

21世纪高等学校规划教材 | 计算机科学与技术

离散数学 与算法化思维

程显毅 李医民 编著



清华大学出版社

21世纪高等学校规划教材 | 计算

离散数学 与算法化思维

程显毅 李医民 编著



清华大学出版社

内 容 简 介

本书从信息科学角度出发,系统全面地阐述了离散数学的基本原理和基本方法,并把算法化的思维贯穿于书中。全书共8章,包括集合论、数论、矩阵、关系、映射、函数、图论和数理逻辑。书中通过近40个算法实例,培养读者算法化思维的意识,且以节为单位布置习题,便于教学和自学。

本书可作为高等学校计算机专业、信息管理等相关专业离散数学课程教材,也可供从事信息科学的有关技术人员学习参考。

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

离散数学与算法化思维/程显毅,李医民编著. --北京:清华大学出版社,2013

21世纪高等学校规划教材·计算机科学与技术

ISBN 978-7-302-33530-6

I. ①离… II. ①程… ②李… 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(2013)第199377号

责任编辑:郑寅堃 王冰飞

封面设计:傅瑞学

责任校对:李建庄

责任印制:李红英

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:北京国马印刷厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:19.75

字 数:495千字

版 次:2013年12月第1版

印 次:2013年12月第1次印刷

印 数:1~2000

定 价:36.00元

产品编号:054387-01

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)是将传统的逻辑学、集合论、初等数论、组合分析、关系理论、函数与映射、图论、代数系统(群、环、域等)、计算模型(语言与自动机)等汇集起来的一门综合学科。离散数学的大部分内容早已出现在数学中,为什么还要将各个数学分支中的一些内容集中起来加以研究,并且冠以一个新的名称(离散数学)呢?这主要是因为计算机科学的产生和发展。

由于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,因此,无论计算机科学本身,还是与计算机科学及其应用密切相关的现代科学研究领域,都面临着如何对离散结构建立相应的数学模型,又如何将已用连续数量关系建立起来的数学模型离散化,从而可由计算机加以处理的问题。

离散数学各个主题的基本理论和基本方法大量地应用在数字电路、编译原理、数据结构、操作系统、数据库系统、算法的分析与设计、人工智能、计算机网络等信息科学专业课程中。离散数学不仅作为计算机科学与技术及相关专业的理论基础及核心主干课,还对后续课程提供必需的理论支持。更重要的是,离散数学旨在“通过对逻辑推理、组合分析、离散化、算法化思维等方面进行专门训练及实现相关应用,使学生养成严谨、完整、规范的科学态度。”

本教材根据作者多年的教学经验和较成熟的教案整理而成。本教材从计算机科学的角度的(而不是从数学的角度)讨论了离散数学主题,选择了最基本、最重要的内容,并努力做到简明扼要、深入浅出,既保持各主题的独立性,又展现出它们之间的密切联系。

本教材的主要特色如下所示:

(1) 算法是用计算机解决问题的方法与步骤,其核心是一种思维方式。本书通过近40个算法实例,培养读者算法化思维的意识,所有算法不涉及任何一种编程语言。

(2) 每章都使用图的方式对该章主题的知识体系进行总结,使学生认识到离散数学重要知识点之间的联系,便于学生从整体上把握众多的抽象概念。

(3) 在内容的组织上,考虑到本科生能接受的程度,突出了与计算机学科应用密切相关的内容(如数理逻辑、组合分析、图论、关系理论),删除或弱化了与计算机关联较弱的纯数学方面的内容(如图的着色、环、域等),使教材的结构更接近专业的实际需要。

(4) 以节为单位布置习题,便于教学和自学,每章都有自测题,便于了解学生对知识的综合掌握情况。习题和自测题参考答案没有安排在本教材内,全部习题答案的电子版资料仅供选用本教材的老师使用(可发E-mail至zhengyk@tup.tsinghua.edu.cn联系索取),以强迫学生进行独立思考。这样做能培养学生抽象思维、主动学习的习惯。

本书可作为计算机科学与技术及信息类本科专业的基础理论教材,也可供有关技术人员学习参考。

我们在编写本书的过程中参阅了许多国内外离散数学教材及专著,在此对相关作者表示感谢。

由于我们的水平和经验有限,书中的错误及不妥之处在所难免,恳请读者批评指正。

编 者

2013 年 10 月于南通大学

目 录

第 1 章 引论	1
1.1 离散化	1
1.1.1 为什么要离散化	1
1.1.2 计算机系统本质上是离散的	2
1.2 离散数学与计算机的关系	3
1.2.1 数学是计算机的基础	3
1.2.2 计算机对数学的贡献	5
1.2.3 离散数学的作用	6
1.2.4 离散数学在计算机学科主干课程中的应用	8
1.3 离散数学主题以及算法化思维	9
1.3.1 离散数学主题	9
1.3.2 算法化思维的重要性	11
1.4 如何学习离散数学	12
1.4.1 离散数学的特点	12
1.4.2 学习离散数学要注意的问题	12
1.5 本章小结	13
习题 1	15
第 2 章 基础知识	16
2.1 集合论	16
2.1.1 集合的基本概念	16
2.1.2 集合论的思想渊源	18
2.1.3 集合表示	21
2.1.4 集合运算及相关算法	23
2.1.5 集合证明技巧	30
习题 2.1	32
2.2 矩阵论	33
2.2.1 矩阵的概念及其基本运算	33
2.2.2 布尔矩阵及布尔积算法	37
习题 2.2	39
2.3 初等数论	39
2.3.1 数的整除性	39

2.3.2 同余	41
习题 2.3	41
2.4 本章小结	42
自测题 2	45

第 3 章 关系 47

3.1 序偶和笛卡儿积	47
习题 3.1	50
3.2 关系及其表示	50
3.2.1 关系的概念	50
3.2.2 几种特殊的关系	51
3.2.3 关系的表示	52
习题 3.2	54
3.3 关系的性质及其判定算法	54
3.3.1 关系的性质	54
3.3.2 关系性质判定算法	55
习题 3.3	57
3.4 复合关系	58
3.4.1 复合关系的定义	58
3.4.2 关系的复合运算的性质	59
3.4.3 复合关系的矩阵表示及图形表示	60
3.4.4 复合关系生成算法	61
习题 3.4	62
3.5 逆关系	63
3.5.1 逆关系的概念及性质	63
3.5.2 逆关系生成算法	65
习题 3.5	65
3.6 关系的闭包运算	66
3.6.1 关系传递闭包	66
3.6.2 关系传递闭包计算的 Warshall 算法	68
习题 3.6	70
3.7 等价关系	70
3.7.1 集合的划分和覆盖	70
3.7.2 等价关系与等价类	71
3.7.3 等价关系相关算法	74
习题 3.7	75
3.8 相容关系	76
习题 3.8	79
3.9 偏序关系	79

3.9.1 偏序关系的定义	79
3.9.2 哈斯图及其构造算法	80
3.9.3 偏序集中特殊位置的元素	81
3.9.4 拓扑排序算法	83
3.9.5 良序	83
习题 3.9	84
3.10 格	85
习题 3.10	91
3.11 关系在计算机科学中的应用	92
3.11.1 关系在关系数据库中的应用	92
3.11.2 关系传递闭包在语法分析中的应用	93
3.12 本章小结	95
自测题 3	97
第 4 章 映射	100
4.1 映射的基本概念	100
4.1.1 映射的概念	100
4.1.2 映射的分类	103
习题 4.1	105
4.2 复合映射和逆映射	106
4.2.1 复合映射	106
4.2.2 逆映射	108
习题 4.2	110
4.3 置换函数	111
习题 4.3	112
4.4 计算机科学中常用的函数	112
4.4.1 特征函数	112
4.4.2 取整函数	114
4.4.3 布尔函数	117
4.4.4 哈希函数	118
4.4.5 算法复杂性分析的数学基础	120
习题 4.4	123
4.5 本章小结	124
自测题 4	125
第 5 章 组合分析	126
5.1 计数	126
5.1.1 基本计数关系式	126
5.1.2 相容排斥原理	126

5.1.3	加法法则和乘法法则	129
习题 5.1		130
5.2	排列	131
5.2.1	无重复排列	131
5.2.2	有重复的排列	131
5.2.3	排列生成算法	132
习题 5.2		134
5.3	组合	135
5.3.1	无重复的组合	135
5.3.2	有重复的组合	136
5.3.3	组合生成算法	138
习题 5.3		139
5.4	生成函数	139
习题 5.4		141
5.5	鸽巢原理	142
5.5.1	一般的鸽巢原理	142
5.5.2	推广的鸽巢原理	142
习题 5.5		143
5.6	组合分析在计算机中的应用	144
5.7	本章小结	145
自测题 5		146
第 6 章	代数系统	148
6.1	代数系统发展史	148
6.2	运算与代数系统	150
6.2.1	运算的概念	150
6.2.2	代数系统的概念	151
6.2.3	运算的性质及性质判定算法	151
6.2.4	单位元、零元和逆元	153
习题 6.2		156
6.3	半群	157
6.3.1	半群及其性质	157
6.3.2	单位半群	159
习题 6.3		160
6.4	群	161
6.4.1	群的基本概念	161
6.4.2	群的性质	162
6.4.3	子群	163
习题 6.4		165

6.5 特殊的群	166
6.5.1 交换群	166
6.5.2 循环群	167
6.5.3 置换群	168
习题 6.5	169
6.6 代数系统的同态与同构	170
习题 6.6	175
6.7 代数系统在计算机中的应用	176
6.7.1 布尔代数及其在电路设计中的应用	176
6.7.2 群论在计算机编码纠错中的应用	177
6.7.3 半群与文法及形式语言	180
习题 6.7	181
6.8 本章小结	182
自测题 6	183
第 7 章 图论	186
7.1 图的基本概念	186
7.1.1 图的定义与分类	186
7.1.2 补图与生成子图	188
7.1.3 握手定理	189
7.1.4 同构图	190
习题 7.1	192
7.2 图的连通性	193
7.2.1 基本路与基本回路	193
7.2.2 图的连通性	195
习题 7.2	196
7.3 图的矩阵表示	197
7.3.1 邻接矩阵	197
7.3.2 可达性矩阵	199
7.3.3 完全关联矩阵	201
7.3.4 与图有关的算法	202
习题 7.3	203
7.4 欧拉图与哈密尔顿图	203
7.4.1 欧拉图	203
7.4.2 哈密尔顿图	206
习题 7.4	208
7.5 二部图及匹配	210
7.5.1 二部图	210
7.5.2 匹配及最大匹配算法	211

习题 7.5	214
7.6 平面图	215
7.6.1 平面图的定义	215
7.6.2 欧拉公式	217
习题 7.6	220
7.7 最短路 Dijkstra 算法	220
7.7.1 问题的提出	220
7.7.2 Dijkstra 算法	221
习题 7.7	221
7.8 树	222
7.8.1 树的基本概念	222
7.8.2 最小生成树 Kruskal 算法	224
7.8.3 二叉树	226
7.8.4 最优二叉树哈夫曼算法	229
习题 7.8	231
7.9 图论在计算机中的应用	232
7.9.1 二叉树的遍历算法及表达式表示	232
7.9.2 前缀码的设计	233
7.9.3 有限状态机的图表示	235
7.10 本章小结	236
自测题 7	240
第 8 章 数理逻辑	243
8.1 命题演算	243
8.1.1 命题与命题联结词	243
8.1.2 命题公式	248
8.1.3 真值表及其真值计算算法	250
8.1.4 命题等值式	252
8.1.5 重言式、矛盾式和可满足式	256
8.1.6 蕴涵式	257
习题 8.1	259
8.2 范式	261
8.2.1 对偶原理	261
8.2.2 范式转换算法	262
8.2.3 主范式形成算法	264
习题 8.2	272
8.3 命题推理	273
8.3.1 直接推理法	274
8.3.2 间接推理法	275

习题 8.3	276
8.4 谓词演算	278
8.4.1 个体、谓词和量词	278
8.4.2 谓词公式和辖域	281
8.4.3 谓词公式的解释及逻辑有效式	283
习题 8.4	284
8.5 谓词等值式和置换规则	286
习题 8.5	288
8.6 谓词推理	289
8.6.1 逻辑有效蕴涵式	289
8.6.2 推理定律	289
8.6.3 推理实例	291
习题 8.6	292
8.7 数理逻辑在计算机中的应用	293
8.7.1 逻辑推理在人工智能中的应用	293
8.7.2 人工智能语言 Prolog 简介	294
习题 8.7	295
8.8 本章小结	295
自测题 8	298
后记	301
参考文献	302

第1章

引论

计算机与数学的联系十分紧密。当今软件开发等信息界的精英,大部分是数学出身,数学在计算机中的应用是不言而喻的。

大部分高校的计算机系所开设的数学课程,无论广度、深度都达到相当的水准。从事计算机软件、硬件开发不仅需要高深的数学知识作为基础,而且需要很强的逻辑思维能力、形象思维能力和空间想象能力,这些离开数学是不可能实现的。

计算机科学就是算法的科学,而算法所处理的对象(数据)是离散的,所以离散对象的处理就成了计算机科学的核心,而研究离散对象的数学恰恰就是离散数学。

离散数学不仅在基础数学研究中具有极其重要的地位,在其他的学科中也有重要的应用,如在计算机科学、物理、化学、生物等学科中均有重要应用。本教材是从计算机科学角度,理解离散数学原理。如果说微积分的发展为工业革命奠定了基础,那么离散数学的发展则奠定了信息革命的基础。计算机之所以可以被称为电脑,就是因为有了离散结构算法,才使人感到计算机好像拥有思维一样。

1.1 离散化

1.1.1 为什么要离散化

离散化(或数字化)就是将许多复杂多变的信息转变为可以度量的数字,把它们转变为一系列二进制代码,进而用计算机进行处理,这也是对象离散化的目的。

离散化是程序设计中一个非常常用的技巧,它可以有效地降低时间复杂度。其基本思想就是在众多可能的情况中“只考虑关键的值”。离散化可以改进一个低效的算法,甚至实现根本不可能实现的算法。

为了说明离散化的思想,下面举一个简单的例子。

【例 1.1】 如图 1.1 所示,给定平面上 n 个点的坐标,求能够覆盖所有这些点的最小矩形面积。

解: 这个问题难就难在,这个矩形可以倾斜放置(边不必平行于坐标轴)。

这里的倾斜放置是很不好处理的,因为我们不知道这个矩形最终会倾斜多少度。假设我们知道这个矩形的倾角是 α ,那么答案就很简单了:矩形面积最小时,四条边一定都挨着某个点。也就是说,四条边的斜率已知的话,只需要让这些边从外面不断逼近这个点集直到

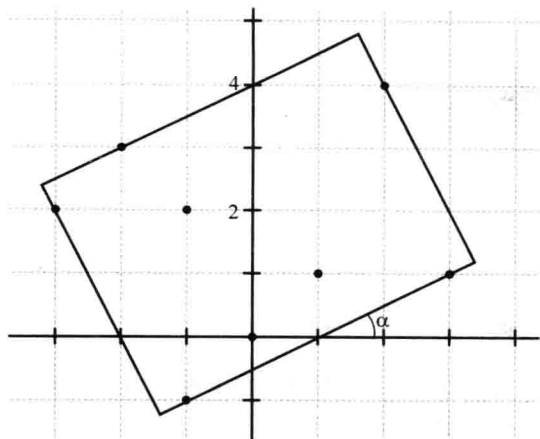


图 1.1 例 1.1 的图

碰到了某个点。你不必知道这个具体应该怎么实现,只需要理解这可以通过某种方法计算出来,毕竟我们的重点在下面的过程。

有了这样一个想法,算法就简化为:枚举矩形的倾角,对于每一个倾角计算出最小的矩形面积,最后取一个最小值。

这个算法是否正确呢?我们不能说它是否正确,因为它根本不可能实现。矩形的倾角是一个实数,它有无数种可能,你永远不可能枚举每一种情况。我们说,矩形的倾角是一个“连续的”变量,我们无法枚举这些倾角。我们需要一种方法,把这个“连续的”变量变成一个离散的离散值,这个过程也就是所谓的离散化。

可以证明,最小面积的矩形不但要求四条边上都有一个点,而且还要求至少一条边上有两个或两个以上的点。试想,如果每条边上都只有一个点,则我们总可以把这个矩形旋转一点使得这个矩形变“松”,从而有余地得到更小的矩形。于是我们发现,矩形的某条边的斜率必然与某两点的连线相同。如果我们计算出了所有过两点的直线的倾角,那么 α 的取值只有可能是这些倾角或它减去 90° 后的角(直线按“\”方向倾斜时),这样就有 $C(n, 2)$ 种情况。至此,我们说这个“倾角”已经被“离散化”了。虽然这个算法仍然有优化的余地,但此时我们已经达到了本文开头所说的目的。

【例 1.2】 计算机是如何处理图像的呢?

首先,要把连续的图像函数 $f(x, y)$, 在 z 轴(量化)和 XY 平面(采样)上数字化。

图 1.2(a)为原始图像,图 1.2(b)为对应图 1.2(a)的离散化结果,图 1.2(c)为图 1.2(b)的图像矩阵表示(局部)。

【例 1.3】 知识表示(符号化): $f(x)$ 在 x_0 点连续。

解: $\forall \varepsilon > 0, \exists \delta > 0, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$ 。

1.1.2 计算机系统本质上是离散的

计算机系统具有以下特点:

- (1) 计算机不能做什么——不能处理连续对象,当然还有其他不能做的事情。
- (2) 计算机体系结构——模块化。
- (3) 计算机工作原理——串行化。

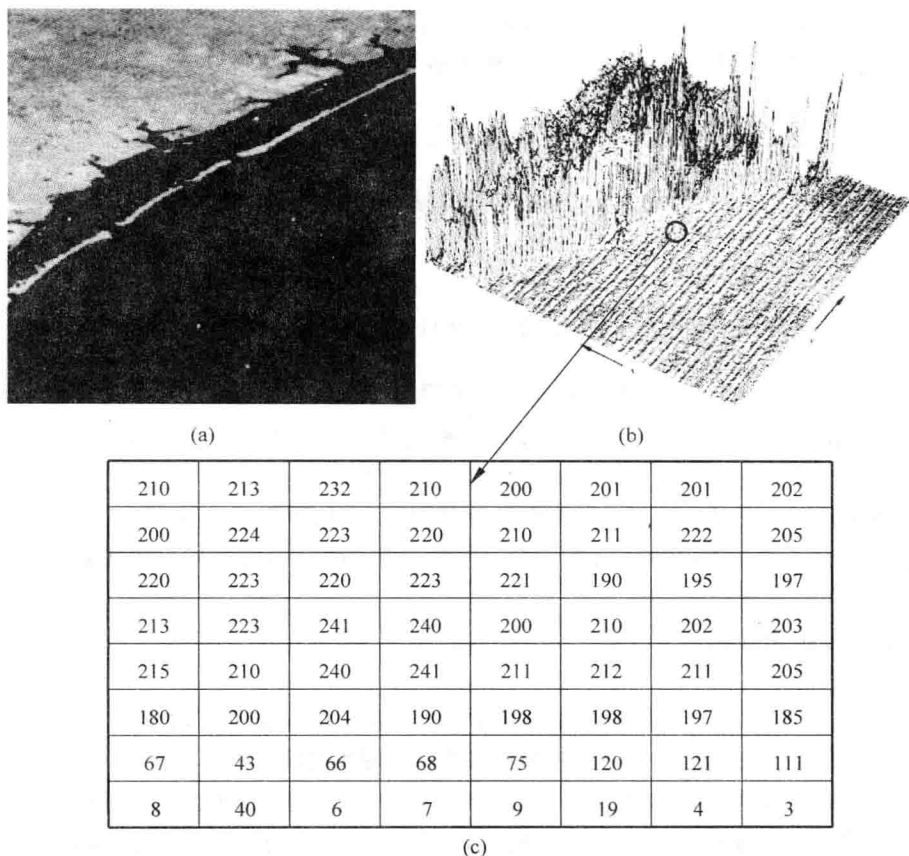


图 1.2 图像离散化的结果

所有这些,归为一点:数字电子计算机系统本质上是离散的。

由于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,因此,无论计算机科学本身,还是与计算机科学及其应用密切相关的现代科学研究领域,都面临着如何对离散结构建立相应的数学模型;又如何将已用连续数量关系建立起来的数学模型离散化,从而可由计算机加以处理的问题。

离散数学的基本概念、基本理论和基本方法大量地应用在数字电路、编译原理、数据结构、操作系统、数据库系统、算法的分析与设计、人工智能、计算机网络等专业课程中;同时,该课程所提供的训练十分有益于学生概括抽象能力、逻辑思维能力、归纳构造能力的提高,十分有益于培养学生严谨、完整、规范的科学态度。

1.2 离散数学与计算机的关系

1.2.1 数学是计算机的基础

计算机自从其诞生之日起,它的主要任务就是进行各种各样的科学计算。文档处理、数据处理、图像处理、硬件设计、软件设计等,可以被抽象为两大类:数值计算与非数值计算。