材的写法上，既注意概念的严谨和清晰，又特别注意采用读者容易理解的方法阐明看似深奥难懂的问题，做到例题丰富，通俗易懂，便于自学。这一点是本丛书一个十分重要的特点。

(4) 采用多样化的形式。除了教材这一基本形式外，有些教材还配有习题解答和上机指导，并提供电子教案。

总之，本丛书的指导思想是内容新颖、概念清晰、实用性强、通俗易懂、教材配套。简单概括为：新颖、清晰、实用、通俗、配套。我们经过多年实践形成的这一套行之有效的创作风格，相信会受到广大读者的欢迎。

本丛书多年来得到了各方面人士的指导、支持和帮助，尤其是得到了全国高等院校计算机基础教育研究会的各位专家和各高校老师们的支持和帮助，我们在此表示由衷的感谢。本丛书肯定有不足之处，希望得到广大读者的批评指正。

欢迎访问谭浩强网站：<http://www.tanhoqiang.com>

丛书主编
全国高等院校计算机基础教育研究会会长
谭浩强

本书第二版是普通高等教育“十一五”国家级规划教材,并被评为2008年度普通高等教育精品教材。

本次修订保持了第二版的特点(见第二版前言),主要在内容上作了如下几方面的调整:

(1) 在第1章中增加了集合方面的基本知识,对算法方面的基本知识内容作了适当的精简。

(2) 在第2章中增加了索引存储结构的内容。

(3) 在第5章中删去了关系数据库语言SQL的内容。

(4) 新增加了编译技术概述一章。

本书内容丰富,通俗易懂,实用性强,书中所有算法程序均上机调试通过。本书可作为非计算机专业的大学生或研究生软件基础课程的教材,也可作为广大从事计算机应用工作的科技人员的参考书。

由于作者水平有限,书中难免有错误或不妥之处,恳请读者批评指正。

作 者

2010年5月

非计算机专业的学生(包括广大科技人员)如何学习和掌握软件技术,不可能像计算机专业的学生那样学习软件的各门课程,因此有必要将主要的软件技术和知识在一门课程中介绍,但又不能是拼盘式的组合,在国外根本没有这样的教材。国内这样的教材也不多,有的教材虽然名称叫“软件技术”,但实际上是一些新软件的使用。本教材是介绍软件技术,而不是软件的使用。

本教材主要是针对非计算机专业的学生对于计算机软件技术的需要,将有关软件基础知识以及应用技术介绍给读者。其特点如下:

(1) 系统性。本教材并不是简单地将计算机专业的各门课程内容拼装在一起,而是以软件应用技术的需要,将它们有机地结合在一起,为读者提供软件开发中所需要的软件知识和技术。全书以数据结构与算法为基础,以软件技术为线索,系统性强。

(2) 强调应用。本教材强调以应用为目的,全书中实例比较丰富。全书几乎是围绕解决软件开发中所遇到的软件技术问题来展开。在介绍系统软件(如操作系统)时,为了便于读者理解,也适当介绍一些原理,但主要的还是介绍实现系统软件中的技术,以便读者将这些技术用到应用软件的开发中。

(3) 可读性强。本书深入浅出,通过例子引出基本概念,便于读者接受。

此次对本书的修订主要是前三章。书中所有的算法均采用C++描述。由于C++语言既可用于面向过程的程序设计,又支持面向对象的程序设计,因此,作者在对算法进行描述时,尽量采用最合适的程序设计方法。例如,对于基本的数据结构(如顺序存储与链式存储的线性表、栈、队列等)采用面向对象的方法,将数据与运算封装成类,以便在其他应用程序中直接使用;而对于同一批数据进行同类操作的各种算法(如对线性表的各种排序方法)采用面向过程的方法,将各种不同的算法用普通函数来描述。这样,书中的所有算法就可以直接在实际应用中方便地使用。

本书内容丰富,通俗易懂,实用性强,书中所有算法程序(C++描述)均上机调试通过。本书可作为非计算机专业的大学生或研究生的软件课程教材,也可作为广大从事计算机应用工作的科技人员的参考书。

由于作者水平有限,书中难免有错误或不妥之处,恳请读者批评指正。

作者

2007年1月

随着计算机技术的深入发展,计算机技术的应用已经渗透到各个领域,特别是计算机软件的设计与开发,已经不只是计算机专业人员的事情了。现在,越来越多的软件需要非计算机专业人员来设计与开发,很多系统软件与应用软件由非计算机专业人员来使用,并在此基础上进行二次开发。因此,普及计算机软件技术已经是大势所趋。

本书是在《软件应用技术基础》(该书由清华大学出版社出版,获电子工业部优秀教材一等奖)一书的基础上改写而成,使本书更适合于广大非计算机专业学习软件设计与开发的需要。作为应用计算机的科技人员,除了要掌握现有计算机软件的使用外,从实际应用出发,还必须要掌握软件设计与开发的基本知识和有关技术,如数据的组织、程序的组织、计算机资源的利用、数据处理技术等,以便得心应手地进行应用软件的设计与开发。

全书共分6章,每章后面都附有一定数量的习题。

第1章介绍算法。主要包括算法的基本概念、算法的基本设计方法、算法的复杂度分析等内容。

第2章介绍基本数据结构及其运算。主要包括数据结构的基本概念,线性表、栈、队列及其在顺序存储结构下的运算和应用,线性链表及其运算,数组,二叉树的概念、存储及其应用,图的存储及其遍历。

第3章介绍常用的查找与排序技术。主要包括基本的查找技术、哈希表技术、基本的排序技术、二叉排序树及其查找、多层索引树及其查找、拓扑分类。

第4章介绍资源管理技术。主要包括操作系统的功能与任务、多道程序设计、存储空间的组织。

第5章介绍数据库技术。主要包括数据库基本概念、关系代数、数据库设计、关系数据库语言SQL。

第6章介绍应用软件设计与开发技术。主要包括软件工程概述、软件详细设计的表达、结构化分析与设计方法、测试与调试基本技术、软件开发新技术。

本书内容丰富,通俗易懂,实用性强,书中所有算法程序均上机调试通过。本书可作为非计算机专业的大学生或研究生的软件课程教材,也可作为广大从事计算机应用工作的科技人员的参考书。

由于作者水平有限,书中难免有错误或不妥之处,恳请读者批评指正。

作 者

2001年9月

目 录



第 1 章 预备知识

1.1 集合	1
1.1.1 集合及其基本运算	1
1.1.2 自然数集与数学归纳法	3
1.1.3 笛卡儿积	5
1.1.4 二元关系	5
1.2 算法	6
1.2.1 算法的基本概念	6
1.2.2 算法设计基本方法	8
1.2.3 算法的复杂度分析	13
习题 1	15



第 2 章 基本数据结构及其运算

2.1 数据结构的基本概念	16
2.1.1 什么是数据结构	16
2.1.2 数据结构的图形表示	19
2.2 线性表及其顺序存储结构	21
2.2.1 线性表及其运算	21
2.2.2 栈及其应用	30
2.2.3 队列及其应用	40
2.3 线性链表	51
2.3.1 线性链表的基本概念	51
2.3.2 线性链表的插入与删除	55
2.3.3 带链的栈与队列	60
2.3.4 循环链表	67
2.3.5 多项式的表示与运算	70
2.4 线性表的索引存储结构	77
2.4.1 索引存储的概念	77
2.4.2 “顺序-索引-顺序”存储方式	79

2.4.3	“顺序-索引-链接”存储方式	80
2.4.4	多重索引存储结构	80
2.5	数组	82
2.5.1	数组的顺序存储结构	82
2.5.2	规则矩阵的压缩	83
2.5.3	一般稀疏矩阵的表示	86
2.6	树与二叉树	111
2.6.1	树的基本概念	111
2.6.2	二叉树及其基本性质	114
2.6.3	二叉树的遍历	117
2.6.4	二叉树的存储结构	118
2.6.5	穿线二叉树	123
2.6.6	表达式的线性化	136
2.7	图	137
2.7.1	图的基本概念	137
2.7.2	图的存储结构	138
2.7.3	图的遍历	142
2.7.4	图邻接表类	143
习题 2		149



第 3 章 查找与排序技术

3.1	基本的查找技术	152
3.1.1	顺序查找	152
3.1.2	有序表的对分查找	152
3.1.3	分块查找	157
3.2	哈希表技术	158
3.2.1	哈希表的基本概念	158
3.2.2	几种常用的哈希表	161
3.3	基本的排序技术	179
3.3.1	冒泡排序与快速排序	179
3.3.2	简单插入排序与希尔排序	184
3.3.3	简单选择排序与堆排序	186
3.3.4	其他排序方法简介	189
3.4	二叉排序树及其查找	193
3.4.1	二叉排序树的基本概念	194
3.4.2	二叉排序树的插入	195
3.4.3	二叉排序树的删除	196
3.4.4	二叉排序树查找	198

3.5 多层索引树及其查找	200
3.5.1 B ⁻ 树	201
3.5.2 B ⁺ 树	211
3.6 拓扑分类	220
习题 3	223



第 4 章 资源管理技术

4.1 操作系统的概念	225
4.1.1 操作系统的功能与任务	225
4.1.2 操作系统的发展过程	226
4.1.3 操作系统的分类	229
4.2 多道程序设计	232
4.2.1 并发程序设计	232
4.2.2 进程	235
4.2.3 进程之间的通信	239
4.2.4 多道程序的组织	244
4.3 存储空间的组织	245
4.3.1 内存储器的管理技术	245
4.3.2 外存储器中文件的组织结构	249
习题 4	256



第 5 章 数据库设计技术

5.1 数据库基本概念	257
5.1.1 数据库技术与数据库系统	257
5.1.2 数据描述	262
5.1.3 数据模型	264
5.2 关系代数	268
5.3 数据库设计	274
5.3.1 数据库设计的基本概念	274
5.3.2 数据库设计的过程	275
5.3.3 数据字典	282
习题 5	283



第 6 章 编译技术概述

6.1 编译程序的工作过程及其基本组成	285
6.1.1 编译程序的工作过程	285
6.1.2 编译程序的基本组成	286
6.2 状态矩阵法的编译过程	287

6.2.1	状态矩阵法的基本原理	287
6.2.2	状态矩阵的压缩	288
6.3	词法分析	290
6.3.1	词法分析的任务	290
6.3.2	读字符程序	291
6.3.3	状态矩阵法的词法分析过程	291
6.3.4	算术常数的识别和翻译	295
6.4	中间语言表示	297
6.4.1	波兰表示	297
6.4.2	三元组表示	300
6.5	语法的分析与加工	303
习题 6		308



第 7 章 应用软件设计与开发技术

7.1	软件工程概述	310
7.1.1	软件工程的概 念	310
7.1.2	软件生命周期	310
7.1.3	软件支援环境	313
7.2	软件详细设计的表达	314
7.2.1	程序流程图	314
7.2.2	NS 图	315
7.2.3	问题分析图 PAD	316
7.2.4	判定表	317
7.2.5	过程设计语言 PDL	318
7.3	结构化分析与设计方法	318
7.3.1	应用软件开发的原则和方法	318
7.3.2	结构化分析方法	320
7.3.3	结构化设计方法	324
7.4	测试与调试基本技术	331
7.4.1	测试	331
7.4.2	调试	337
7.5	软件开发新技术	339
7.5.1	原型方法	339
7.5.2	瀑布模型	340
7.5.3	面向对象技术	341
习题 7		343

参考文献	344
------------	-----

第1章 预备知识

1.1 集 合

1.1.1 集合及其基本运算

1. 集合的基本概念

所谓集合,是指若干个或无穷多个具有相同属性的元(元素)的集体。

通常,一个集合名称用大写字母表示,而集合中的某个元素用小写字母表示。

如果集合 M 由 $n(n \geq 0)$ 个元素 $a_1, a_2, \dots, a_n$ 组成,则称集合 M 为有限集。例如,大于 1 而小于 100 的所有整数构成的集合 A 为有限集。如果一个集合中有无穷多个元素,则称此集合为无限集。例如,所有整数构成的集合 Z ,所有实数构成的集合 R ,大于 0 而小于 1 的所有实数构成的集合 B 等均为无限集。不包括任何元素的集合称为空集。例如,大于 1 而小于 2 的整数构成的集合为空集。空集通常用 $\emptyset$ 表示。

如果 M 是一个集合, a 是集合 M 中的一个元素,则记作 $a \in M$,称元素 a 属于集合 M ;如果 a 不是集合 M 中的元素,则记作 $a \notin M$,称元素 a 不属于集合 M 。

对于一个集合,通常用以下两种方法表示。

(1) 列举法

用列举法表示一个集合是将此集合中的元素全部列出来,或者列出若干项但能根据规律可知其所有的元素。例如,上述举出的几个集合的例子可以用如下列举法表示:

大于 1 而小于 100 的所有整数的集合 A 可以表示为

$$A = \{2, 3, 4, \dots, 99\}, \quad \text{有限集}$$

所有整数构成的集合 Z 可以表示为

$$Z = \{0, \pm 1, \pm 2, \pm 3, \dots\}, \quad \text{无限集}$$

空集表示为

$$\emptyset = \{\}, \quad \text{空集}$$

(2) 性质叙述法

用性质叙述法表示一个集合是将集合中的元素所具有的属性描述出来。例如:

大于 1 而小于 100 的所有整数的集合 A 可以表示为

$$A = \{a \mid 1 < a < 100 \text{ 的所有整数}\}$$

所有整数构成的集合 Z 可以表示为

$$Z = \{z \mid z \text{ 为一切整数}\}$$

大于 0 而小于 1 的所有实数构成的集合 B 可以表示为

$$B = \{b \mid 0 < b < 1 \text{ 的所有实数}\}$$

所有实数构成的集合 R 可以表示为

$$R = \{r \mid r \text{ 为一切实数}\}$$

设 M 与 N 为两个集合。如果集合 M 中的每一个元素也都为集合 N 的元素,则称集合 M 为 N 的子集,记作 $M \subseteq N$ 或 $N \supseteq M$ 。如果 $M \subseteq N$,且 N 中至少有一个元素 $a \notin M$,则称 M 是 N 的真子集,记作 $M \subset N$ 或 $N \supset M$ 。如果 $M \subseteq N$ 且 $N \subseteq M$,则称集合 M 和集合 N 相等,记作 $M = N$ 。

2. 集合的基本运算

(1) 两个集合的并(union)

设有两个集合 M 和 N ,它们的并集记作 $M \cup N$,其定义如下:

$$M \cup N = \{\alpha \mid \alpha \in M \text{ 或 } \alpha \in N\}$$

即两个集合 M 与 N 的并集是指 M 与 N 中所有元素(去掉重复的元素)组成的集合。

(2) 两个集合的交(intersection)

设有两个集合 M 和 N ,它们的交集记作 $M \cap N$,其定义如下:

$$M \cap N = \{\alpha \mid \alpha \in M \text{ 且 } \alpha \in N\}$$

即两个集合 M 与 N 的交集是指 M 与 N 中所有共同元素组成的集合。

两个集合 M 与 N 的并、交均满足交换律,即

$$M \cup N = N \cup M$$

$$M \cap N = N \cap M$$

(3) 两个集合的差(difference)

设有两个集合 M 和 N , M 和 N 的差集记作 $M - N$,其定义如下:

$$M - N = \{\alpha \mid \alpha \in M \text{ 但 } \alpha \notin N\}$$

两个集合的差不满足交换律,即

$$M - N \neq N - M$$

例 1.1 设集合

$$A = \{a, b, c, d, e\}$$

$$B = \{d, e, f, g, h\}$$

则

$$A \cup B = \{a, b, c, d, e, f, g, h\}$$

$$A \cap B = \{d, e\}$$

$$A - B = \{a, b, c\}$$

$$B - A = \{f, g, h\}$$

对于集合的并、交、差有以下几个基本性质:

(1) 结合律

$$(A \cap B) \cap C = A \cap (B \cap C)$$

$$(A \cup B) \cup C = A \cup (B \cup C)$$

(2) 分配律

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$(3) (A - B) \cup (B - A) = (A \cup B) - (A \cap B)$$

$$(4) B \cap (A - B) = \emptyset$$

$$(A \cap B) \cup (A - B) = A$$

3. 映射

定义 1.1 设 A, B 是两个非空集。如果根据一定的法则 f , 对于每一个 $x \in A$, 在 B 中都有唯一确定的元素 y 与之对应, 则称 f 是定义在 A 上而在 B 中取值的映射, 记作 $f: A \rightarrow B$ 。并将 x 与 y 的关系记作 $y = f(x)$ 。 x 称为自变元, y 称为在 f 作用下 x 的像。

集合 A 称为 f 的定义域, $f(A) = \{f(x) | x \in A\}$ 称为 f 的值域。

定义 1.2 设给定映射 $f: A \rightarrow B$, 且 $B = f(A)$ (即 f 的像充满整个 B)。如果对于每个 $y \in B$, 仅有唯一的 $x \in A$ 使 $f(x) = y$, 则称 f 有逆映射 f^{-1} (它是定义在 $f(A)$ 上而取值于 A 的映射)。当映射 $f: A \rightarrow f(A)$ 有逆映射时, 则称 f 是一一映射。

定义 1.3 若 A, B 两集合有一一映射 f 存在, 使 $f(A) = B$, 则称 A 与 B 成一一对应。

如果集合 A 与 B 为一一对应, 则称它们互相对等, 并记作 $A \sim B$ 。当两个集合互相对等时, 称它们有相等的浓度 (或相等的元素个数)。

例 1.2 设两集合为

$$A = \{1, 2, \dots, 10\} = \{x | 1 \leq x \leq 10 \text{ 的整数}\}$$

$$B = \{1, 2, \dots, 20\} = \{y | 1 \leq y \leq 20 \text{ 的整数}\}$$

若映射 $f: A \rightarrow B$ 为

$$y = 2x$$

其中定义域为 A , 值域为

$$f(A) = \{2, 4, 6, \dots, 20\}$$

显然, 映射 f 不是一一映射。

例 1.3 设两集合为

$$A = \{x | 0 \leq x \leq 4 \text{ 的所有实数}\}$$

$$B = \{y | 0 \leq y \leq 2 \text{ 的所有实数}\}$$

考虑映射 $f: A \rightarrow B$

$$y = \sqrt{x}$$

其中定义域为 A , 值域为 $f(A) = B$ 。

显然, f 为一一映射, 集合 A 与 B 互相对等, 即 $A \sim B$ 。

集合的对等满足以下性质:

- (1) 自反性。即 $A \sim A$ 。
- (2) 对称性。即若 $A \sim B$, 则 $B \sim A$ 。
- (3) 传递性。即若 $A \sim B$ 且 $B \sim C$, 则 $A \sim C$ 。

1.1.2 自然数集与数学归纳法

由所有自然数所组成的集合

$$\{1, 2, 3, \dots\}$$

称为自然数集。自然数集是一个无限集。

由自然数组成的集合均是自然数集的子集。自然数集的子集可以是有限集, 也可以是无限集。

与自然数集对等 (即具有相等浓度) 的集合称为可列集 (或可数集)。任一可列集中的元

素排列时可标以正整数下标,即任意可列集 M 均可写成

$$M = \{a_1, a_2, \dots, a_n, \dots\}$$

关于自然数集及其子集有以下两个命题成立。

定理 1.1 在自然数集的任一非空子集 M 中,必定有一个最小数。即在集合 M 中不大于其他任意数的数。

证明: 因为 M 非空,所以在 M 中可以取得一自然数 n 。

显然, M 中所有不大于 n 的自然数形成的非空子集 N 包含在 M 中,即 $N \subset M$ 。如果 N 中有最小数,则此最小数就是 M 的最小数。

而在 N 中最多有 n 个自然数($1 \sim n$),因此, N 中有一个最小数。

综上所述,在自然数集的任一非空子集 M 中,必定有一个最小数。

定理得证。

定理 1.2 设 M 是由自然数形成的集合,如果它含有 $1, 2, \dots, k$, 并且当它含有数 $n-1, n-2, \dots, n-k$ ($n > k$) 时,也含有数 n , 那么它含有所有的自然数,即 M 是自然数集。

证明: 设 N 是所有不属于 M 的自然数形成的集合,则 $1, 2, \dots, k \notin N$ 。

现假设 N 不是空集,则由定理 1.1 可知: 在 N 中必定有一个最小数。设此最小数为 c 。

由于 c 是 N 中的最小数,即 $c \in N$, 因此, $c \notin M$, 且 $c \neq 1, 2, \dots, k$, 同时, $c-1, c-2, \dots, c-k$ 均为自然数。又由于 c 是 N 中的最小数,所以自然数 $c-1, c-2, \dots, c-k \notin N$, 即 $c-1, c-2, \dots, c-k \in M$, 而根据定理中的条件应有 $c \in M$ 。

由上所述,一方面有 $c \notin M$, 另一方面又有 $c \in M$, 这就导致矛盾。这个矛盾是由于一开始假定 N 不是空集所造成的。因此, N 只能为空集,即所有自然数均在 M 中, M 为自然数集。

定理得证。

定理 1.2 是数学归纳法的基础。通常,为了证明一个命题对于所有的自然数是真,采用数学归纳法证明的步骤如下:

(1) 证明命题对于自然数 $1, 2, \dots, k$ 是真的;

(2) 假设命题对于自然数 $n-k, n-k+1, \dots, n-1$ ($n > k$) 是真的(这一步称为归纳假设);

(3) 证明命题对于自然数 n 也是真的。

上述步骤(1)中 k 值的选取决定于在步骤(3)的证明过程中要用到归纳假设的最小自然数。在步骤(3)中,为了证明命题对于自然数 n 是真,要用到归纳假设的最小自然数如果为 $n-k$,则在步骤(1)中要对 1 到 k 中的所有自然数证明命题为真。

例 1.4 证明下列命题:

对于任意的自然数 n ,存在一对非负整数 (i, j) 有

$$7 + n = 3i + 5j$$

证明:

下面用数学归纳法证明这个命题。

(1) 当 $n=1$ 时,有 $7+1=3+5$,即 $i=1, j=1$ 。命题成立。

当 $n=2$ 时,有 $7+2=3 \times 3+5 \times 0$,即 $i=3, j=0$ 。命题成立。

当 $n=3$ 时,有 $7+3=3 \times 0+5 \times 2$,即 $i=0, j=2$ 。命题成立。

(2) 假设命题对于自然数 $n-1, n-2, n-3 (n>3)$ 成立(归纳假设)。

(3) 考虑自然数 n , 有

$$7+n = [7+(n-3)] + 3$$

根据归纳假设, 对于自然数 $n-3$ 命题成立, 设存在一对非负整数 (i_1, j_1) 有

$$7+(n-3) = 3i_1 + 5j_1$$

则有

$$7+n = [7+(n-3)] + 3 = 3(i_1+1) + 5j_1 = 3i + 5j$$

其中 $i=i_1+1, j=j_1$ 均为非负整数。即对于自然数 n 命题也成立。

由此得出结论, 对于所有的自然数 n 命题成立。

在这个例子中, 由于步骤(3)的证明过程中, 要用到归纳假设的最小自然数为 $n-3$, 因此在此步骤(1)中取 $k=3$ 。

1.1.3 笛卡儿积

在 1.1.1 节中介绍了集合的并、交、差运算。对于集合, 还有一种很重要的运算——笛卡儿积(Cartesian Product)。

设有 n 个集合 $D_1, D_2, \dots, D_n$, 此 n 个集合的笛卡儿积定义为

$$D_1 \times D_2 \times \dots \times D_n = \{(d_1, d_2, \dots, d_n) \mid d_i \in D_i, i = 1, 2, \dots, n\}$$

其中 $(d_1, d_2, \dots, d_n)$ 称为 n 元组(n -tuple), d_i 称为 n 元组的第 i 个分量。

由笛卡儿积的定义可以看出, n 个集合的笛卡儿积是以 n 元组为元素的集合, 而每一个 n 元组中的第 i 个分量取自于第 i 个集合 D_i 。

例 1.5 设有 3 个集合

$$A = \{a_1, a_2, a_3\}, \quad B = \{b_1, b_2\}, \quad C = \{c_1, c_2\}$$

则它们的笛卡儿积为

$$\begin{aligned} A \times B \times C = & \{(a_1, b_1, c_1), (a_1, b_1, c_2), (a_1, b_2, c_1), (a_1, b_2, c_2), \\ & (a_2, b_1, c_1), (a_2, b_1, c_2), (a_2, b_2, c_1), (a_2, b_2, c_2), \\ & (a_3, b_1, c_1), (a_3, b_1, c_2), (a_3, b_2, c_1), (a_3, b_2, c_2)\} \end{aligned}$$

如果 n 个集合 $D_1, D_2, \dots, D_n$ 中的元素个数分别为 $m_1, m_2, \dots, m_n$, 则其笛卡儿积中共有 $m_1 \times m_2 \times \dots \times m_n$ 个 n 元组。即 n 个集合的笛卡儿积是所有 n 元组组成的集合。

1.1.4 二元关系

定义 1.4 设 M 和 N 是两个集合, 则其笛卡儿积

$$M \times N = \{(x, y) \mid x \in M \text{ 且 } y \in N\}$$

的每一个子集称为在 $M \times N$ 上的一个二元关系。

如果 $M=N$, 则其笛卡儿积

$$M \times M = \{(x, y) \mid x, y \in M\}$$

的每一个子集称为在集合 M 上的一个二元关系, 简称为在集合 M 上的一个关系。

例 1.6 设集合 M 为

$$M = \{a, b, c, d, e, f\}$$

则下列每一个二元组的集合是在集合 M 上的一个关系: