

21世纪高等学校规划教材 | 计算机科学与技术

国家精品课程“离散数学”教学用书

“计算机科学与技术”国家特色专业教学用书

国家级“电子信息类工程应用型人才培养模式创新实验区”教学用书

离散数学

Discrete Structure

古天龙 常亮 编著

清华大学出版社



21世纪高等学校规划教材 | 计算机科学与技术

离散数学

Discrete Structure

古天龙 常亮 编著

清华大学出版社

北京

内 容 简 介

离散数学又称离散结构,是研究离散对象的模型、性质及操作的一门学科,是现代数学的一个重要组成部分,是计算机科学与技术的理论基础。本书依据 ACM 和 IEEE-CS 发布的 CC2005 教程,以及教育部高等学校计算机科学与技术教学指导委员会制订的计算机科学与技术专业规范,着力使内容和知识体系的设计达到理论与实际结合、抽象与直观统一、局部与整体协调。全书共 9 章,主要内容包括集合、关系、函数、命题逻辑、谓词逻辑、半群和群、环和域、格和布尔代数、图、树等。

本书体系严谨、结构新颖、内容翔实,可作为高等院校计算机及相关专业本科生、研究生“离散数学”课程的教材,也可作为从事计算机及相关领域研究和应用开发人员的参考用书。

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

离散数学/古天龙,常亮编著.--北京:清华大学出版社,2012.7

(21 世纪高等学校规划教材·计算机科学与技术)

ISBN 978-7-302-28820-6

I. ①离… II. ①古… ②常… III. ①离散数学—高等学校—教材 IV. ①O158

中国版本图书馆 CIP 数据核字(2012)第 101936 号

责任编辑:郑寅堃 张为民

封面设计:傅瑞学

责任校对:焦丽丽

责任印制:张雪娇

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:23.5

字 数:571 千字

版 次:2012 年 7 月第 1 版

印 次:2012 年 7 月第 1 次印刷

印 数:1~3000

定 价:35.00 元

产品编号:044080-01

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和教学方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人: 魏江江

E-mail: weijj@tup.tsinghua.edu.cn

前言

离散数学(discrete mathematics)又称离散结构(discrete structure),是以离散变量特征的对象为主要目标,研究其模型、性质及操作的一个数学分支。从人类社会历史的发展过程来看,18世纪以前的数学主要讨论整数、整数的比(有理数),甚至把几何图形也看作是由很多孤立的“原子”组成的。因而,那时的数学被看作是研究离散的或离散化了的数量关系的科学,基本上属于离散数学的范畴。此后,随着数学理论不断发展(不可通约线段的发现,对无限概念的深入探讨),加之天文学、物理学中遇到的物体运动等相关问题的需求推动,出现了连续数量的实数概念以及处理连续数量关系的微积分学。20世纪80年代后,随着计算机日益渗透到现代社会的各个方面,工业革命时代以微积分为代表的连续数学占主流的地位逐渐发生了变化,离散数学又重新受到高度的重视。

离散数学与计算机科学二者之间有着相辅相成的关系。一方面,离散数学是计算机科学与技术的理论基础,为计算机及其应用的诞生和发展提供了必要的理论支撑。例如,图灵(Alan Mathison Turing, 1912—1954)针对可计算所建立的图灵机是计算机的理论模型,这个理念导致了计算机的诞生;布尔(George Boole, 1800—1864)的逻辑代数是计算机硬件分析与设计的基础;谓词逻辑演算为人工智能学科提供了一种重要的知识表示和推理方法;代数系统的域为信息安全提供了一种重要的椭圆曲线密码体制,代数系统的格是计算机硬件和软件模型检验验证的理论基础。另一方面,数字电子计算机是一个离散结构,它只能处理离散的或离散化了的数量关系。随着计算机科学的迅猛发展,在计算技术、计算机软硬件和计算机应用等各个领域提出了许多有关离散变量为特征的理论问题,迫切需要适当的数学工具来描述和深化,从而,使得人们重新认识到离散变量对象的研究意义,重新重视讨论离散数量关系的数学分支,并取得新的发展。离散数学也因此作为一门学科应运而生,并成为现代数学的一个重要组成部分。

离散数学在计算机科学与技术及相关领域有着广泛的应用。它是计算机科学与技术及相关专业的一门核心基础课程,是许多其他后续课程,如数字电路、程序设计语言、数据结构、操作系统、编译原理、软件工程、人工智能、数据库系统、算法设计与分析、计算机网络、密码学、运筹学等必不可少的先行课程。通过离散数学的学习,不但可以掌握处理离散结构的描述工具和研究方法,为后续课程的学习创造条件,而且可以提高学生的抽象思维、逻辑推理和归纳构造能力,为将来参与创新性的研究和开发工作打下坚实的基础。

离散数学是数理逻辑、集合论、关系论、函数论、组合学、数论、代数结构、图论等汇集起来的一门综合学科。它跨越了数学的诸多分支,并与整个计算机科学紧密联系,所涉及的概念和方法,采用的符号和工具都远远超出其他任一门学科,所以不少读者会对这门课程或多或少产生畏惧、厌倦情绪。鉴于上述情况,我们根据多年来对离散数学的研究与教学实践的经验总结,从学生的学习情况及相关专业人员的自学特点出发编写了本书。在编写过程中,我们以“够用”为主,重点突出,并配合丰富的例题进行讲解,以达到内容翔实而结构清晰,语

言严谨而通俗易懂。同时,在课程的内容和知识体系的设计上,依据 ACM 和 IEEE-CS 发布的 CC2005 教程,以及教育部高等学校计算机科学与技术教学指导委员会制订的计算机科学与技术专业规范,着力达到理论与实际结合、抽象与直观统一、局部与整体协调。全书共分四篇,第一篇为集合论,包括集合、关系、函数;第二篇为数理逻辑,包括命题逻辑、谓词逻辑;第三篇为抽象代数,包括代数系统的基本概念与性质、半群和群、环和域、格和布尔代数;第四篇为图论基础,包括图的基本概念、赋权图、欧拉图、哈密顿图、二部图、平面图、树等。

本书在撰写过程中参阅了国内外大量的离散数学书籍和相关文献资料,从中汲取了许多好的思想,摘取了不少有用的素材,对此向相关作者表示感谢。徐周波、孟瑜、危前进、陈光喜等参与了书稿的讨论及部分文字整理,在此一并表示感谢。

限于作者水平,书中错误和疏漏之处在所难免,恳请广大读者批评指正。

作 者

2012 年 4 月

目录

第1篇 集合论

第1章 集合	3
1.1 集合的概念及表示	3
1.1.1 基本概念	3
1.1.2 集合的表示	4
1.2 特殊集合	6
1.2.1 子集合	6
1.2.2 幂集合	7
1.2.3 补集合	7
1.3 集合的运算	8
1.3.1 基本运算	8
1.3.2 运算的性质	9
1.4 计数问题	11
1.4.1 基本计数原理	11
1.4.2 排列与组合	12
1.4.3 容斥原理	13
1.5 集合的应用	17
习题	20
第2章 关系	24
2.1 关系的概念及表示	24
2.1.1 序偶与笛卡儿积	24
2.1.2 关系的定义	26
2.1.3 关系的表示	28
2.2 关系的性质	30
2.2.1 性质的定义	30
2.2.2 性质的判别	33
2.3 关系的运算	35
2.3.1 基本运算	35
2.3.2 复合运算	36
2.3.3 逆运算	39

2.3.4	幂运算	41
2.3.5	闭包运算	44
2.3.6	关系性质的运算封闭性	49
2.4	特殊关系	51
2.4.1	等价关系	51
2.4.2	相容关系	56
2.4.3	偏序关系	58
2.5	关系的应用	65
	习题	68
第3章	函数	75
3.1	函数的概念	75
3.1.1	函数的定义	75
3.1.2	特殊函数	78
3.2	函数的运算	80
3.2.1	复合运算	80
3.2.2	逆运算	82
3.3	函数的应用	83
	习题	86

第2篇 数理逻辑

第4章	命题逻辑	91
4.1	命题逻辑的基本概念	91
4.1.1	命题	91
4.1.2	联结词	93
4.2	命题逻辑公式	98
4.2.1	命题公式及其解释	98
4.2.2	命题公式的分类	104
4.2.3	命题公式的等值式	107
4.2.4	命题公式的范式	111
4.3	命题逻辑推理	122
4.3.1	推理的基本概念	122
4.3.2	简单证明推理	125
4.3.3	构造证明推理	129
4.4	命题逻辑的应用	135
	习题	140

第 5 章 谓词逻辑	145
5.1 谓词逻辑的基本概念	145
5.1.1 个体词	145
5.1.2 谓词	146
5.1.3 函词	147
5.1.4 量词	148
5.2 谓词逻辑公式	149
5.2.1 谓词公式及其解释	149
5.2.2 谓词公式的分类	156
5.2.3 谓词公式的等值式	158
5.2.4 谓词公式的范式	164
5.3 谓词逻辑推理	168
5.4 谓词逻辑的应用	176
习题	179

第 3 篇 抽象代数

第 6 章 代数系统	189
6.1 代数系统的基本概念	189
6.1.1 代数运算	189
6.1.2 代数系统	191
6.2 代数运算的性质	192
6.2.1 基本性质	192
6.2.2 特殊元素	198
6.3 相互联系的代数系统	202
6.3.1 同构代数系统	202
6.3.2 同态代数系统	206
6.3.3 商代数系统	210
6.4 代数系统的应用	215
习题	217
第 7 章 典型代数系统	221
7.1 半群和群	221
7.1.1 半群	221
7.1.2 群	227
7.1.3 特殊群	244
7.1.4 群的应用	250
7.2 环和域	254

7.2.1	环	254
7.2.2	域	259
7.2.3	域的应用	261
7.3	格和布尔代数	265
7.3.1	格	265
7.3.2	特殊格	271
7.3.3	布尔代数	274
7.3.4	格的应用	280
	习题	284

第4篇 图论基础

第8章	图	291
8.1	图的概念与表示	291
8.1.1	基本概念	291
8.1.2	图的连通性	298
8.1.3	图的操作	302
8.1.4	图的表示	303
8.2	赋权图	307
8.2.1	赋权图的定义	307
8.2.2	最短通路问题	308
8.3	欧拉图	309
8.3.1	欧拉图的定义	309
8.3.2	欧拉图的判定	311
8.3.3	中国邮路问题	313
8.4	哈密顿图	315
8.4.1	哈密顿图的定义	315
8.4.2	哈密顿图的判定	316
8.4.3	货郎担问题	319
8.5	二部图	321
8.5.1	二部图的定义	321
8.5.2	二部图的判定	322
8.5.3	匹配问题	322
8.6	平面图	326
8.6.1	平面图的定义	326
8.6.2	平面图的判定	328
8.6.3	图的着色问题	330
	习题	334

第 9 章 树	340
9.1 无向树	340
9.1.1 基本概念	340
9.1.2 生成树	343
9.1.3 最小生成树问题	345
9.2 有向树	348
9.2.1 基本概念	348
9.2.2 根树	349
9.2.3 二叉树	353
9.2.4 最优树问题	354
习题	360
参考文献	362

集合论是德国著名数学家康托尔(Georg Cantor, 1845—1918)于19世纪末创立的。17世纪,数学中出现了一门新的分支:微积分。在之后的一二百年中,这个学科获得了飞速发展并结出了丰硕成果。19世纪初,微积分相关的许多迫切问题得到解决后,出现了一场重建数学基础的运动。在这场运动中,康托尔意识到:微积分本质上是一种“无限数学”,以注数学基础中的许多问题都与无限集合有关,那么无限集合的本质是什么?它是否具备有限集合所具有的性质?基于此,康托尔探讨了前人从未碰过的实数点集,开辟了集合论研究的先河。

康托尔对集合(set)所下的定义是:把若干确定的有区别的(不论是具体的或抽象的)事物合并起来,看作一个整体,就称为一个集合,其中各事物称为该集合的元素。人们把康托尔于1873年12月7日给戴德金(Julius W. R. Dedekind, 1831—1916)的信中最早提出集合论思想的那一天定为集合论的诞生日。经历了20多年后,到20世纪初集合论才得到数学家们的赞同,并最终获得了世界公认。康托尔的集合论的建立,不仅是数学发展史上一座高耸的里程碑,甚至还是人类思维发展史上的一座里程碑。它标志着人类经过几千年的努力,终于基本上弄清了无限的性质,找到了制服无限“妖怪”的法宝。苏联著名数学家柯尔莫戈洛夫(Andrey Nikolaevich Kolmogorov, 1903—1987)指出:“康托尔的不朽功绩在于向无限冒险迈进”。德国数学大师希尔伯特(David Hilbert, 1862—1943)赞扬康托尔的理论是“数学思想最惊人的产物,在纯粹理性的范畴中人类活动最美的表现之一”。

1900年前后,数学家们普遍乐观地认为借助集合论的概念,从算术公理系统出发,便可以建造起整个数学的大厦。在1900年第二次国际数学大会上,著名数学家庞加莱(Jules Henri Poincaré, 1854—1912)就曾兴高采烈地宣布:“……数学已被算术化了。今天,我们可以说绝对的严格已经达到了”。正当康托尔的思想逐渐被人接受,并成功地把集合论应用到许多其他数学领域,大家认为数学的“绝对严格性”有了保证的时候,一系列完全没有想到的逻辑矛盾,在集合论的边缘被发现了。1903年英国哲学家兼数学家罗素(Bertrand A. W. Russell, 1872—1970)提出了一个著名的罗素悖论(paradox):“一个以具有某一特征的个体 r 为元素的集合 R ,集合 R 中的元素由非个体自身 r 的所有个体组成。现在问个体 r 是否属于集合 R ?如果个体 r 属于集合 R ,则集合 R 应满足 R 的定义,因此个体 r 不应属于 R ;另一方面,如果个体 r 不属于集合 R ,则集合 R 不满足 R 的定义,因此个体 r 应属于集合 R 。这样,不论何种情况都存在着矛盾。”

罗素悖论可以通过如下理发师悖论得到直观理解。在某个城市中有一位理发师,他的广告词是这样写的:“本人的理发技艺十分高超,誉满全城。我将为本城所有不给自己刮脸的人刮脸,我也只给这些人刮脸。我对各位表示热诚欢迎!”来找他刮脸的人络绎不绝,自然都是那些不给自己刮脸的人。可是,有一天,这位理发师从镜子里看见自己的胡子长长了,他本能地抓起了剃刀,你们看他能不能给他自己刮脸呢?如果他不给自己刮脸,他就属于“不给自己刮脸的人”,他就要给自己刮脸;而如果他给自己刮脸,他又属于“给自己刮脸的人”,他就不该给自己刮脸。如果每个人对应一个集合,这个集合的元素被定义成这个人刮脸的对象。那么,理发师宣称:他的元素,都是城里不属于自身的那些集合,并且城里所有不属于自身的集合都属于他。那

么他是否属于他自己?这样就由理发师悖论得到了罗素悖论。反过来的变换也是成立的。

罗素悖论一提出就在当时的数学界与逻辑学界引起了极大震动。不久,集合论存在漏洞的消息迅速传遍了数学界。著名逻辑学家弗雷格(Ludwig Gottlob Frege, 1848—1925)在他的关于算术的基础理论完稿付印时,收到了罗素关于这一悖论的信。他立刻发现,自己忙了很久得出的一系列结果却被这条悖论搅得一团糟。他只能在自己的著作《算术基础法则》的末尾写道:“一个科学家所碰到的最倒霉的事,莫过于是在他的工作即将完成时却发现所做的工作的基础崩溃了。”罗素悖论仅涉及“集合”与“属于”两个最基本的概念,它是如此简单明了,以至于根本没有留下为集合论漏洞辩解的余地。罗素悖论发表之后,接着又发现一系列悖论(后来归入所谓语义悖论):理查德悖论、培里悖论、格瑞林和纳尔逊悖论等。绝对严密的数学陷入了自相矛盾之中。这就是数学史上著名的第三次数学危机。

危机产生后,众多数学家纷纷提出了自己的解决方案,希望能够对康托尔的集合论进行改造,并通过对集合定义加以限制来排除悖论,为此需要建立新的原则:“这些原则必须足够狭窄,以保证排除一切矛盾;另一方面又必须充分广阔,使康托尔集合论中一切有价值的内容得以保存下来”。在这一原则基础上,策梅洛(Ernst Ferdinand Zermelo, 1871—1953)于1908年提出了第一个公理化集合论体系。弗兰克尔(Abraham Adolf Fraenkel, 1891—1965)将该公理系统改进成为无矛盾的集合论公理系统,简称为ZF公理系统(Zermelo-Fraenkel Axioms)。公理化集合系统的建立,使得原本直观的集合概念建立在严格的公理基础之上,成功排除了集合论中出现的悖论,形成了集合论发展的第二个阶段:公理化集合论。与此相对应,在1908年以前由康托尔创立的集合论被称为朴素(古典)集合论。公理化集合论是对朴素集合论的严格处理。它保留了朴素集合论的有价值的成果,并消除了其可能存在的悖论,从而比较圆满地解决了第三次数学危机。

关系(relation)理论最早出现在豪斯多夫(Felix Hausdorff, 1868—1942)于1914年出版的《集论》的序型理论中。等价关系概念的发展历程难以追溯。然而,这个概念的核心思想可以在拉格朗日(Joseph Louis Lagrange, 1735—1813)和高斯(Carl Gauss, 1777—1855)的工作中发现,他们定义了整数集上的同余关系。类似的思想在皮亚诺(Giuseppe Peano, 1858—1932)1889年的著作《几何原理》中也出现过。

莱布尼茨(Gottfried Wilhelm von Leibniz, 1646—1716)于1692年首先使用了函数(function)这个词,他用函数表示那些用来描述曲线的代数关系的量。1748年,欧拉(Leonhard Euler, 1707—1783)在他的《无穷小分析引论》中写到:“一个变量的函数是一个解析表达式,它由该变量以任何方式组成……”。我们正是从欧拉和克莱罗(Alexis Claude Clairaut, 1713—1765)的著作中,继承了 $f(x)$ 的记号,并一直沿用至今。1837年,狄利克雷(Peter Gaustav Lejeune Dirichlet, 1805—1859)对变量、函数和 $f(x)$ 中自变量与因变量之间的对应关系等给出了更为严谨的表述。这个定义不依赖于某个代数关系,而是考虑到更抽象的关系,以定义实体间的联系。基于狄利克雷的工作,布尔巴基(Nicolas Bourbaki)数学小组于20世纪30年代在现代集合论中把函数定义为笛卡儿(Rene Descartes, 1596—1650)积的子集。

20世纪以来的研究表明,不仅微积分的基础——实数理论奠定在集合论的基础上,而且各种复杂的数学概念都可以用“集合”概念定义出来,而各种数学理论又都可以“嵌入”集合论之内。因此,集合论就成了全部数学的基础,而且有力地促进了各个数学分支的发展。现代数学几乎所有的分支都会用到集合这个概念。

计算机科学及其应用的研究与集合论有着极其密切的关系。集合不仅可以用来表示数及其运算,更可以用于非数值信息的表示和处理,如数据的增加、删除、修改、排序,以及数据间关系的描述。有些很难用传统的数值计算来处理的问题,可以用集合运算得到方便处理。集合论在程序语言、数据结构、编译原理、数据库与知识库、形式语言和人工智能等领域中都得到了广泛的应用,并且还得到了发展,如扎德(Lotfi A. Zadeh, 1921—)的模糊集(fuzzy set)理论和帕夫拉克(Zdzislaw Pawlak, 1926—2006)的粗糙集(rough set)理论等。

1.1 集合的概念及表示

1.1.1 基本概念

集合作为数学中的基本概念,如同几何中的点、线、面等概念一样,是不能用其他概念精确定义的原始概念。集合是什么呢?下面是由康托尔首先给出的经典定义。

定义 1.1 集合就是由人们直观上或思想上能够明确区分的一些对象所构成的一个整体。集合里含有的对象或客体称为集合的**元素**(elements)或**成员**(members),集合是指总体,而元素是指组成总体的个体。

在日常生活和科学实践中经常会遇到各种各样的集合。例如,下列这些都可构成集合:

- ① 计算机学院的全体学生;
- ② 教室中的课桌;
- ③ 所有门电路;
- ④ 程序语言 Pascal 的全部数据类型;
- ⑤ 一个人的思想观点;
- ⑥ 张三同学所有选修的课程;
- ⑦ 计算机键盘上的所有符号;
- ⑧ 一个汉字的所有笔画;
- ⑨ 坐标平面上所有的点;
- ⑩ 离散数学课程中的所有概念。

可以看到,集合元素所表示的事物可以是具体的,如学生、课桌等,也可以是抽象的,如概念、观点、数据类型等。集合的元素可以是任意的,例如,一个汉堡、一张桌子、一个字母、一双鞋子、离散数学及漓江可以组成一个集合,尽管这样的集合可能没有人关心,但是的确符合集合的概念,是可以接受的。

集合的元素又必须是确定的和可区分的。例如,“授课的中年教师”这样的客体就不能组成集合,这是因为“中年”是一个界定不清的概念,教师到底是什么年纪才算中年呢?这个概念并不能明确界定集合的元素,所以不能构成集合。

一般用大写字母 $A, B, C, \dots$ 表示集合,用小写字母 $a, b, c, \dots$ 表示集合的元素。

定义 1.2 一个集合中的元素个数称为集合的**基数**(cardinality)。集合 A 的基数用

$|A|$ 或 $\text{card}(A)$ 表示。

例如,一个汉堡、一张桌子、一个字母、一双鞋子、离散数学及漓江组成的集合的元素个数是 6,所以,这个集合的基数就是 6。

定义 1.3 如果组成一个集合的元素个数是有限的,则称该集合为**有限集合**(finite set),简称为**有限集**,否则称为**无限集合**(infinite set),简称为**无限集**。

例如,英语字母组成的集合就是有限集,实数组成的集合就是无限集。

1.1.2 集合的表示

表示一个集合的主要方法有如下 3 种。

1. 枚举法

枚举法列出集合的全体元素,元素之间用逗号隔开括起来。例如: $A = \{1, 2, 3, 4, 5\}$, $B = \{3, 2, 1, 4, 5\}$, $C = \{a, b, c, d\}$, $D = \{a, c, d, d, b, a, c\}$ 。

一般来说,各元素出现的先后次序并不重要,所以集合 A 和 B 相同,集合 A 和 B 的基数分别为 $\text{card}(A)$ 和 $\text{card}(B)$,有 $\text{card}(A) = \text{card}(B) = 5$ 。一般地,集合中的元素可以以任意先后次序出现,称为集合元素的**无次序性**。

集合中元素如果有重复(即多次出现),是没有意义的,所以,集合 C 和 D 是相同的,即, $\text{card}(C) = \text{card}(D) = 4$ 。一般地,如果没有特别说明,集合中元素的重复出现和单次出现表示的含义相同,这个称为集合元素的**互异性**。

定义 1.4 如果集合 P 和 Q 由完全相同的元素组成,则称这两个集合**相等**,记为 $P = Q$ 。否则,称这两个集合**不相等**,记为 $P \neq Q$ 。

例如,上面的集合 C 和 D 就是两个相等的集合,即 $C = D$; 集合 A 和 C 就是两个不相等的集合,即 $A \neq C$ 。

当集合的元素数目较少,可以将所有的元素都列举出来。但是,当集合中的元素个数很多,甚至集合是无限集时,就不可能将集合的元素一一列举出来,这时可以只列出部分元素,而没有列出的元素则由已有的元素及其前后关系确定。但是,要注意列出的元素要足够多,否则会难以根据前后关系确定其他元素。从计算机角度看,枚举法是一种“静态”表示法,如果同时将所有的“数据”都输入到计算机中去,将占据大量的“内存”。

例 1.1 下面是枚举法给出的集合的例子。

- ① $A = \{1, 3, 5, 7, \dots\}$;
- ② $B = \{2, 4, 6, 8, \dots, 100\}$;
- ③ $P = \{a+1, a+2, a+3, \dots, a+999\}$;
- ④ $Q = \{a, A, b, B, c, C, \dots, Z\}$ 。

解释 ① 集合 A 由所有正奇数组成,是一个无限集;

② 集合 B 由 2 到 100 之间的 50 个偶数组成,是一个有限集,集合的基数为 $\text{card}(B) = 50$;

③ 集合 P 由 $a+1$ 到 $a+999$ 的表达式组成,是一个有限集,集合的基数为 $\text{card}(P) = 999$;

④ 集合 Q 由大、小写英文字母组成,是一个有限集,集合的基数为 $\text{card}(Q) = 52$ 。

2. 描述法

描述法通过刻画集合中元素所具备的某种特性来表示集合,通常用符号 $P(x)$ 表示不同对象 x 所具有的性质或属性 P ,又称为属性表示法。可表示为 $A=\{x|P(x)\}$,即集合 A 是由满足特性 P 的全体 x 组成。

例 1.2 下面是描述法给出的集合的例子。

① $A=\{x|x \text{ 是“discrete structure”中的所有英文字母}\}$;

② $B=\{x|x \text{ 是偶数,且 } x \geq 100\}$;

③ $P=\{x|x \text{ 是整数,且 } x^2+1=0\}$;

④ $Q=\{x|x \text{ 是计算机科学与技术专业的本科必修课程}\}$ 。

解释 ① A 由“discrete structure”中的英文字母 d,i,s,c,r,e,t,e,s,t,r,u,c,t,u,r 和 e 组成,但根据集合元素的互异性,不同字母为 d,i,s,c,r,e,t 和 u,所以, $A=\{d,i,s,c,r,e,t,u\}$,它是一个有限集合,集合的基数为 $\text{card}(A)=8$;

② B 由大于等于 100 的偶数组成,是一个无限集;

③ 没有任何整数满足 $x^2+1=0$,所以 P 中没有元素,集合的基数为 $\text{card}(P)=0$;

④ Q 由“高等数学”、“大学物理”、“数字逻辑”、“离散数学”、“编译原理”、“操作系统”、“软件工程”、“计算机组成原理”等组成,是一个有限集。

描述法的特点是便于对具有复杂特性的集合尤其是由无穷多个元素组成的无限集合进行表示。从计算机角度看,描述法是一种“动态”表示法,计算机在处理“数据”时,不用占据大量的“内存”。

注意: 对于给定的属性 P 和任意的元素 x , x 或者满足 P 或者不满足 P 。描述法是通过限定对象或客体是否在某一集体里来表示集合。一个给定的对象或客体是否在某一集合里,这是集合论中的一个基本问题。元素与集合的关系是集合的一类基本关系,称为属于关系或成员关系。

定义 1.5 一个对象 a 是集合 A 的元素,记为 $a \in A$,读作“ a 属于集合 A ”;如果一个对象 a 不是集合 A 的元素,记为 $a \notin A$,读作“ a 不属于集合 A ”。

对于给定集合 A 和元素 a ,或者 $a \in A$,或者 $a \notin A$,二者必居其一并且仅居其一,称为集合元素的确定性。

例如,对于例 1.1 中集合 A 和 B ,9 是集合 A 中的元素,所以 9 属于集合 A ,记为 $9 \in A$;9 不是集合 B 中的元素,所以 9 不属于集合 B ,记为 $9 \notin B$ 。

再如,对于例 1.2 中集合 A ,d 是集合 A 中的元素,所以 d 属于集合 A ,记为 $d \in A$;a 不是集合 A 中的元素,所以 a 不属于集合 A ,记为 $a \notin A$ 。

3. 图形法

图形法利用平面上点的对应元素的封闭区域对集合进行图解表示,一般地,用平面上的方形或圆形表示一个集合,又称为文氏图(Venn Diagrams)法。图 1.1 就是集合 A, B, C 和 D 的图形表示。

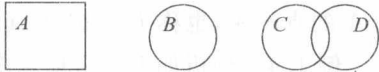


图 1.1 集合的图形表示