应用型本科信息大类专业“十三五”规划教材

基于计算思维的

离散数学基础

秦明 主编

提供 课件PPT 教案 www.ibook4us.com

提供 习题/试题(含答案) www.ibook4us.com

提供 课程标准 教学计划 www.ibook4us.com

华中科技大学出版社
http://www.hustp.com

应用型本科信息大类专业"十三五"规划教材

基于计算思维的
离散数学基础

主　编　秦　明
副主编　王岭玲　郑立平　石守礼　孙丽云
参　编　冯　睿　万其君　蔡文中　侯青青
　　　　朱道专　陶鸿敏　陈建国　卢　龙
　　　　张兆民

华中科技大学出版社
http://www.hustp.com
中国·武汉

内容简介

离散数学是计算机科学的理论基础，是计算机专业的核心课程，对于培养学生的逻辑思维能力，尤其是计算思维能力起着至关重要的作用。

相对于传统类型的离散数学教材，本书的最大特点是将计算思维融入各部分，力图使读者不仅能理解和掌握这门课程的基本概念和基本原理，而且通过对全书的学习，掌握怎样通过计算思维分析来解决实际的应用问题。本书系统地介绍了离散数学四大部分的内容：集合论、抽象代数、图论和数理逻辑。全书共分为9章，主要包括集合、关系、函数，代数系统、群论、格与布尔代数，图论，命题逻辑、谓词逻辑。本书内容的安排具有内在的逻辑联系，并在每一章都给出了通过计算思维来分析和解决实际应用问题的经典实例，以便于学生更好地理解和掌握分析问题和解决问题的方法。

为了方便教学，本书还配有电子课件等教学资源包，任课教师和学生可以登录“我们爱读书”网(www.ibook4us.com)免费注册浏览，或者发邮件至 hustpeiit@163.com 免费索取。

本书可以作为高等院校计算机科学、智能科学、信息安全等相关专业的本科生教材，也可以作为从事计算机科学及相关专业领域的从业人员的计算机专业理论参考书。

图书在版编目(CIP)数据

离散数学基础/秦明主编. —武汉：华中科技大学出版社，2016.6
应用型本科信息大类专业“十三五”规划教材
ISBN 978-7-5680-1669-8

Ⅰ.①离… Ⅱ.①秦… 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(2016)第 073663 号

离散数学基础
Lisan Shuxue Jichu

秦 明 主编

策划编辑：康 序
责任编辑：康 序
封面设计：原色设计
责任校对：何 欢
责任监印：朱 玢
出版发行：华中科技大学出版社(中国·武汉)
武昌喻家山 邮编：430074 电话：(027)81321913
录 排：武汉正风天下文化发展有限公司
印 刷：武汉鑫昶文化有限公司
开 本：787mm×1092mm 1/16
印 张：18.25
字 数：500 千字
版 次：2016 年 6 月第 1 版第 1 次印刷
定 价：38.00 元

华中出版

本书若有印装质量问题，请向出版社营销中心调换
全国免费服务热线：400-6679-118 竭诚为您服务

前言 PREFACE

随着计算机科学与技术的飞速发展和不断完善，离散数学已经成为计算机科学以及信息安全等相关专业的专业基础理论的核心课程。然而，一方面，由于我国高等教育发展的现状是几乎每所高校开设的计算机专业对于这门课所选定的教材五花八门，这些教材讲解的侧重点也不尽相同，造成了对这门课究竟应该讲授哪些内容，应该怎样讲授等诸多方面的问题并未达成一致的观念，形成统一的认识；另一方面，计算机专业或与计算机相关专业（如信息安全等专业）的学生认为这门课程与自己所学计算机专业的联系并不十分密切，甚至有相当一部分学生把这门课程仅仅当成是一门数学课程来学习，但是与数学课程相比较而言，这门课程的内容显得更加复杂，这是因为这门课程的四个部分（集合论、抽象代数、图论和数理逻辑）中的任何一个部分单独抽取出来都可以作为一个独立的数学学科，以至于使他们感到学习这门课程不仅难度很大，而且枯燥乏味。因此，编者一直在思考，是否正是由于这两个方面的原因，导致了在全国研究生统一入学考试的初试阶段，计算机专业统考的课程中竟然没有离散数学这门课程的考试内容，这与离散数学这门课程在计算机科学与技术中所占有的专业基础理论的核心课程的地位是极不相称的。正因为如此，编者觉得有必要编写一本关于讲解和学习离散数学这门课程的教材，力图通过这本教材给国内各高校的计算机专业师生带来一些启发，起到抛砖引玉的作用。

作为一名长期在离散数学这门课程教学第一线的教师来说，编者深知这门课程对计算机专业的学生大学四年的学习来说意味着什么；作为一名多年从事高等教育的工作者来说，编者深知“授人以鱼不如授人以渔”的道理，特别是对于计算机专业的学生来说尤其如此。由于这个专业的知识更新的速度实在是太快，因此，在大学四年中仅仅只教给学生一些现成的计算机专业的理论与实践的知识是远远不够的，教给学生受用终生的“渔”，对于每个从事计算机专业的教师来说是责无旁贷的。那么，在计算机专业中的“渔”究竟指的是什么呢？在编者看来应该是计算思维（computational thinking）。“计算思维”这个概念是由美国卡内基梅隆大学的周以真教授于 2006 年 3 月在美国计算机权威期刊《Communications of the ACM》首次提出的。计算思维是指运用计算机科学的基础概念

进行问题求解、系统设计，以及人类行为理解（如何用计算机模拟人类行为）等涵盖计算机科学之广度的一系列思维活动。

因此，编者在2011年对该课程进行了教学改革，也就是将计算思维这种全新的思维方式融入整个离散数学课程的教学中来。为什么在计算机专业核心课程的教学过程中要融入计算思维呢？如果我们将计算机科学理论体系看成一座建筑师设计的大厦，那么计算思维无疑就是建造这座大厦所必需的一根柱子。事实上，通过这种教学改革实践，不仅使教师在讲授离散数学课程中能够切身体会到这门课程在计算机专业课程中所处的核心地位，而且使学生能够更加深入理解什么是计算思维，以及使学生逐步认识到计算思维对于计算机专业的学习具有哪些决定性的指导作用。通过三年的教学改革实践，取得了一系列丰硕的教学成果，不仅使学生理解了什么是计算思维，并且能够带着这样的一种思维方式（思维习惯）去学习计算机专业中的其他课程，进而能够通过计算思维分析和解决一些典型的实际应用问题，甚至可以去涉猎计算机科学中的一些前沿的领域（如大数据处理、量子计算等）。与此同时，我们还发表了数篇关于将计算思维融入计算机专业课程的教学改革实践的论文，获得了计算机科学界和计算机教育界的一致认可。

在进行教学改革的同时，我们也专门研究了传统的与离散数学这门课程相配套的国内教材，虽然这些教材将离散数学这门课程所涉及的范畴、内容和问题都一一介绍了，甚至有些教材中所举的例题数量非常多，但是我们认为这些教材的编者都忽视了一个重要的问题，就是这些教材所讲的内容、所举的例题似乎跟计算机专业没有密切地结合起来，以至于使学生在学习这门课程时至少产生两个疑虑：这门课程与计算机专业究竟有什么联系？为什么要学这门课程？因此，我们认为，除了在教学方法上要进行改革实践之外，还应在教学改革的另一个层面，即教材上也要狠下一番功夫。因此，为了把这三年以来教学改革的阶段性成果进行总结，我们编写了这本教材。与传统教材相比，本教材的最主要特点在于将计算思维融入各个章节中。也就是说，读者在阅读本书时，将会有一个比较直观的感受，即本书在介绍该课程的基本概念和基本原理的同时，还介绍了构成这些知识的来源。通过阅读此书，读者可以进一步领悟到离散数学中有许多概念和许多处理问题的方法是随着计算机科学的发展而发展起来的。从这个意义上来说，该课程是与计算机专业的发展有着密切联系的，这种密切的联系在本教材的每个章节中都通过具体的例子体现出来了。通过对这些例子的学习，读者能逐步理解并掌握怎样通过计算思维分析和解决实际应用问题。

全书由四大部分（共9章）组成，即集合论（第1章～第3章）、抽象代数（第4章～第6章）、图论（第7章）、数理逻辑（第8章、第9章）。这四大部分有些部分之间的联系比较密切（如集合论与抽象代数），而有些部分相对来说自成体系，比较独立（如数理逻辑）。教师在选择这本教材进行授课时需要注意这个特点。在每一章的开头都有内容提要，在这一部分内容中，除了介绍本章的主要内容之外，还特别指出了学习本章的方法，读者需要仔细领会。在每一章的最后，都配有经典例题选编和一定数量的习题。在本章经典例题选编中，将本章中所介绍

的基本概念通过对综合性较强的例题的解答，力图使读者对全章的精华有更为深入的理解。

本书可以作为高等院校计算机科学、智能科学、信息安全等相关专业的本科生教材，也可以作为从事计算机科学及其相关领域的计算机专业理论参考书。建议本书授课学时数为80～90学时，如果学时数不够，可以根据各个学校计算机专业发展的实际情况进行相应调整。

在本书编著的过程中得到了文华学院等相关院校与专家的大力支持和帮助；感谢华中科技大学出版社的全体编辑，正是因为有了他们的辛勤工作，才使得本书在较短的时间内完稿并出版；同时对在本书编写过程中给予编者默默支持和帮助的家人和朋友一并表示感谢。

为了方便教学，本书还配有电子课件等教学资源包，任课教师和学生可以登录“我们爱读书”网（www.ibook4us.com）免费注册浏览，或者发邮件至hustpeiit@163.com免费索取。

本书由文华学院秦明担任主编，由武汉工程科技学院王岭玲、哈尔滨远东理工学院郑立平、石家庄铁道大学四方学院石守礼、燕京理工学院孙丽云担任副主编。

在本书的编写过程中，有许多对基本概念和基本原理的全新的诠释、内容、观念和体会都是我们这三年来教学改革实践成果的总结，现在将这些成果展示出来，其主要目的是借此机会与从事计算机科学领域的专家、学者共同分享，同时也通过该教材的出版向更多的计算机爱好者传播和普及计算思维这种全新的教学思维模式。

欢迎广大读者与编者联系，E-mail地址：noahqm@126.com。

秦　明

2016年3月

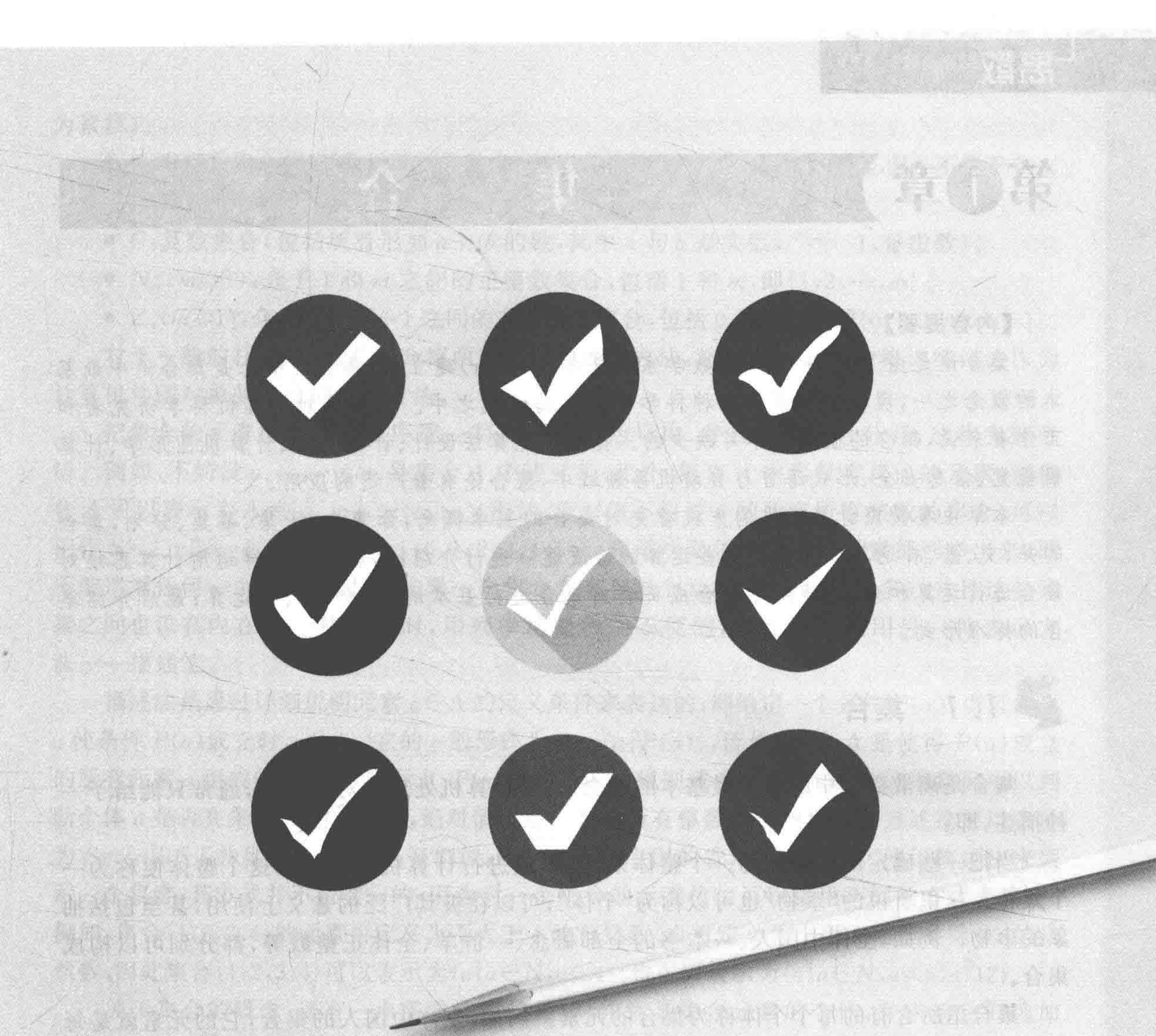
目录 CONTENTS

第1部分　集合论

第 2 部分 抽象代数

第 3 部分 图论

第4部分 数理逻辑



第1部分

Part 1 集合论

JIHELUN

A 为城里所有由该理发师理发的人的集合，稍加考虑就会明白，该理发师与集合 A 之间只可能存在两种关系：要么该理发师在集合 A 中，要么该理发师不在集合 A 中。二者必居其一。据此，我们可以进行下面的推理。若该理发师在集合 A 中，则表明该理发师给自己（理发师）理发，然而又根据由理发师理发的人的共性是他们都不给自己理发可知，该理发师不在集合 A 中；也就是说，由理发师在集合 A 中这个前提出发，通过正确的逻辑推理（演绎），最后得到了一个与前提相矛盾的结论，这只能说明前提假设不正确。于是，我们只能换一种前提，由此产生了第二种推理。若该理发师不在集合 A 中，则表明该理发师不给自己（理发师）理发。然而又根据“理发师为且只为城里所有不给自己理发的人理发”这个前提可知，该理发师在由城里所有不给自己理发的人组成的集合之中的。因此可以推得，该理发师在集合 A 中。也就是说，由理发师不在集合 A 中这个前提出发，通过正确的逻辑推理（演绎），最后也得到了一个与前提相矛盾的结论，这只能说明第二种前提假设不正确。而理发师与集合 A 的关系只有两种，即要么理发师在集合 A 中，要么理发师不在集合 A 中。二者必居其一。因此，在集合论中，某理发师跟且只跟城里所有不给自己理发的人理发是一个逻辑悖论。这个集合论中著名的悖论也是由英国的集数学家与逻辑学家于一身的罗素（1872—1970）提出的。通常将该悖论称为理发师悖论。通常有两种方法解决这种逻辑悖论。其一是将产生这种逻辑悖论的因素去掉，也就是将理发师从所讨论的集合 A 中剔除出去，于是就可以产生一个新的集合，这种集合的特点就是将自身排除在外的寻常集；其二是仍然保留产生这种逻辑悖论的因素，并将这种含有逻辑悖论的集合称为非寻常集。今后，为了更好地建立集合论这种公理化体系，我们只研究寻常集，而对于非寻常集就不再进行专门讨论了。为了使集合的运算更加方便，我们还需要引入空集这个基本的概念。

定义 1-1 不含有任何元素的集合，称为空集，记作∅。

空集从表面上看，虽然是一个很不自然的概念，但是却很有用。其用处体现在以下两个方面。

(1) 对于某些集合的证明问题，如果使用空集的概念，往往可以使得推理过程清晰、简捷。例如，一般来说，如果想要证明命题 $P(x)$ 对于论域中所有的 x 均不成立，则只需要证明 $\{x \mid P(x)\}$ 是空集即可。关于这一问题的研究，将在第 4 部分数理逻辑中进行详细讨论。

(2) 可用于计算机处理集合的基本运算上，使用空集的概念，通常可以使计算过程变得简捷明了。关于这部分的内容，我们在后面的相关章节中再进行深入的分析与讨论。

下面介绍几个与集合有关的概念。集合 A 中不同元素的数目，称为集合 A 的基数，通常用 $\# A$ 表示。当集合 A 拥有有限数目的不同元素，亦即 $\# A$ 有限时，称集合 A 为有限集，当集合 A 拥有的不同元素数目是无限个，亦即 $\# A$ 无限时，称集合 A 为无限集。前面提及的自然数集、非负整数集、整数集、素数集、有理数集、实数集和复数集都是无限集；而集合 N_m 与 Z_m 是有限集，因为 N_m 与 Z_m 的基数均为数 m，m 是有限的。由于计算机存储数据的机器字长是有限的，所以计算机处理的对象通常是有限集，无限集仅仅是在数学家的大脑中进行逻辑思维时的对象。关于集合的基数，后面还将会进行比较详细的讨论。

1.2 集合的包含和相等

集合的包含关系和集合的相等关系是两个集合之间最基本的两种关系。

定义 1-2 设有集合 A、B，如果集合 A 的每一个元素都是集合 B 的元素（即如果

为素数)。

- Q:有理数集合(有理数是可以表示成 i/j 形式的数,这里 i 和 j 都是整数,且 j 不等于 0)。
- R:实数集合(包括全体有理数和无理数)。
- C:复数集合(包括所有形如 $a+ib$ 的数,其中 a 与 b 是实数,$i^2=-1$,是虚数)。
- $N_m(m\geqslant 1)$:介于 1 和 m 之间的正整数集合,包括 1 和 m,即$\{1,2,\cdots,m\}$。
- $Z_m(m\geqslant 1)$:介于 0 和 $m-1$ 之间的非负整数集合,包括 0 和 $m-1$,即$\{0,1,2,\cdots,m-1\}$。

对于一般的朴素集合,首先介绍两种常用的表示方法,然后在 1.5 节中将详细讨论作为计算机处理对象的集合的表示方法。

把集合的元素按照任意顺序逐一写在一个花括号内,并且使用逗号分开,这称为列举法。例如,不妨设 $a_1,a_2,\cdots,a_n$ 是集合 A 中的元素,此外,集合 A 中再没有其他的元素,则集合 A 可以表示为 $A=\{a_1,a_2,\cdots,a_n\}$;又例如,绝对值不超过 5 的所有所有整数的集合,可以记作 $S=\{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$。列举法必须将所有的元素都列举出来,而不能遗漏任何一个元素。因此,如果一个集合含有有限个元素(元素数目很多),并且这些元素之间也没有内在的一些联系时,用列举法是极为不便的,因此可以采用另一种表示方法——描述法。

描述法是通过详细说明元素 $a\in A$ 的定义条件来表达的,即给定一个条件 $P(x)$,当且仅当 a 使条件 $P(a)$成立时,$a\in A$。它的一般形式为 $A=\{a|P(a)\}$,读作"集合 A 是使得 $P(a)$成立的所有元素 a 组成的集合"。实际上,$P(a)$描述了一种规则或者一个公式,它使得我们可以判断个体 a 是否在集合 A 中。例如,绝对值不超过 5 的所有整数组成的集合用描述法可以表示为 $S=\{a|a\in I$ 并且 $5\geqslant a\geqslant -5\}$;又例如,$B=\{a|a$ 是中国的省$\}$。不难看出,使用描述法来表示一个集合,其方式并不是唯一的,因为对一个集合的元素往往可以采用不同的方式来确定。例如,集合$\{1,2,3,4\}$的元素可定义为不大于 4 的自然数,也可定义为小于 6 而能整除 12 的自然数,因此集合$\{1,2,3,4\}$可以表示为$\{a|a\in N,a\leqslant 4\}$,也可以表示为$\{a|a\in N,a<6,a|12\}$。

关于集合的概念,还有一点需要注意的是,对作为某集合元素的个体,并未给它们施加一些限制。因此,常常需要研究一些集合,其元素本身也是集合。例如,集合 $A=\{78,\{1,2,4\},c,\{p,q\}\}$,集合 $B=\{\{1,2\},\{2,3\},\{3,1\}\}$。对于这种情形,重要的是将元素$\{a\}$与元素 a 区分开来。例如,集合$\{p,q\}$是集合 A 的元素,即$\{p,q\}\in A$,但是个体 p 是集合$\{p,q\}$的元素,即 $p\in\{p,q\}$,可是个体 p 却不是集合 A 的元素,即 $p\notin A$。

然而,对于"包罗一切的集合"或"由一切集合组成的集合"等类似的术语,应该避免使用,因为它们会导致集合论中的逻辑悖论(logic paradox)。所谓逻辑悖论,就是指一种导致逻辑矛盾的命题。在逻辑学上指无论从怎样的前提假设出发,通过正确的逻辑推理(演绎),都会得出与前提假设相互矛盾的命题的理论体系或命题。例如,著名的罗素悖论。将不包含自身作为元素的集合称为寻常集,而将包含自身作为元素的集合称为非寻常集。于是可知,一个集合要么是寻常集,要么是非寻常集,这二者必居其一,并且只居其一。不妨设集合 T 是由所有寻常集组成的集合,即 $T=\{A|A$ 是集合,并且 $A\notin A\}$。

现在考虑,集合 T 是寻常集还是非寻常集?如果 T 是寻常集,那么由集合 T 的定义,T 必包含自身为元素,因此,集合 T 是非寻常集。这与假设矛盾,故集合 T 不是寻常集,也就是说集合 T 是非寻常集。然而根据非寻常集的定义,就必须应有 $T\in T$,因此集合 T 中包含了一个非寻常集作为它的元素,这又与集合 T 的定义矛盾。这就是说,由于假定集合 T 的存在,无论 T 是寻常集或者非寻常集都将引出矛盾。

又例如,研究下述情况:某理发师为且只为城里所有不给自己理发的人理发。定义集合

A 为城里所有由该理发师理发的人的集合，稍加考虑就会明白，该理发师与集合 A 之间只可能存在两种关系：要么该理发师在集合 A 中，要么该理发师不在集合 A 中。二者必居其一。据此，我们可以进行下面的推理。若该理发师在集合 A 中，则表明该理发师给自己（理发师）理发，然而又根据由理发师理发的人的共性是他们都不给自己理发可知，该理发师不在集合 A 中；也就是说，由理发师在集合 A 中这个前提出发，通过正确的逻辑推理（演绎），最后得到了一个与前提相矛盾的结论，这只能说明前提假设不正确。于是，我们只能换一种前提，由此产生了第二种推理。若该理发师不在集合 A 中，则表明该理发师不给自己（理发师）理发。然而又根据“理发师为且只为城里所有不给自己理发的人理发”这个前提可知，该理发师在由城里所有不给自己理发的人组成的集合之中的。因此可以推得，该理发师在集合 A 中。也就是说，由理发师不在集合 A 中这个前提出发，通过正确的逻辑推理（演绎），最后也得到了一个与前提相矛盾的结论，这只能说明第二种前提假设不正确。而理发师与集合 A 的关系只有两种，即要么理发师在集合 A 中，要么理发师不在集合 A 中。二者必居其一。因此，在集合论中，某理发师跟且只跟城里所有不给自己理发的人理发是一个逻辑悖论。这个集合论中著名的悖论也是由英国的集数学家与逻辑学家于一身的罗素（1872—1970）提出的。通常将该悖论称为理发师悖论。通常有两种方法解决这种逻辑悖论。其一是将产生这种逻辑悖论的因素去掉，也就是将理发师从所讨论的集合 A 中剔除出去，于是就可以产生一个新的集合，这种集合的特点就是将自身排除在外的寻常集；其二是仍然保留产生这种逻辑悖论的因素，并将这种含有逻辑悖论的集合称为非寻常集。今后，为了更好地建立集合论这种公理化体系，我们只研究寻常集，而对于非寻常集就不再进行专门讨论了。为了使集合的运算更加方便，我们还需要引入空集这个基本的概念。

定义 1-1 不含有任何元素的集合，称为空集，记作 $\varnothing$。

空集从表面上看，虽然是一个很不自然的概念，但是却很有用。其用处体现在以下两个方面。

(1) 对于某些集合的证明问题，如果使用空集的概念，往往可以使得推理过程清晰、简捷。例如，一般来说，如果想要证明命题 $P(x)$ 对于论域中所有的 x 均不成立，则只需要证明 $\{x \mid P(x)\}$ 是空集即可。关于这一问题的研究，将在第 4 部分数理逻辑中进行详细讨论。

(2) 可用于计算机处理集合的基本运算上，使用空集的概念，通常可以使计算过程变得简捷明了。关于这部分的内容，我们在后面的相关章节中再进行深入的分析与讨论。

下面介绍几个与集合有关的概念。集合 A 中不同元素的数目，称为集合 A 的基数，通常用 $\# A$ 表示。当集合 A 拥有有限数目的不同元素，亦即 $\# A$ 有限时，称集合 A 为有限集，当集合 A 拥有的不同元素数目是无限个，亦即 $\# A$ 无限时，称集合 A 为无限集。前面提及的自然数集、非负整数集、整数集、素数集、有理数集、实数集和复数集都是无限集；而集合 N_m 与 Z_m 是有限集，因为 N_m 与 Z_m 的基数均为数 m，m 是有限的。由于计算机存储数据的机器字长是有限的，所以计算机处理的对象通常是有限集，无限集仅仅是在数学家的大脑中进行逻辑思维时的对象。关于集合的基数，后面还将会进行比较详细的讨论。

1.2 集合的包含和相等

集合的包含关系和集合的相等关系是两个集合之间最基本的两种关系。

定义 1-2 设有集合 A、B，如果集合 A 的每一个元素都是集合 B 的元素（即如果

$a \in A$，则必有 $a \in B$)，则称集合 A 是集合 B 的子集，或者称集合 A 包含于集合 B 中(或者集合 B 包含集合 A)，记作 $A \subseteq B$ 或者 $B \supseteq A$。

这个定义通常可以用来证明两个集合相等或者判断两个集合是否相等。下面，我们举几个例子进行说明。

例 1-1 设集合 $A=\{a,c,d,e\}$，集合 $B=\{a,b,c,x,y\}$，集合 $C=\{a,b\}$，则有集合 C 是集合 B 的子集，但是集合 C 却不是集合 A 的子集。注意区别属于关系和包含关系。属于关系 $a \in A$ 是指集合 A 的元素 a 与集合 A 的关系，然而包含关系 $C \subseteq A$ 是指集合 A 与另一个集合 C 之间的关系。

例 1-2 设集合 $A=\{a,b,c,d\}$，则有 $a \in A$，并且 $\{a\}$ 是集合 A 的子集。由属于关系和包含关系的定义可知，并不排斥有 $A \in B$ 和 $A \subseteq B$ 同时成立的可能性。

例 1-3 设集合 $A=\{a,b,c\}$，集合 $B=\{\{a,b,c\},a,b,c\}$，则显然应有集合 A 是集合 B 的元素，同时有集合 A 是集合 B 的子集。

关于两个集合之间的包含关系，有以下几个重要的性质。

(1) 对于任意的集合 A，有 $\varnothing \subseteq A$。

(2) 对于任意的集合 A，有 $A \subseteq A$。

(3) 对于任意的集合 A,B,C，如果 $A \subseteq B$，$B \subseteq C$，那么有 $A \subseteq C$。

其中，性质(2)和性质(3)是显而易见的。下面仅证明性质(1)，现使用反证法证明，不妨设空集 $\varnothing$ 不是集合 A 的子集，即 $\varnothing \not\subseteq A$，则必存在元素 $x \in \varnothing$，可是 $x \notin A$，而这与空集的定义相矛盾，因此，有 $\varnothing \subseteq A$。

以上我们讨论了集合与集合之间的第一种关系——包含关系。下面讨论集合与集合之间的第二种关系——相等关系。

定义 1-3 设有集合 A 和集合 B，如果集合 A 中的每一个元素都是集合 B 的元素，反过来，集合 B 的每一个元素也都是集合 A 的元素，则称集合 A 与集合 B 相等，记作 $A=B$。

由定义 1-3 不难看出，所谓集合 A 与集合 B 相等，即意味着集合 A 与集合 B 具有完全相同的元素。由定义 1-2 和定义 1-3 可知，当且仅当 $A \subseteq B$ 并且 $B \subseteq A$ 时，有 $A=B$。这是用来证明两个集合相等的主要方法之一，即运用集合相等的定义证明两个集合相等。这是一种按照逻辑思维证明两个集合相等的方法。后面我们还将会讨论如何按照计算思维证明两个集合相等。

下面用一些例子来说明如何理解集合之间的相等与不相等。

例 1-4 令集合 $A=\{1,2,4\}$，集合 $B=\{1,2,2,4\}$，则根据定义 1-3，$A=B$。

因为集合 A 中的任意一个元素都在集合 B 中，同时，集合 B 中的每个元素都在集合 A 中，根据定义 1-3，则 $A=B$。这个例子说明，在集合的列举法表示法中，某个元素的符号是可以重复出现的，不会改变这个集合。通常将类似于集合 B 这样的集合称为多重集合，多重集合在计算机科学研究中有其特殊的意义和价值，特别是在讨论具有相同关键字记录的排序过程中需要用到多重集合的概念。关于多重集合的讨论，我们将在 1.9 节详细展开讨论。为了叙述方便起见，本书中如果没有做特别说明，主要讨论的集合仍然是朴素集合，并且这些集合中的元素是互不相同的。

例 1-5 令集合 $C=\{1,4,2\}$，集合 $D=\{2,1,4\}$，则根据定义 1-3 可知，$C=D$。

这个例子说明在集合的列举法表示法中，如果将集合中元素的次序任意改变，该集合不变。

例 1-6 设集合 $P=\{\{1,2\},4\}$，集合 $Q=\{1,2,4\}$，则 $P\neq Q$。

如果集合 $A=\{x|x(x-1)=0\}$，集合 $B=\{0,1\}$，则 $A=B$。这说明集合的表示方法不是唯一的。

定义 1-4 设有两个集合 A 与 B，如果 $A\subseteq B$ 并且 $A\neq B$，则称集合 A 是集合 B 的真子集，用 $A\subset B$ 表示。

例如，集合 $\{1,2,3,4,5\}$ 是集合 $\{x|x\in I,-5\leqslant x\leqslant 5\}$ 的真子集。由于空集是任意一个集合的子集，因此可以推导出定理 1-1。

定理 1-1 空集是唯一的。

证明 假设有两个空集合 $\varnothing_1$ 和 $\varnothing_2$，由于空集被包含于每一个集合中，因此有 $\varnothing_1\subseteq\varnothing_2$，同时有 $\varnothing_2\subseteq\varnothing_1$，根据集合相等的定义 1-3，可以得出 $\varnothing_1=\varnothing_2$。即空集是唯一的。

这个定理将在后面的相关集合证明的推理论证过程中起着很重要作用。

下面，我们将介绍在计算机科学中有着非常广泛应用的一类集合——幂集。

1.3 幂集

任意给定一个集合 A，我们知道空集和集合 A 都是集合 A 的子集，对于任何元素 $a\in A$，集合 $\{a\}$ 也是集合 A 的子集。类似地，还可以举出集合 A 的其他子集。下面讨论关于集合 A 的全部集合的子集的集合——幂集。幂集是在形式语言与自动机中需要使用的一个重要概念，也是计算机专业的后续课程（编译原理）中需要用到的重要概念。下面，我们给出幂集的定义。

定义 1-5 设有集合 A，由集合 A 的所有子集组成的集合，称为集合 A 的幂集，记作 2^A，即 $2^A=\{S|S\subseteq A\}$。之所以称其为幂集，主要是因为任何一个集合的幂集中元素的个数为 2 的整数次幂。

例如，设集合 $A=\{a\}$，则幂集 $2^A=\{\varnothing,\{a\}\}$，该集合中有 2 个元素；若集合 $B=\{a,b\}$，则幂集 $2^B=\{\varnothing,\{a\},\{b\},\{a,b\}\}$，该集合中有 4 个元素。若集合 $C=\{a,b,c\}$，则幂集 $2^C=\{\varnothing,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$，该集合中有 8 个元素。由于空集 $\varnothing$ 中没有元素，因此，空集 $\varnothing$ 的子集也是 $\varnothing$，即空集 $\varnothing$ 的幂集 $2^{\varnothing}=\{\varnothing\}$。通常将集合中元素的个数称为集合的基数。集合 A 的基数通常使用 $\#A$ 表示。根据集合的基数定义，不难得出以下的结论：任何一个集合的幂集一定是非空集合。换句话说，空集不可能作为幂集。这是幂集的重要性质之一。这条性质对于以后很多关于幂集问题的讨论至关重要。特别是对于形式语言与自动机中的有限状态自动机理论尤为重要。它是计算机专业课程“编译原理”的基本理论之一。

从以上的例子不难看出，当集合的基数增加时，集合的幂集的基数也随之增加。对于有限集，下面的定理给出了二者之间的关系。

定理 1-2 设集合 A 是具有基数 $\#A$ 的有限集，则 $\#(2^A)=2^{\#A}$。

证明 设集合 A 的基数 $\#A=n$，从 n 个元素中选取 i 个不同的元素的方法共有 C_n^i 种，这里，

$$C_n^i=\frac{n!}{i!\ (n-i)!}$$

所以集合 A 的不同子集的数目(包括空集$\varnothing$)为：

$$\#(2^A)=C_n^0+C_n^1+C_n^2+\cdots+C_n^n$$

由二项式定理可知：

$$(x+y)^n=C_n^0x^n+C_n^1x^{n-1}y+C_n^2x^{n-2}y^2+\cdots+C_n^ny^n$$

令 $x=y=1$,应有 $2^n=C_n^0+C_n^1+C_n^2+\cdots+C_n^n$,因此,$A$ 的幂集的基数 $\#(2^A)=2^n$,又由于集合 A 的基数 $\#A=n$,故有 $\#(2^A)=2^{\#A}$,证毕。

当集合 A 中的元素个数较多时,要毫无遗漏地列出集合 A 的所有子集是一件相当困难的事情。现在,我们引入一种表示法,按照这种表示法,能够毫无遗漏地列出一个有限集的每一个子集。为此,需要对所给集合的元素规定某种次序,使得某个元素可以被称为第一个元素,另一个元素称为第二个元素等(虽然在集合的定义中,并没有这样一种次序),即给每一个元素附加一个标号,以便描述这个元素相对于该集合其他元素的位置。这里又一次体现出了计算思维的思考方式,这是因为作为计算机处理对象的集合来说,它在计算机中通常是按照数组来存储的,而对于数组中的各个元素来说,它们具有一种前后顺序关系。实际上,这种表示法就是从计算机科学这个学科背景的角度出发考虑产生的。例如,在集合 $A=\{a,b,c\}$中,可以令元素 a 是第一个元素,元素 b 是第一个元素,元素 c 是第三个元素。在集合 A 的子集中,通常会有一些元素出现,而其余的元素则不出现。根据这一情况以及指定给集合中各个元素的顺序关系,就可以使用以下方式来表示集合 A 所有的子集。例如,集合 A 的各个子集可以表示为 $B_{000}=\varnothing,B_{001}=\{c\},B_{010}=\{b\},B_{011}=\{b,c\},B_{100}=\{a\},B_{101}=\{a,c\}$,$B_{110}=\{a,b\},B_{111}=\{a,b,c\}$,因此,$A$ 的幂集 $2^A=\{B_{000},B_{001},B_{010},B_{011},\cdots,B_{110},B_{111}\}$。其中,$B$ 的下标是一个三位的二进制数,每一位对应集合 A 中的一个元素,左边第一位是 1 还是 0 表示第一个元素 a 在集合 A 的子集中出现与否。类似地,第二位和第三位是 1 还是 0 分别表示第二个元素 b 和第三个元素 c 在集合 A 的子集中出现与否。于是,集合 A 的任意一个子集都可以使用 000～111 中的某一个下标来表示;反之,如果给出这 8 个(即 2^3 个)下标中的任何一个,就能够确定出相应的子集。

假设集合 $J=\{j\,|\,j$ 是二进制数,$000\leqslant j\leqslant 111\}$,则有 $2^A=\{B_j\,|\,j\in J\}$。可以看到,这里只用了下标来确定子集的各个元素,而表示这些子集时用到的字母 B 则是至关重要的。上面的表示法可以推广到一般的情形,用来表示具有任意 n 个不同元素的集合的各个子集。用来表示这些子集的下标是十进制数 0 到 2^n-1 的二进制表示。为了凑足 n 个数位,一定要在这些二进制数的左边插入所需数目的零。也可以使用从 0 到 2^n-1 的十进制数来作为子集的下标,则只在需要确定所对应子集的元素时才转换为二进制数。例如,令集合 $A_6=\{a_1,a_2,\cdots,a_6\}$,不难看出,集合 A_6 有 $2^6=64$ 个子集,可以称它们为 $B_0,B_1,B_2,\cdots,B_{63}$。下面我们讨论怎样确定集合 A_6 的任何子集的各个元素。

例如,
$$B_{24}=B_{011000}=\{a_2,a_3\}$$
$$B_{47}=B_{101\,111}=\{a_1,a_3,a_4,a_5,a_6\}$$

与集合的幂集相类似,我们还可以定义以下类型的集合,即这类集合的任何一个元素本身也是集合,这种类型的集合在后面还会经常遇到,这种集合称为集合族。例如,其和为 6 的不同正整数所构成的集合的集合$\{\{6\},\{1,5\},\{2,4\},\{1,2,3\}\}$就是一个集合族。按照这个定义,任意一个集合的幂集也是一个集合族。

通常用记号$\{A_i\}_{i\in K}$来表示所有集合 $A_i(i\in K)$所构成的集合族,也就是说：

$$\{A_i\}_{i\in K}=\{A_i\,|\,i\in K\}$$

其中,集合 K 是指标集。例如,集合族$\{A_0,A_1,A_2,A_3,A_4\}$可以表示为$\{A_i\}_{i\in K}$,在这

里，$K=\{0,1,2,3,4\}$。而当指标集 $K=\{i|\ i\in J, i_a\leqslant i\leqslant i_b\}$时，又可以将集合族$\{A_i\}_{i\in K}$表示为$\{A_i\}_{i=i_a}^{i_b}$。例如，集合族$\{A_0,A_1,A_2,A_3,A_4\}$也可以表示为$\{A_i\}_{i=0}^{4}$。

1.4 集合的运算

集合的运算不仅仅是集合论中关于集合的基本运算，也是计算机科学相关研究领域（如计算机图形学、人工智能等）中基本理论的出发点。在正式讨论集合的运算之前，还必须首先定义一个特殊的集合，它包含了所需要讨论的每一个集合。

定义 1-6 如果一个集合包含了某个问题中所讨论的一切集合，则称其为该问题的全域集合，或简称为全集合，记作 U。

根据上面的定义，全集合 U 并非是唯一的，为方便起见，通常总是取一个较为方便的集合作为全集合 U。例如，如果在实数范围内讨论问题，就可以将实数集 R 取作全集合 U，若在正整数范围内讨论问题，则可以将正整数集 N 取作全集合 U 的子集。全集合 U 应在问题讨论之初便被确定，以后在该讨论中涉及的任何一个集合均需要看成该全集合 U 的子集。因此，以后在讨论任何有关集合的运算或讨论集合的性质时，需要在全集合这个背景下进行。

本小节将讨论集合的几种运算。使用这些运算，通过对给定集合元素进行组合，就可以构成新的集合。

定义 1-7 设有集合 A、B，属于集合 A 或者属于集合 B 的所有元素组成的集合，称为集合 A 和集合 B 的并集，记作 $A\cup B$，即 $A\cup B=\{u|u\in A$ 或 $u\in B\}$。

定义 1-8 设有集合 A、B，属于集合 A 同时又属于集合 B 的所有元素组成的集合，称为集合 A 和集合 B 的交集，记作 $A\cap B$，即 $A\cap B=\{u|u\in A$ 且 $u\in B\}$。

下面，我们举几个具体例子来进行说明。

例 1-7 设集合 $A=\{a,b,c,d,e,f,g\}$，集合 $B=\{d,e,f,g,h,i\}$，则：

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

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

例 1-8 设全集 $U=N$，集合 $A=P$（素数集），集合 B 为正整数集 N 中所有奇数组成的集合，则 $A\cup B=\{$正奇数和 2$\}$；$A\cap B=\{$正奇素数$\}$。

特别地，如果集合 A 与集合 B 没有公共元素，即 $A\cap B=\varnothing$，就称集合 A 与集合 B 是不相交的。

例 1-9 设集合 $A_1=\{\{1,2\},\{3\}\}$，集合 $A_2=\{\{1\},\{2,3\}\}$，集合 $A_3=\{\{1,2,3\}\}$。则 $A_1\cap A_2=\varnothing$，$A_2\cap A_3=\varnothing$，$A_3\cap A_1=\varnothing$。因此，我们说集合 A_1、A_2、A_3 两两互不相交。

按照以上的定义，显然可以得到以下的关系式。

$$A\subseteq A\cup B,\quad B\subseteq A\cup B$$

$$A\cap B\subseteq A,\quad A\cap B\subseteq B$$

如果集合 A 是集合 B 的子集，则可以证明 $A\cup B=B$，$A\cap B=A$。读者可以自己证明。

定义 1-9 设有集合 A 和集合 B 两个集合，由属于集合 B 而不属于集合 A 的所有元素组成的集合，称为集合 A 关于集合 B 的相对补集，记作 $B-A$，即

$$B-A=\{u|u\in B\text{，且 }u\notin A\}$$