



21世纪高等院校教材·地理信息系统教学丛书



地理信息系统算法基础

◎ 张 宏 温永宁 刘爱利 等 编著



科学出版社
www.sciencep.com

21世纪高等院校教材·地理信息系统教学丛书

地理信息系统算法基础

张 宏 温永宁 刘爱利 等 编著

科学出版社

北京

内 容 简 介

本书全面、系统地收集和整理了当前地理信息系统算法领域的相关资料,以地理信息系统设计与实现为线索,内容涉及地理空间数据的描述、检索、存储和管理,以及地理空间信息分析基本方法的设计和实现。

本书可作为地理信息系统专业的本科生和研究生教材,也可作为从事地理信息系统软件开发和应用的人员的学习资料,并可供地理信息系统的理论研究人员参考。

图书在版编目(CIP)数据

地理信息系统算法基础/张宏等编著.一北京:科学出版社,2006

(21世纪高等院校教材·地理信息系统教学丛书)

ISBN 7-03-016868-2

I . 地… II . 张… III . 地理信息系统-算法理论-高等学校-教材
IV . P208

中国版本图书馆 CIP 数据核字(2006)第 009556 号

责任编辑:郭 森 杨 红 李久进 / 责任校对:刘小梅

责任印制:张克忠 / 封面设计:高海英

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮 政 编 码: 100717

<http://www.sciencep.com>

新 葡 印 刷 厂 印 刷

科 学 出 版 社 发 行 各 地 新 华 书 店 经 销

*

2006 年 6 月第 一 版 开本: B5(720×1000)

2006 年 6 月第一次印刷 印张: 22

印数: 1—3 500 字数: 410 000

定 价: 35.00 元

(如有印装质量问题,我社负责调换(路通))

《地理信息系统教学丛书》编委会

顾问 陈述彭 王家耀 孙九林 李小文 李德仁
承继成 高俊 童庆禧 廖克

主编 阎国年

副主编 王桥 汤国安 盛业华 黄家柱

委员 (按姓氏笔画排序)

丁一	王春	雷伟	王卫国	王富兰	刘明媛	安年	中斌	梅霞	伟英	盼	卫苗	苗旺	丽峰	海明	苗
王平	韦春	伟本	邓勇	毕硕	刘基	余江	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中
建毅	田冉	本剋	勇	刘基	如	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
龙伟	乔春	机武	邓	任利	刘爱	余江	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中
乔峰	延冉	利宏	邓	刘宏	孙利	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
刘学	晓春	庆文	邓	刘宏	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
许婷	艳琴	文鹏	邓	孙亚	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
严荣	琴平	邦婷	邓	苏乐	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
李云	平源	婷鸣	邓	李发	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
杨旭	昕生	邦婷	邓	杨平	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
吴长	吴平	婷鸣	邓	吴平	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
宋亚	张源	利鸣	邓	张亦	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
张书	张亦	利鸣	邓	张亦	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
张海	张亦	利鸣	邓	陈林	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
陈惠	陈林	伟忠	邓	郑林	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
周国	周伟	忠涛	邓	郑贺	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
闾年	闾永	伟涛	邓	徐贺	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
徐敏	徐永	伟发	邓	徐徐	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
高黎	高晓	鹏发	邓	唐陶	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
龚霞	晓黎	鹏发	邓	盛常	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
蒋琴	晓敏	鹏发	邓	东曾	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	
缪深	深潘	鹏发	邓	巧玲	利宏	江硕	波军	华鑑	善踊	山翎	丁华柱	中宁	秀家	中	

序

南京师范大学地理科学学院发起并组织编写地理信息系统专业系列教材,奋斗三载,先后问世,这是我国第一套全面阐述地理信息系统理论、方法、技术和应用的教科书。对于地理学科的现代化,信息科学新型人才的培训,落实科教兴国战略,深化教学改革来说,都是值得庆贺的。

据中国科学院地学部调查(2002),全国综合性大学共有150个地理学科机构,在地学领域中居首位,而地理信息系统专业脱颖而出,发展最快。21世纪初,已设置地理信息系统专业的学校有70多个,仅江苏省内就有12个。这是经济发展、社会进步的客观需求。面对全社会数字化的浪潮,“数字地球”、数字化城市、省区与流域,百舸争流。地理信息系统作为人口、资源与环境问题的公共平台,作为国家推动信息化、实现现代化的重要组成部分,正在与电子政务、电子商务信息系统相融合,愈来愈显示出其跨行业、多功能的优势,不断开拓新的应用领域。一些涉及地理分布现象的数据采集、时空分析,涉及城市或区域规划、管理与决策的过程,都喜欢用上地理信息系统这种新的技术手段,来提高办公自动化的水平,提高企业科学管理的效率和透明度,加强面对国际市场的开放力度和竞争能力。近20年来,全国范围从事地理信息系统的事业、企业单位,迅猛增长,已超过400个,而且方兴未艾,与时俱进。

中国科学院地学部地学教育研究组在咨询报告(2002)中指出:“随着社会和科技的发展,地学的内涵、性质和社会功能也在变化。这在最近20年中尤为明显:遥感、信息技术和各种实时观测、分析技术的发展,使地球科学进入了覆盖全球、穿越圈层,即地球系统科学的新阶段,从局部现象的描述,推进到行星范围的推理探索,获得了全球性和系统性的信息。”这就是说,从学科的本质及其自身发展的规律来看,地理信息系统不仅仅是技术,而且是科学,是发展地球系统科学不可缺少的部分。

地理信息系统之所以一枝独秀,并非偶然!主要是由于它本身具备多样化的社会功能。社会信息化的主要内容包括三个方面:一是信息基础设施的建设,地理信息系统正是地图测绘的数字化产品,同时又是兼收并容遥感、定位系统的缓冲区,起着调节网络信息流的作用;二是产业结构调整,地理信息系统起着润滑剂的作用,以信息流调控物流、能流和人流,以信息化促进现代化;三是信息服务,地理信息系统是电子政务、电子商务信息系统不可分割的组成部分。在航天事业、电信

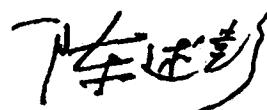
网络和电脑技术日新月异的 21 世纪,地理信息系统如虎添翼,广泛地渗透到各行各业之中,提供无微不至的信息服务。

地理信息系统教材,前人多以综论形式出版。例如,英文教材先后有 D. R. Taylor(1991),J. C. Autenucci et al. (1991),M. D. I. Goodchild(1991),M. M. Fisher(1993),Murai Shuji(1996),D. Rhind(2000);中文教材先后有黄杏元、汤勤(1989),边馥苓(1996),陈述彭、鲁学军、周成虎(1999),龚健雅(1999),邬伦(1999),闾国年、吴平生、周晓波(1999),李德仁、关泽群(2000),马蔼乃(2000),王家耀(2001)等。这些教材对地理信息系统的科学与哲学性质,及其与邻近学科的相互关系,均有精辟论述。地理信息系统应用专论方面,城市:曹桂发等(1991),宋小冬、叶嘉安(1995),宫鹏(1996),陈述彭(1999),张新长等(2001);林业:李芝喜、孙俊平(2000);农业:王人潮(1999)。这些专论密切结合相关行业,具有中国特色。现在,闾国年教授等主持编写的地理信息系统专业系列教材,是在前人工作的基础上,博采众家之所长,继往开来,推陈出新,拓展为系列教材。基础是扎实的,时机是成熟的。

这套系列教材的编写,紧密结合地理信息系统专业的课程设置。在理论方面,又推出了一部新作,从哲学的高度来探讨地理信息系统中的虚拟时空。系列教材的重点侧重于方法、技术,总结了数据集成、知识发现的最新进展;率先推出数据共享、虚拟环境与网络三部分,反映地理信息系统的生长点。在应用方面,主要是结合作者们近年参与建设项目的实践,加以总结和提高,是来自生产第一线的“新知”,目前已涉及土地与水资源管理、城市规划、环境保护以及设备设施管理与房产管理等,今后随着应用领域的拓展,还会有旅游、物流等地理信息系统教材相继问世。

同学们可以根据课程设置计划,循序渐进,在理论方面广泛涉猎,解放思想,开阔眼界。在方法、技术方面,配合辅导教材和实习大纲,刻苦钻研,掌握关键技术,学以致用。在应用方面结合个人志趣、专长与就业需求,选修其中一二门,理清不同行业的应用特点,举一反三。系列教材是面向整个专业的,并不要求每位同学都把全部教材囫囵吞咽下去,食而不化。编写系列教材,正是为同学们提供了更加宽阔的学习园地、更加宽松的学习环境。祝同学们健康成长,时刻准备着,与时俱进,开拓创新,为祖国信息化和现代化多做贡献。

中国科学院院士



2003 新年

前　　言

地理信息系统(GIS)自 20 世纪 60 年代问世以来,已历经 40 多个春秋。随着计算机技术、通信技术的普及, GIS 已经走出实验室,在各行业内得到了广泛的应用并取得了良好的经济效益和社会效益。越来越多的来自不同行业、不同专业的人士开始利用 GIS 技术解决生产和社会实践中的问题。

本书旨在对 GIS 基础软件、应用软件,以及 GIS 应用过程中涉及的基本的 GIS 算法及其应用做一较为全面系统的介绍和分析。“算法”一词源于公元 9 世纪波斯数学家比阿勒·霍瓦里松的一本影响深远的著作《代数对话录》。英国数学家图灵在 20 世纪提出了著名的图灵论点,并抽象出了一台机器,这台机器被我们称为图灵机。图灵的思想对算法的发展起到了重要的作用。算法是指完成一个任务所需要的具体步骤和方法。完成同样任务,不同的算法可能使用不同的时间和空间。算法是计算机处理信息的本质,因为计算机程序本质上是一个算法,告诉计算机确切的步骤来执行一个指定的任务。一般地,当算法在处理信息时,数据会从输入设备读取,写入输出设备,可能保存起来以供后期使用。随着地理信息系统在各个领域中更为广泛的应用,对 GIS 算法的研究变得日益重要。GIS 算法作为处理地理科学领域中各种问题的分析求解方法,有着鲜明的特点。首先,GIS 算法是用来解决地学领域中的问题,但许多算法都不是孤立的,不是无源之水,无本之木,而是借鉴和发展了其他学科的研究成果;其次,GIS 算法处理的往往是海量的地理信息,涉及许多复杂的空间运算,不同于简单的数据查询、编辑操作;再次,地理信息系统与实际应用、工程开发有着密切的关系,GIS 算法与一般算法很重要的一个区别就是要处理问题的不确定性,它无法被定性、定量成一个非常明确的纯算法问题。例如,标注一条河流,标注必须靠近河流但是不能与其相交,标注的走向与河流一致但又不能覆盖地图上的其他要素,标注必须间隔一致但又不允许太大也不能太小等,不一而足。本书将更多地从计算机算法的角度来阐述 GIS 算法的设计原则、分析方法、技巧以及相关的评价。

GIS 与各类学科都有密切的联系, GIS 算法与地理科学、计算机科学、数学等同样有着千丝万缕的关系。GIS 的许多算法都是从计算几何、计算图形学、离散数学演化而来的。它是整个地理信息科学的核心,不管是从基本的 GIS 空间数据结构到空间数据模型,还是从必需的 GIS 空间关系的表达与描述到各种各样的空间拓扑关系,抑或是从高级的时态多维 GIS 到 GIS 空间数据挖掘与知识发现, GIS

算法作为地理信息系统的基石当之无愧。GIS 算法是一个富集优雅技术和复杂数学分析结果的领域,一个好的算法或数据结构可能使某个原来需要数月才能完成的问题在顷刻之间得到解决。在实际的应用中,往往需要考虑各方面的因素,可能需要利用时间来换取空间,也有可能必须牺牲效率来获得最佳的存储性能,换句话来说,即可能没有放之四海而皆准的最优算法。因此掌握 GIS 算法的设计原则,剖析各种 GIS 分析方法,灵活应用一些技巧自有其用武之地。

本书第 1 章由张宏、温永宁编写;第 2 章由张宏、张强、刘二年编写;第 3 章由张宏、丁一编写;第 4 章由张宏、乔延春编写;第 5 章由张宏、冯文钊、乔延春编写;第 6 章由张宏、温永宁编写;第 7 章由张宏、刘二年编写;第 8 章由张宏、温永宁编写;第 9 章由张宏、温永宁编写;第 10 章由刘二年、张宏、丰江帆编写;第 11 章由张宏、陈洋编写;第 12 章由汤国安、刘爱利编写;第 13 章由刘爱利、张宏、毕硕本编写;第 14 章由张宏、徐洁编写。张强、刘瑜、巢俊杰、丰江帆、徐洁帮助整理了书中的插图和图表。最后由张宏、温永宁、刘爱利统稿,张宏定稿。

在全书的撰写过程中,始终得到了闾国年教授、黄家柱教授、盛业华教授、汤国安教授、王桥教授、刘晓艳博士的指导和帮助。他们为本书的撰写提供了充裕的科研和设备条件,以及丰富的参考资料,帮助作者拟定教材提纲、审阅相关章节,并提出宝贵的修改意见。特别是闾国年教授严谨求实的作风使作者不敢有丝毫的懈怠。本书的完成得到了实验室诸多同事和学生的热心支持,他们是李斌、陆娟、张亦含、陈踊、田冉、蒋海琴、马刚、张金善、朱明媛、蒋文明、谈帅、王帮进、邓勇伟,在此谨向他们表示诚挚的谢意。

本书的内容引用了很多前人的工作成果,在此谨向他们表示诚挚的敬意。同时由于作者水平有限,书中不足之处在所难免,敬请读者批评指正。

作 者
2005 年 11 月于随园

目 录

序

前言

第1章 算法设计和分析	(1)
1.1 概述	(1)
1.2 算法设计原则	(1)
1.3 算法复杂性的度量	(2)
1.3.1 时间复杂性	(2)
1.3.2 空间复杂性	(8)
1.4 最优算法	(8)
1.5 算法的评价	(9)
1.5.1 如何估计算法运行时间	(9)
1.5.2 最坏情况和平均情况的分析	(11)
1.5.3 平摊分析	(13)
1.5.4 输入大小和问题实例	(13)
思考题	(14)
第2章 GIS 算法的计算几何基础	(15)
2.1 维数扩展的 9 交集模型	(15)
2.1.1 概述	(15)
2.1.2 模型介绍	(15)
2.1.3 空间关系的判定	(17)
2.2 矢量的概念	(21)
2.2.1 矢量加减法	(21)
2.2.2 矢量叉积	(22)
2.3 折线段的拐向判断	(22)
2.4 判断点是否在线段上	(23)
2.5 判断两线段是否相交	(23)
2.6 判断矩形是否包含点	(25)
2.7 判断线段、折线、多边形是否在矩形中	(25)
2.8 判断矩形是否在矩形中	(25)

2.9 判断圆是否在矩形中.....	(25)
2.10 判断点是否在多边形内	(25)
2.10.1 射线法	(26)
2.10.2 转角法	(28)
2.11 判断线段是否在多边形内	(30)
2.12 判断折线是否在多边形内	(32)
2.13 判断多边形是否在多边形内	(32)
2.14 判断矩形是否在多边形内	(33)
2.15 判断圆是否在多边形内	(33)
2.16 判断点是否在圆内	(33)
2.17 判断线段、折线、矩形、多边形是否在圆内.....	(33)
2.18 判断圆是否在圆内	(33)
2.19 计算两条共线的线段的交点	(33)
2.20 计算线段或直线与线段的交点	(34)
2.21 求线段或直线与圆的交点	(35)
2.22 中心点的计算	(36)
2.23 过点作垂线	(37)
2.24 作平行线	(37)
2.25 过点作平行线	(38)
2.26 线段延长	(38)
2.27 三点画圆	(39)
2.28 线段打断	(39)
2.29 前方交会	(40)
2.30 距离交会	(41)
2.31 极坐标作点	(43)
思考题	(44)
第3章 空间数据的变换算法	(45)
3.1 平面坐标变换.....	(45)
3.1.1 平面直角坐标系的建立	(45)
3.1.2 平面坐标变换矩阵	(45)
3.1.3 平移变换.....	(46)
3.1.4 比例变换.....	(47)
3.1.5 对称变换.....	(47)
3.1.6 旋转变换.....	(47)
3.1.7 错切变换.....	(48)

3.1.8 复合变换	(48)
3.1.9 相对(x_f, y_f)点的比例变换	(49)
3.1.10 相对(x_f, y_f)点的旋转变换	(49)
3.1.11 几点说明	(50)
3.2 球面坐标变换	(50)
3.2.1 球面坐标系的建立	(50)
3.2.2 确定新极 Q 地理坐标中 φ_0, λ_0	(51)
3.3 仿射变换	(54)
3.4 地图投影变换	(55)
3.4.1 概述	(55)
3.4.2 地球椭球体的相关公式	(57)
3.4.3 兰勃特投影	(62)
3.4.4 墨卡托投影	(64)
3.4.5 高斯-克吕格投影	(65)
3.4.6 通用横轴墨卡托投影	(68)
思考题	(69)
第 4 章 空间数据转换算法	(70)
4.1 矢量数据向栅格数据转换	(70)
4.1.1 矢量点的栅格化	(70)
4.1.2 矢量线的栅格化	(71)
4.1.3 矢量面的栅格化	(73)
4.2 栅格数据向矢量数据转换	(77)
4.2.1 栅格点坐标与矢量点坐标的关系	(77)
4.2.2 栅格数据矢量化的基本步骤	(78)
4.2.3 线状栅格数据的细化	(78)
4.2.4 多边形栅格转矢量的双边界搜索算法	(84)
4.2.5 多边形栅格转矢量的单边界搜索算法	(86)
思考题	(89)
第 5 章 空间数据组织算法	(90)
5.1 矢量数据的压缩	(90)
5.1.1 间隔取点法	(90)
5.1.2 垂距法和偏角法	(90)
5.1.3 道格拉斯-普克法	(91)
5.1.4 光栏法	(93)
5.1.5 曲线压缩算法的比较	(95)

5.1.6 面域的数据压缩算法	(96)
5.2 棚格数据的压缩.....	(98)
5.2.1 链式编码.....	(98)
5.2.2 游程长度编码	(99)
5.2.3 块式编码	(101)
5.2.4 差分映射法	(101)
5.2.5 四叉树编码	(102)
5.3 拓扑关系的生成	(102)
5.3.1 基本数据结构	(103)
5.3.2 弧段的预处理	(106)
5.3.3 结点匹配算法	(112)
5.3.4 建立拓扑关系	(114)
思考题.....	(117)
第6章 空间度量算法.....	(118)
6.1 直线和距离	(118)
6.1.1 直线	(118)
6.1.2 直线方程	(119)
6.1.3 点到直线的距离	(120)
6.2 角度量算	(125)
6.3 多边形面积的量算	(125)
6.3.1 三角形面积量算	(125)
6.3.2 四边形面积量算	(127)
6.3.3 任意二维平面多边形面积量算	(129)
6.3.4 任意三维平面多边形面积量算	(131)
思考题.....	(134)
第7章 空间数据索引算法.....	(135)
7.1 B树与 B ⁺ 树	(135)
7.1.1 B树索引结构	(136)
7.1.2 B ⁺ 树索引结构	(140)
7.2 R树结构	(141)
7.2.1 R树定义	(141)
7.2.2 R树索引的主要操作算法	(142)
7.2.3 R [*] 树算法	(146)
7.3 四叉树结构	(147)
7.3.1 常规四叉树	(147)

7.3.2 线性四叉树	(149)
7.3.3 线性四叉树的编码	(149)
7.3.4 Z 曲线和 Hibert 曲线算法	(155)
思考题.....	(158)
第 8 章 空间数据内插算法.....	(159)
8.1 概述	(159)
8.1.1 几何方法	(159)
8.1.2 统计方法	(159)
8.1.3 空间统计方法	(160)
8.1.4 函数方法	(160)
8.1.5 随机模拟方法	(160)
8.1.6 确定性模拟	(161)
8.1.7 综合方法	(161)
8.2 分段圆弧法	(161)
8.3 分段三次多项式插值法	(162)
8.3.1 三点法	(162)
8.3.2 五点法	(162)
8.4 趋势面插值算法	(163)
8.5 反距离权重插值算法	(166)
8.6 双线性插值算法	(166)
8.7 薄板样条函数法	(167)
8.7.1 薄板样条函数法	(167)
8.7.2 规则样条函数	(168)
8.7.3 薄板张力样条法	(168)
8.8 克里金法	(169)
8.8.1 普通克里金法	(169)
8.8.2 通用克里金法	(171)
思考题.....	(171)
第 9 章 Delaunay 三角网与 Voronoi 图算法.....	(172)
9.1 概述	(172)
9.2 Voronoi 图	(173)
9.3 Delaunay 三角形	(174)
9.4 Voronoi 图生成算法	(174)
9.4.1 半平面的交	(174)
9.4.2 增量构造方法	(175)

9.4.3 分治算法	(178)
9.4.4 减量算法	(180)
9.4.5 平面扫描算法	(182)
思考题.....	(184)
第 10 章 缓冲区分析算法	(185)
10.1 概述.....	(185)
10.2 缓冲区边界生成算法基础.....	(185)
10.3 点缓冲区边界生成算法.....	(187)
10.4 线缓冲区边界生成算法.....	(188)
10.5 面缓冲区边界生成算法.....	(195)
10.6 多目标缓冲区合并算法.....	(195)
思考题.....	(200)
第 11 章 网络分析算法	(201)
11.1 概述.....	(201)
11.2 网络数据模型.....	(201)
11.3 路径分析算法.....	(203)
11.3.1 单源点的最短路径	(204)
11.3.2 单目标最短路径问题	(208)
11.3.3 单结点对间最短路径问题	(208)
11.3.4 多结点对间最短路径问题	(209)
11.3.5 次短路径求解算法	(209)
11.4 最佳路径算法.....	(210)
11.4.1 最大可靠路径	(210)
11.4.2 最大容量路径	(211)
11.5 连通性分析算法.....	(212)
11.5.1 Prim 算法	(212)
11.5.2 Kruskal 算法	(214)
11.6 资源分配算法.....	(216)
思考题.....	(218)
第 12 章 地形分析算法	(219)
12.1 数字地面模型的生成算法.....	(219)
12.1.1 基于离散点的 DEM 规则网格的生成	(219)
12.1.2 基于不规则三角网的 DEM 生成	(221)
12.1.3 DEM 数据结构的相互转换	(221)
12.2 基本地形因子分析算法.....	(235)

12.2.1 坡面因子提取的算法基础	(235)
12.2.2 坡度、坡向	(239)
12.2.3 坡形	(241)
12.3 地形特征提取算法	(244)
12.3.1 地形特征点的提取	(245)
12.3.2 基于规则格网 DEM 数据提取山脊与山谷线的典型算法	(246)
12.4 通视分析算法	(248)
12.4.1 判断两点之间的可视性的算法	(248)
12.4.2 计算可视域的算法	(248)
思考题	(249)
第 13 章 空间数据挖掘算法	(250)
13.1 概述	(250)
13.2 分类算法	(250)
13.2.1 数据分类的基本过程	(250)
13.2.2 决策树分类概述	(251)
13.2.3 决策树的特点	(251)
13.2.4 二叉决策树算法与分类规则的生成	(252)
13.2.5 决策树分类算法	(253)
13.2.6 决策树属性的选取	(254)
13.2.7 改进决策树性能的方法	(255)
13.3 泛化规则算法	(256)
13.3.1 概念层次	(256)
13.3.2 面向属性泛化的策略与特点	(258)
13.3.3 基于规则的面向属性泛化方法	(260)
13.4 相关分析	(263)
13.4.1 两要素间的相关分析	(263)
13.4.2 多要素之间的相关分析	(265)
13.4.3 关联规则算法	(267)
13.5 回归分析	(274)
13.5.1 一元线性回归模型	(274)
13.5.2 多元线性回归模型	(277)
13.5.3 非线性回归模型	(281)
13.5.4 回归分析与相关分析	(282)
13.6 系统聚类分析	(283)
13.6.1 概述	(283)

13.6.2 聚类要素预处理	(283)
13.6.3 分类统计量	(284)
13.6.4 系统聚类法	(286)
13.6.5 其他聚类方法概述	(292)
13.7 判别分析.....	(293)
13.7.1 距离判别	(294)
13.7.2 费歇判别法	(297)
13.7.3 贝叶斯判别法	(300)
13.7.4 判别分析应注意的问题	(302)
13.8 主成分分析.....	(303)
13.8.1 主成分分析的基本原理	(303)
13.8.2 主成分分析的方法	(304)
思考题.....	(306)
第14章 数据输出算法	(308)
14.1 概述.....	(308)
14.1.1 地图符号构成元素组成	(308)
14.1.2 地图符号几何特征	(309)
14.1.3 基于SVG的地图符号描述模型.....	(310)
14.2 点状地图符号的绘制.....	(312)
14.2.1 圆的绘制	(312)
14.2.2 椭圆的绘制	(314)
14.2.3 多边形的绘制	(314)
14.2.4 五角星的绘制	(316)
14.3 线状地图符号的绘制.....	(318)
14.3.1 平行线绘制	(319)
14.3.2 虚线绘制	(321)
14.3.3 短齿线的绘制	(322)
14.3.4 铁路线的绘制	(324)
14.3.5 境界线的绘制	(326)
14.4 面状地图符号的绘制.....	(328)
思考题.....	(333)
主要参考文献.....	(334)

第1章 算法设计和分析

1.1 概述

算法是解决问题方法的精确描述,但是并不是所有问题都有算法,有些问题经研究可行,则相应有算法,但这并不是说问题就有结果。上述的“可行”,是指对算法的研究。

1. 待解问题的描述

待解问题的描述应精确、简练、清楚,使用形式化模型刻画问题是最恰当的。例如,使用数学模型刻画问题是最简明、严格的,一旦问题形式化了,就可依据严格的模型对问题求解。

2. 算法设计

算法设计的任务是对各类具体问题设计良好的算法及研究设计算法的规律和方法。常用的算法有穷举搜索法、递归法、回溯法、贪心法、分治法等。

3. 算法分析

算法分析的任务是对设计出的每一个具体的算法,利用数学工具,讨论各种复杂度,以探讨某种具体算法适用于哪类问题,或某类问题宜采用哪种算法。

1.2 算法设计原则

设计的真谛,就是在一些互相冲突的需求和约束条件之间寻找平衡点。什么是一个良好的算法?这个问题可以从许多方面来阐述,不管是哪种情况,都无法放之四海而皆准,但总有一些原则是在不断的探索中被发现并潜移默化地影响着我们的思维。算法的设计应当遵循以下一些原则:

1. 正确性

算法的正确性是指对于一个问题,之所以将其放在第一位是因为如果一个算法自身有缺陷,或者不适合于问题的求解,那么该算法将不会解决问题。