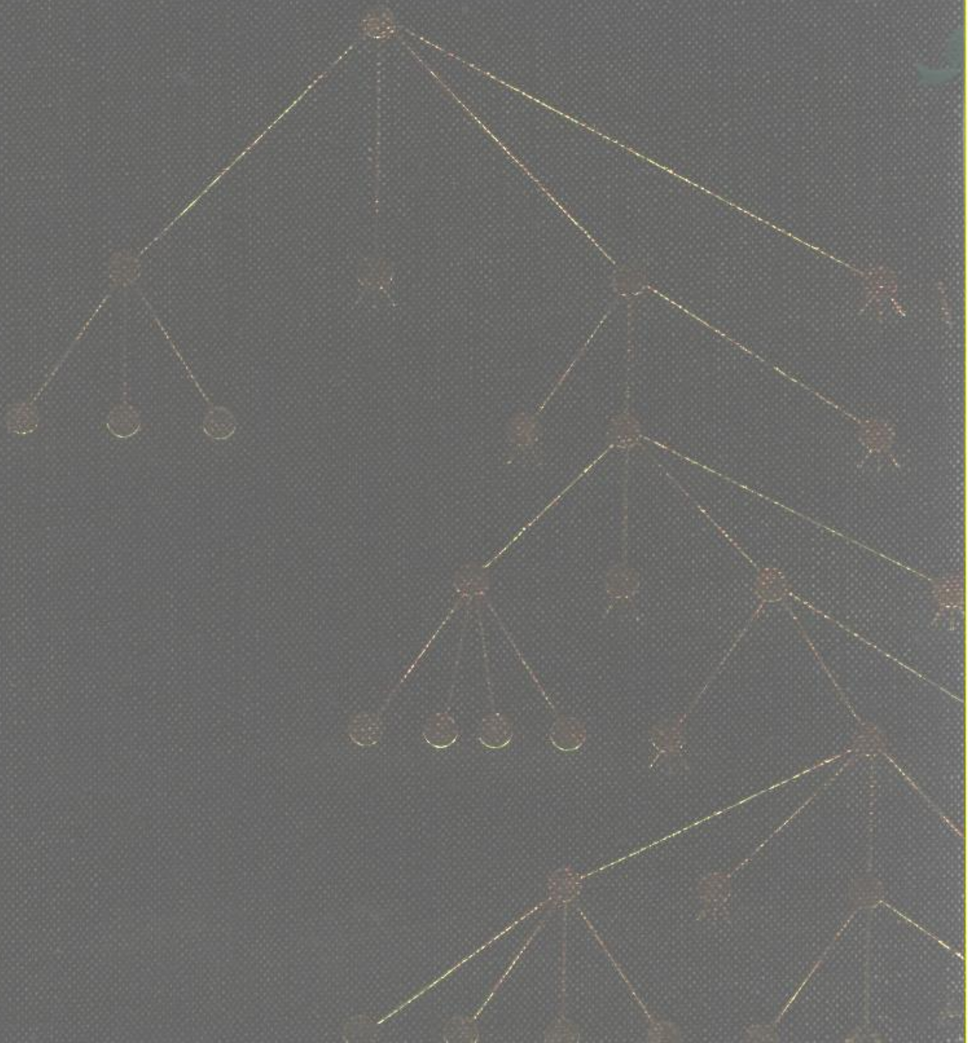
树的枚举与算法复杂性分析

王振宇 著



树的枚举与算法复杂性分析

王振宇　著

国防工业出版社

(京)新登字106号

内 容 简 介

本书系统总结了作者近10年来在树枚举及其在算法复杂性分析的应用中的研究成果。全书共7章：第一章绪论，主要介绍与树枚举及与算法复杂性有关的基本概念；第二章介绍研究树枚举问题的预备知识；第三章阐述各种树枚举问题；第四章论述树上的各种枚举问题；第五章研究树上的可加枚举问题；第六章探讨树结构上的可加复杂性算法；第七章给出几个具体算法的复杂性分析；附录部分列出12种枚举值表。

树枚举不仅在图论、离散数学、组合论、渐近分析方面具有理论意义，而且是对算法设计、程序语言、编译技术、程序设计方法学、数据处理等做算法复杂性分析的有力工具。本书可供上述各方面的科技人员、高等院校教师、研究生、高年级学生参阅。

树的枚举与算法复杂性分析

王振宇 著

责任编辑 周烈强

*

国防工业出版社 出版发行

（北京市海淀区紫竹院南路23号）

（邮政编码 100044）

新华书店经售

国防工业出版社印刷厂印装

*

850×1168 1/32 印张 $8^5/_8$ 220千字

1991年12月第一版 1991年12月第一次印刷 印数：0001— 1500册

ISBN 7-118-00896-6/TP·118 定价：9.80元

致读者

本书由国防科技图书出版基金资助出版。

国防科技图书出版工作是国防科技事业的一个重要方面。优秀的国防科技图书既是国防科技成果的一部分，又是国防科技水平的重要标志。为了促进国防科技事业的发展，加强社会主义物质文明和精神文明建设，培养优秀科技人才，确保国防科技优秀图书的出版，国防科工委于1988年初决定每年拨出专款，设立国防科技图书出版基金，成立评审委员会，扶持、审定出版国防科技优秀图书。

国防科技图书出版基金资助的对象是

1．学术水平高，内容有创见，在学科上居领先地位的基础科学理论图书；在工程技术理论方面有突破的应用科学专著。

2．学术思想新颖，内容明确、具体，有突出创见，对国防科技发展具有较大推动作用的专著；密切结合科学技术现代化和国防现代化需要的高科技内容的专著。

3．有重要发展前景和有重大开拓使用价值，密切结合科学技术现代化和国防现代化需要的新技术、新工艺内容的科技图书。

4．填补目前我国科学技术领域空白的薄弱学科的科技图书。

国防科技图书出版基金评审委员会在国防科工委的领导下开展评审工作，职责是：负责掌握出版基金的使用方向，评审受理的图书选题，决定资助的图书选题和资助金额，以及决定中断或取消资助等。经评审给予资助的图书，由国防工业出版社列选出版。

国防科技事业已经取得了举世瞩目的成就。国防科技图书承担着记载和弘扬这些成就，积累和传播科技知识的使命。在改革

开放的新形势下，国防科工委率先设立出版基金，扶持出版科技图书，这是一项具有深远意义的创举。此举势必促使国防科技图书的出版随着国防科技事业的发展而更加兴旺。

设立出版基金是一件新生事物，是对出版工作的一项改革。因而，评审工作需要不断地摸索、认真地总结和及时地改进，这样，才能使有限的基金发挥出巨大的效能。评审工作更需要国防科技工业战线广大科技工作者、专家、教授，以及社会各界朋友的热情支持。

让我们携起手来，为祖国昌盛、科技腾飞、出版繁荣而共同奋斗。

国防科技图书出版基金
评审委员会

国防科技图书出版基金
第一届评审委员会组成人员

序

树的枚举或计数，在数学中已不是一个新题目，著名的卡塔兰（Catalan）数就诞生在上个世纪。波里亚（Pólya）计数定理也有半个世纪的历史了。计算机的出现和计算机算法的深入研究又给树枚举理论研究以新的推动，注入了新的活力。这里，最深刻的原因是：树，不再仅仅是一种抽象的数学对象（以及大自然中生长的一种植物），而且是计算机算法中的常用非线性分级数据结构。树结构上的算法研究不仅在理论上处于计算机科学的核心，而且有着实用的技术背景。这类算法的性能分析，尤其是平均复杂性分析，对算法的设计、比较、选择使用有着指导意义。树的枚举和树上属性的枚举正是此类算法分析的有力且必不可少的工具。

1980年，对一个算法的分析引起了我对树枚举研究的兴趣。当我得到了具有n个内节点和m个叶子的有序树的个数$O_{n,m}$的生成函数$O(x, y)$，并发现它关于x，y的对称性，从而得出结论$O_{n,m}=O_{m,n}$和“有序树上平均有一半节点是叶子”时，我深深被这种不可言状的和谐和优美所陶醉，这也使我深刻认识到用分析方法处理离散问题的巨大潜力。

自此以后，对树的研究兴趣10年不衰。得益于早年数学工作的一点功底，加上算法复杂性分析的明确应用目标，树枚举研究不断找到新的研究课题。在解决一些具体枚举问题的基础上，提炼出了可加枚举问题的概念。这一概念覆盖了树上数量众多、形态各异的枚举问题。在建立了统一求解模式的理论框架之后，又发现了树结构上算法的复杂性分析与树枚举问题的平行性，并提出了可加复杂性算法的概念。原来，树结构上相当多的递归算法的复杂性（尤其是时间复杂性）具有可加性。当把这些工具用于

递归程序和非递归程序的复杂性分析时，许多复杂性的计算竟然变成简单的查表和线性组合。

这就是我对这项研究的简单回顾，本书就是这项研究的总结。现在，我把本书奉献给读者和同行——数学领域的和计算机领域的，恳请不吝指正。

书中部分结果在这里是第一次发表。

本书各章的内容安排体现如下的思路：第一章绪论引入主题，即树、树枚举、树算法的复杂性分析；第二章介绍几个主要常用的数学工具；第三章至第七章系统地归纳本书作者、合作者及有关国内外同行在树的枚举方面的工作——引进的一些重要概念、一系列具体结果、树的枚举和树结构上算法复杂性分析之间的平行方法学，论述这些结果在算法复杂性分析方面的平行结果和应用。几个附录给出了树枚举问题和树上的枚举问题的一系列计算结果。

需要强调指出的是：在树枚举方面的结果不仅是作为算法分析的工具，而且有其独立的数学结果的意义。正因为如此，这些结果有不少曾在数学刊物上发表，并为国际数学评论刊物所收录和评介。

致　谢

这项研究开始时是自由选题，以后得到中国船舶工业总公司的支持，后期(1988～1990年)又得到国家自然科学基金的资助。实际上，本书也是此项资助课题的成果。因此，这里首先向中国船舶工业总公司和国家自然科学基金委员会表示感谢。

1983年7月，中国船舶工业总公司派遣我出席第10届国际自动机、语言和程序设计讨论会（ICALP 83′，西班牙巴塞罗那市)。承蒙会议主席、东道主J.迪亚斯（Diaz）教授邀请，并为我安排一个会后报告会，使我有机会同巴塞罗那工业大学的众多同行、法国INRIA的Ph.弗列若里特（Flajolet）教授、意大利米兰大学的G.毛利（Mauri）教授等得以进行深入的学术思想交流，以后又多次收到他们寄来新发表的论文。加拿大阿尔伯塔

（Alberta）大学的J. W. 穆恩（Moon）教授就我所提出的问题再作研究并著文发表，还寄来他的多篇新作（见参考文献〔1〕）。

在工作过程中，当时在学的一些研究生，如臧小林、孙潮义、赵一心都不同程度地参与一部分工作、孙潮义和赵一心的工作更多些，书中也把这些工作收集进来了。

我国计算机界的前辈慈云桂教授、武汉大学曾宪昌教授、康立山教授、戴大为教授、华中理工大学刘键教授、黄文奇教授、清华大学郑维民副教授都曾对此项研究工作给以支持。

在此，向国内外的这些同行们一并表示感谢。

我还要感谢我的妻子周豫婷副教授。她有自己的事业，但她一如既往，又在我这项研究工作过程中，从生活和工作方面给了我无法枚举或计数的支持。在成书过程中，她又给了我许多帮助。可以说，没有这些，也就不可能有这本书。

王振宇　识

1990年 9 月，武昌

目 录

第一章　绪　　论

1.1　数学中的树和计算机科学中的树

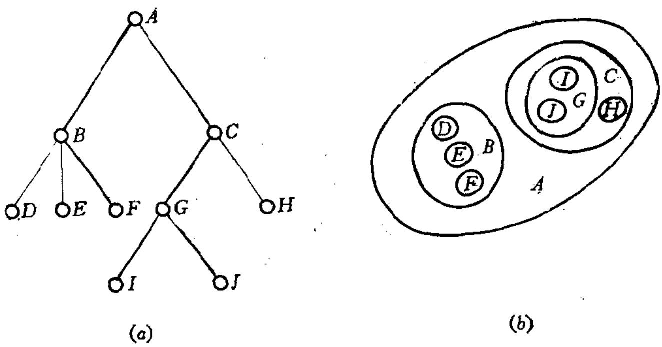
树，在数学的图论或所谓离散数学领域中是一类重要的研究对象，并已有了相当长久的历史（当然要比它在大自然中的历史短得多)。树，就是不含有回路的连通图。如所周知，克希霍夫(Kirchoff)1874年在解决电路理论中求解联立方程的问题时首先提出了树的概念。之后，树的概念被逐步用于研究代数公式的结构，以及化学中的同分异构体问题。但总的来说，树的研究并未受到太多的重视。

计算机科学和信息科学的出现极大地促进了树的研究和应用。实际上，连包括树的研究在内的整个离散数学学科在近几十年中的迅猛发展也主要是由于受了计算机科学和信息科学的有力推动。总的来说，计算机科学和信息科学中的树是由数学中的树演变过来的，它主要用来表达各种分级结构（父亲节点和儿子节点在不同的级上)。客观世界中的大量信息以及对信息的处理都具有分级的（或层次的）特征，而这些分级特征可以有多种外观很不相同的表示形式：可以是嵌套的集合形式，可以是括号形式，可以是缩进形式，可以是编码形式，也可以是真正的树形式。例如，图1-1(a) 中的树还至少可以用另外四种形式表示。图1-1中表示的这种多样性，正反映出客观世界中分级结构的丰富程度：这些表示本身也都代表一种常见的分级结构。

在计算机科学和信息科学中树的广泛应用表现在以下几个方面。

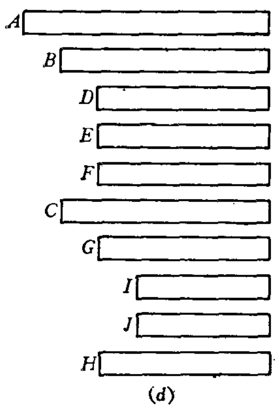
1. 树作为数据和信息结构

客观世界中有大量数据和信息。当用计算机去处理它们时，必须为之选择合适的数据结构表示——这主要说的不是它们在计



(A(B(D)(E)(F))(C(G(I)(J))(H)))

(c)



1 A
1.1 B
1.1.1 D
1.1.2 E
1.1.3 F
1.2 C
1.2.1 G
1.2.1.1 I
1.2.1.2 J
1.2.2 H

(e)

图1-1　树及其不同表示
(a) 常用表示；(b) 嵌套集合表示；(c) 括号表示；
(d) 缩进表示；(e) 编码表示。

算机内的表示，而是在处理它们的算法和程序中的表示。究竟什么叫“合适”，则取决于数据和信息本身的属性和要进行的处理的属性。在计算机科学中，树是一种动态的非线性数据结构，它非常适合于存放按字典式顺序的有序对象集合，以进行按字典式顺序的数据检索。例如，图1-2就是一棵二叉树，在它的各个节

点上按一定规则存放了一个程序语言（实际上就是本书描述算法所用的语言SL，见1.6节）中所用的所有保留字。要决定一给定字是否在这棵树上，可以从根节点开始进行比较，按字典顺序决定再选树的左子树或右子树。如果在某一步上匹配成功，那么检索就停止，否则说明这个字不在此二叉树上，这时又可以根据需要方便地把这个字加到此树上。这种树结构对于树上信息的查找(检索)、遍历、插入、删除和排序等操作都是相当灵活而高效的。已经深入研究并设计了一些细腻的算法，例如，在这种树上插入一个字时的策略和算法。这时要考虑以后再检索时的方便和效率，需要使树保持其平衡性。

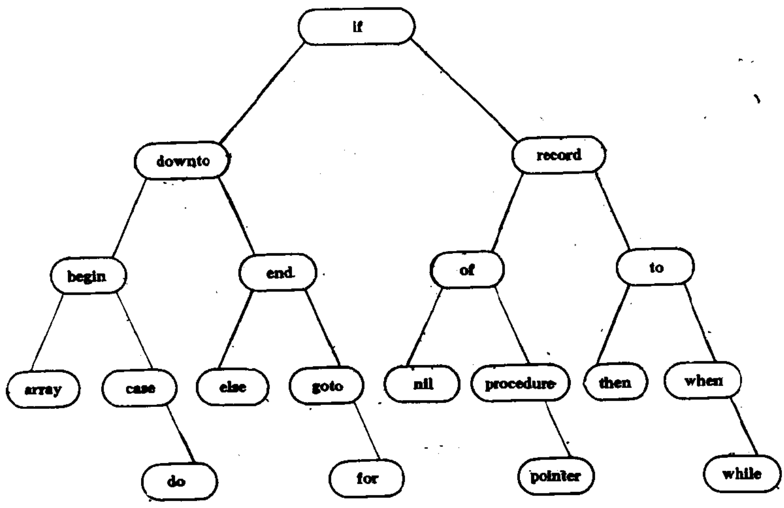


图1-2 作为数据结构的二叉树

正因为这种二叉树主要用于信息和数据结构的搜索（查找或检索)，故又称之为二叉搜索树。

2．表示代数公式

由于代数公式中多用双目运算符，因而用二叉树表示极为方便，例如图1-3就是两个代数公式的二叉树表示。

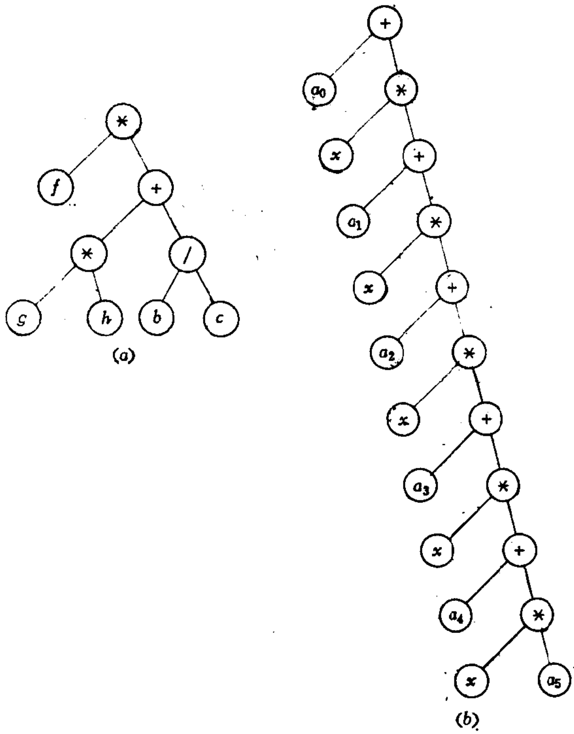


图1-3 算术表达式的树表示

(a) $f*(g*h+b/c)$; (b) $a_0+a_1x+a_2x^2+\cdots+a_5x^5$。

3. 用于表示程序结构

按照结构化的观点，一个程序系统由若干子系统组成，一个子系统由若干个模块组成，一个模块内部又由几种基本控制结构的分层嵌套组成，这是一种典型的分级结构，因而非常适合用树表示。实际上，目前有许多软件开发工具，特别是一些结构化的工具（例如，结构编辑器），都把树作为程序或软件系统的一种自然表达形式，这种表达形式概念清楚，处理方便。如图 1-4 所示，一个软件系统可表达成一棵树[2]，而程序结构自然也能用图1-5中的树形表示。