HZ BOOKS
华章教育

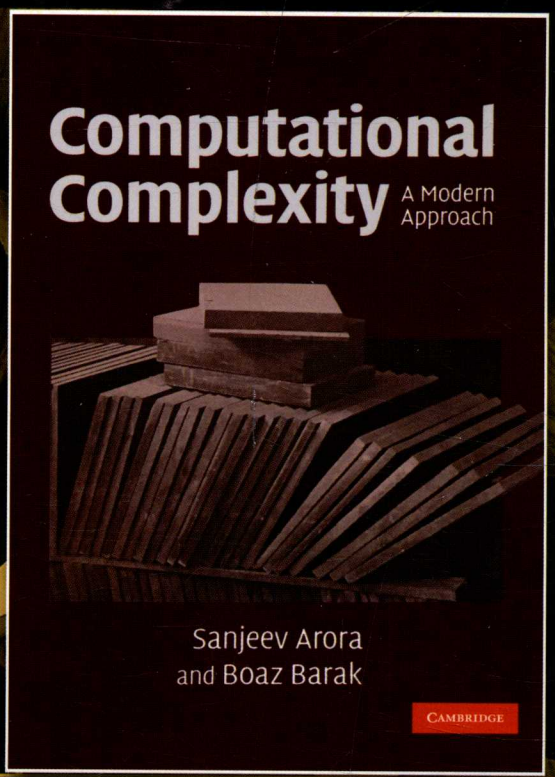
计 算 机 科 学 丛 书

CAMBRIDGE

计算复杂性
现代方法

[美] 桑杰夫·阿罗拉（Sanjeev Arora）
博阿兹·巴拉克（Boaz Barak）著
骆吉洲 译

Computational Complexity
A Modern Approach



机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

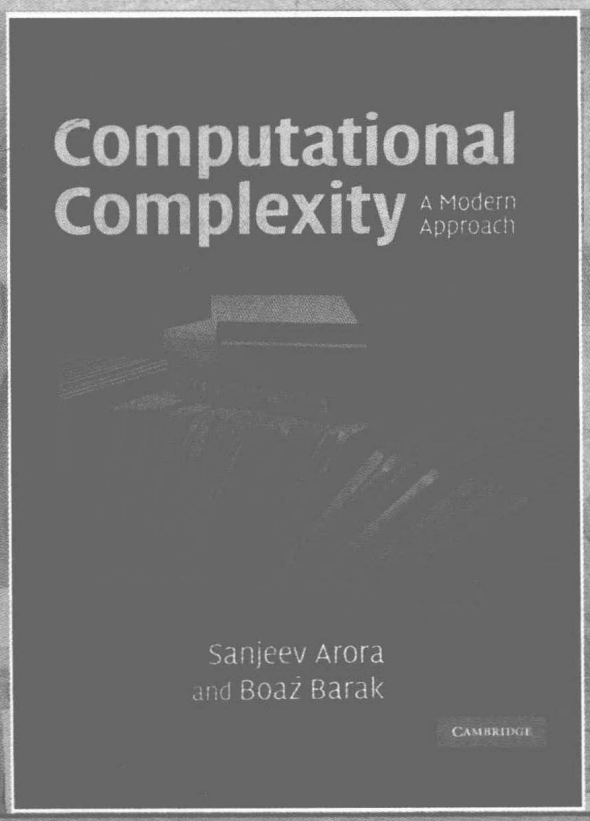
计算复杂性

现代方法

[美] 桑杰夫·阿罗拉（Sanjeev Arora）
博阿兹·巴拉克（Boaz Barak） 著
骆吉洲 译

Computational Complexity

A Modern Approach



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

计算复杂性：现代方法 /（美）阿罗拉（Arora, S.），（美）巴拉克（Barak, B.）著；骆吉洲译．
—北京：机械工业出版社，2015.11
（计算机科学丛书）
书名原文：Computational Complexity：A Modern Approach

ISBN 978-7-111-51899-0

I. 计… II. ①阿… ②巴… ③骆… III. 计算复杂性 IV. TP301.5

中国版本图书馆 CIP 数据核字（2015）第 253023 号

本书版权登记号：图字：01-2012-3791

本书系统地介绍计算复杂性理论的经典结果和近 30 年来取得的新成果，旨在帮助读者了解和掌握复杂性理论中的基本结果、思维方法、主要工具、研究前沿和待决问题。本书分三部分。第一部分（第 1 ~ 11 章）较宽泛地介绍了复杂性理论，包括复杂性理论的经典结果和一些现代专题。第二部分（第 12 ~ 16 章）讨论了各种具体计算模型上的计算复杂性下界。第三部分（第 17 ~ 23 章）主要是 1980 年以后人们在复杂性理论方面获得的进展，内容包括计数复杂性、平均复杂性、难度放大、去随机化和伪随机性、PCP 定理的证明以及自然证明。

本书内容丰富，结构灵活，语言流畅，是从事计算复杂性理论及相关领域的研究人员必不可少的参考书，非常适合作为打算进入该研究领域的研究生、博士生快速接触研究前沿的参考资料，还非常适合作为普通高校计算机科学与技术、数学专业本科生、研究生相关课程的教材，其中的高级专题还可以作为博士生相关讨论班的素材。

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：和 静　　责任校对：殷 虹
印　刷：北京市荣盛彩色印刷有限公司　　版　次：2016 年 1 月第 1 版第 1 次印刷
开　本：185mm × 260mm 1/16　　印　张：31.25
书　号：ISBN 978-7-111-51899-0　　定　价：129.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991 88361066　　投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259　　读者信箱：hzjsj@hzbook.com

文艺复兴以来，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的优势，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自 1998 年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson，McGraw-Hill，Elsevier，MIT，John Wiley & Sons，Cengage 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Andrew S. Tanenbaum，Bjarne Stroustrup，Brain W. Kernighan，Dennis Ritchie，Jim Gray，Afred V. Aho，John E. Hopcroft，Jeffrey D. Ullman，Abraham Silberschatz，William Stallings，Donald E. Knuth，John L. Hennessy，Larry L. Peterson 等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力相助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专门为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

华章网站：www. hzbook. com
电子邮件：hzjsj@hzbook. com
联系电话：(010) 88379604
联系地址：北京市西城区百万庄南街 1 号
邮政编码：100037

华章教育
华章科技图书出版中心

译者序

Computational Complexity, A Modern Approach

阿兰·图灵（Alan Turing）等人对计算的精确定义导致了现代电子计算机的诞生。如今，计算机早已融入社会管理、经济活动、工程实践、军事活动、休闲娱乐等现代生活的方方面面，各种计算机软件精彩纷呈、层出不穷。可以说，计算无时无刻地发生在我们的周围。对各种计算问题的计算过程所消耗的时间/空间等资源数量的多少进行量化，进而对各种计算问题进行分类，并研究各类计算问题之间的相互联系，研究近似求解无法精确求解的问题的难度，力争最终解决计算中最核心的问题，围绕这些任务逐渐发展和形成的理论、技术和方法，形成了理论计算机科学中的一门基础性学科——计算复杂性理论。它是为各种计算问题合理地选择算法、配置资源并进行软件开发活动的基础。

计算复杂性理论形成于20世纪五六十年代。1960年，哈特马尼斯（Hartmanis）和斯特恩斯（Stearns）在他们的开创性论文“On the computational complexity of algorithms”中引入了时间复杂性类并利用对角线方法证明了时间分层定理，由此奠定了计算复杂性理论的基础。在其后的三十年中，人们逐渐得到了各种基本复杂性类、**NP**完全理论等经典结论，并提出了计算复杂性理论中最核心的问题**P**≠**NP**。在过去三十多年中，计算复杂性理论发展迅速。自1990年以来，人们取得了大量出人意料的结果和基础性的结果，这些结果涉及的领域非常广泛，包括：经典复杂性类的概率型新定义（**IP**＝**PSPACE**和各种**PCP**定理）以及它们在近似算法中的应用，肖尔（Shor）为量子计算机设计的整数因数分解算法，对人们目前处理著名的**P**≠**NP**问题的各种方法为什么未能获得成功的理解，去随机化理论和基于计算难度的伪随机性，以及随机性提取器和扩张图等伪随机对象的优美构造。

作为计算机科学与技术相关专业的学生，全面系统地学习计算复杂性中的概念、基本结果、思维方法和重要工具并了解一些悬而未决的问题是十分必要的。本书正是适合于上述目标的一部优秀教科书。

本书作者桑杰夫·阿罗拉（Sanjeev Arora）和博阿兹·巴拉克（Boaz Barak）都是在普林斯顿大学计算机科学系一直从事复杂性理论研究的著名教授。桑杰夫·阿罗拉在概率可验证明**NP**-难问题的可近似性方面取得了基础性的研究成果。博阿兹·巴拉克在计算复杂性理论和密码学方面，特别是“非黑盒”技术方面，也取得了基础性的研究成果。本书已经逐步成为国内外计算复杂性理论课程的标准教材，其翻译和出版对国内读者学习和应用复杂性理论具有重要的意义。有幸承担该书的翻译工作，我们感到十分荣幸。

本书旨在介绍计算复杂性理论的基本概念、经典结果、近年来取得的有用的结果，帮助读者理解和掌握计算复杂性理论中的思维方法、主要工具和研究前沿。基础概念和经典结果可以帮助读者建立计算复杂性理论的知识框架，掌握复杂性理论的思维方法和证明技巧。高级专题是经典结果的有益补充和延伸，而最新的研究成果和悬而未决的问题则可以帮助读者接触计算复杂性理论研究的最前沿。此外，本书还涉及了一些学术界尚存的争论，这些深入分析和深刻见解也是本书的精髓所在。全书特别强调计算复杂性理论的各种概念的直观含义，阐述它们在何种场合下是有用的，以及为什么这些概念要这样定义。全书围绕两条主要线索进行组织：其一是人们所尝试的用于处理**P**≠**NP**问题的各种方法以

及对这些方法的局限性的阐释；其二是逐步准备证明 **PCP** 定理所需的各种素材，最终完成 **PCP** 定理的证明。这种组织使得本书内容丰富，结构灵活，可供不同层次的读者使用。

译者翻译时在深刻理解全书内容的基础上力求准确，对于发现的多处笔误和印刷错误进行了更正。在本书的翻译过程中，译者得到了多位同事、朋友和家人的支持和帮助，他们对译稿提出了很多中肯的意见和建议，使译者受益匪浅。在此一并向他们表示感谢！

限于译者水平，译文中疏漏和错误难免，敬请读者批评指正。如有任何建议，请发送邮件至 luojizhou@hit. edu. cn。

译者

2015 年 10 月

译者简介

Computational Complexity, A Modern Approach

骆吉洲，男，1975 年生，博士，副教授。2006 年 5 月毕业于哈尔滨工业大学计算机科学与技术学院软件与理论专业，获工学博士学位。1999 年、2001 年在哈尔滨工业大学数学系基础数学专业分别获得理学学士学位和理学硕士学位。现就职于哈尔滨工业大学计算机科学与技术学院海量数据计算研究中心，讲授“算法设计与分析”“数学建模”“编译原理”等课程。出版教材《算法设计与分析》一部，出版译著《图论导引》一部。近年来一直从事生物信息学、压缩数据库技术、传感器网络、算法理论等领域的研究。主持和参加多项国家自然基金、863 计划、973 项目、国防预研等项目等多项；2001 年 9 月至 2003 年 5 月参加“计算机机群并行数据库系统”的研制，该项目获得了 2004 年度国家科学技术进步二等奖。近年来发表 30 余篇论文。

计算复杂性理论在过去三十多年中发展迅速。自 1990 年以来取得的出人意料的结果和基础性的结果本身就可以写出一部书。这些结果涉及的领域非常广泛，包括：经典复杂性类的概率型新定义（**IP＝PSPACE** 和各种 **PCP** 定理）以及它们在近似算法中的应用，肖尔（Shor）为量子计算机设计的整数因数分解算法，对人们目前处理著名的 **P ≠ NP** 问题[⊖]的各种方法为什么未能获得成功的理解，去随机化理论和基于计算难度的伪随机性，以及随机性提取器和扩张图等伪随机对象的优美构造。

本书的目标就是为了在介绍复杂性理论经典结果的同时阐述近年来取得的新成果。写作本书的出发点是让它既可以作为教科书使用，也可以作为自学的参考书使用。这意味着我们在写作本书时必须兼顾广泛的读者。为实现这一目标，我们对全书进行了精心的设计。我们实际上还假设读者不具备关于计算的任何背景而且只具备附录 A 中概述的最少数学背景。我们为本书提供了一个网站 http://www.cs.princeton.edu/theory/complexity。网站上列出了相关的辅助材料，包括用本书作为教材时的详细教学计划、全书各个章节的草稿，以及涵盖相关主题的其他资源的超链接。全书始终强调各个概念在何种场合下是有用的，以及为什么这些概念要这样定义。在一些关键的定义上，我们还用一些例子进行了阐释。为了使行文流畅，我们力争尽可能少地引用参考文献。参考文献的引用有两种情况，其一是当前的结果用到了文献中的标准术语，其二是我们觉得为特定的结果提供一些历史信息将有助于阐明其动机和适用的场合。每章末尾有一个单独的注记小节，它简明扼要地讨论了更多的相关工作。当一个概念有多种定义时，我们会选择相对简单的定义；当一个结果有多种证明时，我们会选择能证得更具一般性的结论或者最优结论的证明。

全书分为三个部分。

- *第一部分：基本复杂性类。*这个部分是对复杂性理论的广泛介绍。从图灵机的定义和可计算理论的基本概念开始，这个部分涵盖了各种基本的时间复杂性类和空间复杂性类，还包含了更现代的一些专题，包括概率算法、交互式证明、密码学、量子计算机和 **PCP** 定理及其应用。
- *第二部分：具体计算模型的下界。*这个部分讨论在线路和判定树等具体计算模型上用算法求解各种计算任务所需的计算资源的下界。这些计算模型初看起来与图灵机有很大的区别，但更深入研究将得到它们与图灵机之间的有趣的相互联系。
- *第三部分：高级专题。*这个部分主要是 1980 年以后人们在复杂性理论方面获得的进展。内容包括计数复杂性、平均复杂性、难度放大、去随机化和伪随机性、**PCP** 定理的证明以及自然证明。

本书的每一章几乎都可以单独进行阅读，但是第 1 章、第 2 章和第 7 章不能跳过。正是这种设计，使得本书可以适用于下面各种不同的读者。

- *物理学家、数学家和其他科学家。*这个读者群对计算复杂性理论越来越感兴趣，他们特别感兴趣的是那些高调的研究结果，例如肖尔算法（Shor algorithm）和最

⊖ 译文用“**P ≠ NP**”来表示原文中的“**P versus NP**”。——译者注

近取得的确定型素性测试算法。这个读者群的知识储备丰富，他们可以快速通读第一部分，然后迅速进入第二部分和第三部分，也可以单独阅读各个章节并找到理解当前研究结果所需的每个知识点。

- 本身不从事计算复杂性理论研究的计算机科学家。他们既可以用本书来自学，也可以将本书作为参考书，还可以用本书来讲授本科生或研究生的计算理论或计算复杂性理论课程。
- 从事计算复杂性理论研究或者打算从事这种研究的任何人，包括教授和学生。本书讲解最新研究结果和高级专题的详细程度可以让打算从事复杂性理论和相关领域研究的读者具有充足的知识储备。

本书可以作为如下几类课程的教科书。

- 本科生的计算理论课程。很多计算机科学系都用西普赛尔（Sipser）的书［Sip96］来为本科生开设计算复杂性理论这门课。本书可以用作对西普赛尔的教材在一些更现代的专题上的补充，这些专题包括概率算法、密码学和量子计算。相比于自动机理论和可计算理论的精细划分，本科生可能会发现这些专题更能令人耳目一新。所需的数学背景是能够比较自然地阅读数学证明以及离散数学知识，这些知识通常涵盖于“离散数学”或“计算机数学”等课程中，而目前多数计算机系都已经开设了这样的课程。
- 为高年级本科生和新入学的研究生开设的计算复杂性导论课程。本书还可以用来为计算机科学专业的高年级本科生和新入学的研究生开设计算复杂性导论课程，以替代 1994 年帕帕迪米特里奥（Papadimitriou）撰写的教材［Pap94］（该书未包含很多最近的研究成果）。这门课程可以讲授第一部分的多数专题，再零星地讲授第二部分和第三部分的内容，并且假设学生具备了一定的算法知识和计算理论的知识。
- 研究生的计算复杂性课程。本书也可以作为研究生的计算复杂性课程的教材，以培养学生在复杂性理论或者算法和机器学习等相关领域开展研究的能力。这门课程可以用第一部分来复习基本知识，然后进入第二部分和第三部分的高级专题中。本书的内容多于一个学期的教学内容，网站上提供了这门课程的其他几种教学大纲。
- 研究生讨论班或高级课程。第二部分和第三部分中的各个独立章节都可以用于复杂性理论的讨论班和高级课程，比如关于去随机化、**PCP** 定理和下界的讨论班或高级课程。

本书网站为这些课程提供了几种教学计划和素材。如果你在课程中采用了本书，我们乐意了解情况并得到你的反馈。我们要求你不要在网上发布本书习题的答案，这样其他人才可以用这些习题给学生留作业或出考题。

在写作本书的过程中，我们清醒地意识到我们不得不舍弃对一些重要结果的讲述。我们希望本书对其他教材的大量引用有助于读者的进一步阅读。同时，我们还计划对本书的网站进行周期性的更新，以帮助读者了解和浏览他们感兴趣的新结果。

最重要的是，我们希望通过本书将计算复杂性中激动人心的研究结果以及它们对其他学科的深刻影响传递给读者。

让我们一起为彻底解决 $\mathbf{P} \neq \mathbf{NP}$ 问题而努力吧！

我们对复杂性理论的理解是在与同行交流的过程中逐步形成的。我们从他们身上学到了太多的东西，要感谢的人也远不止下面提到的这些人。Boaz 首先要感谢两位导师 Oded Goldreich 和 Avi Wigderson，是他们把他引入了理论计算机科学的世界并不断影响他在这一领域的思想。

我们感谢 Luca Trevisan，他（从 8 年前开始!）对本书的写作提供了持续不断的鼓励，并为第一版草稿中若干章节的撰写提供了大量帮助。一些同行毅然地同意并审阅了本书部分章节的早期草稿，他们是：Scott Aaronson，Noga Alon，Paul Beame，Irit Dinur，Venkatesan Guruswami，Jonathan Katz，Valentine Kavanets，Subhash Khot，Ji ří Matoušek，Klaus Meer，Or Meir，Moni Naor，Alexandre Pinto，Alexander Razborov，Oded Regev，Omer Reingold，Rosen Shaltiel，Madhu Sudan，Amnon Ta-Shma，Iannis Tourlakis，Chris Umans，Salil Vadhan，Dieter van Melkebeek，Umesh Vazirani 和 Joachim von zur Gathen. 特别感谢 Jiří、Or、Alexandre、Dieter 和 Iannis，他们对本书的很多章节给出了具体而有用的评述。

很多人指出了本书的拼写错误和缺陷，他们给出的评述帮助我们改进了文字质量，还有一些人回答了我们在特定证明中提出的问题或者给出了相应的参考文献。在此，一并对他们表示感谢，他们是：Emre Akbas，Eric Allender，Djangir Babayev，Miroslav Balaz，Arnold Beckmann，Ido Ben-Eliezer，Siddharth Bhaskar，Goutam Biswas，Shreeshankar Bodas，Josh Bronson，Arkadev Chattopadhyay，Bernard Chazelle，Maurice Cochand，Nathan Collins，Tim Crabtree，Morten Dahl，Ronald de Wolf，Scott Diehl，Dave Dody，Alex Fabrikant，Michael Fairbank，Joan Feigenbaum，Lance Fortnow，Matthew Franklin，Rong Ge，Ali Ghodsi，Parikshit Gopanlan，Vipul Goyal，Stephen Harris，Johan Håstad，Andre Hernich，Yaron Hirsch，Thomas Holenstein，Xiu Huichao，Moukarram Kabbash，Bart Kastermans，Joe Kilian，Tomer Kotek，Michal Koucy，Sebastian Kuhnert，Katrina LaCurts，Chang-Wook Lee，James Lee，John Lenz，Meena Mahajan，Mohammad Mahmoody-Ghidary，Shohei Matsuura，Mauro Mazzieri，John McCullough，Eric Miles，Shira Mitchell，Mohsen Momeni，Kamesh Munagala，Rolf Neidermeier，Robert Nowotniak，Taktin Oey，Toni Pitassi，Emily Pitler，Aaron Potechin，Manoj Prabhakaran，Yuri Pritykin，Anup Rao，Saikiran Rapaka，Nicola Rebagliati，Johan Richter，Ron Rivest，Sushant Sachdeva，Mohammad Sadeq Dousti，Rahul Santhanam，Cem Say，Robert Schweiker，Thomas Schwentick，Joel Seiferas，Jonah Sherman，Amir Shpilka，Yael Snir，Nikhil Srivastava，Thomas Starbird，Jukka Suomela，Elad Tsur，Leslie Valiant，Vijay Vazirani，Suresh Venkatasubramanisn，Justin Vincent-Foglesong，Jirka Vomlel，Daniel Wichs，Avi Wigderson，Maier Willard，Roger Wolff，Jureg Wullschleger，Rui Xue，Jon Yard，Henry Yuen，Wu Zhanbin 和 Yi Zhang. 谢谢你们!

毫无疑问，上述列表仍有可能遗漏了这些年向我们的创作提供过帮助的人，我们对那

些被遗漏的人表示感谢和歉意。

本书用 LATEX 进行排版，为此特别感谢 Donald Knuth 和 Leslie Lamport。Stephen Boyd 和 Lieven Vanderberghe 爽快地向我们提供了他们写作《凸优化》一书时使用的 LATEX宏定义。

最重要的是，感谢我们的家人——Silvia、Nia、Rohan Arora 和 Ravit、Alma Barak。

Sanjeev 要感谢父亲，父亲是他创作本书的源泉。

如果一个科学分支还能找到大量的研究问题，则这个分支还活着；研究问题的缺失则预示着这个科学分支的消亡或者独立发展的终结。

——戴维·希尔伯特（David Hilbert），1900

我演讲的主题或许可以直接用两个简单的问题来揭示。第一个问题是，乘法比加法更难吗？而第二个问题则是，为什么？……我（想）要证明不存在乘法的算法在计算上同加法的算法一样简单，这就证明了乘法计算中确实存在某种绊脚石。

——阿兰·科巴姆（Alan Cobham），1964

几千年来，人们在账目管理和天文学中不断地进行着各种计算，因此计算的概念以这种形式一直存在着。例如，利用计算可以求解下面的任务。

- 给定两个整数，计算它们的乘积。
- 给定 n 个变量上 n 个方程构成的方程组，找出它的一个解（如果解存在的话）。
- 给定一个熟人列表和这些人中彼此不能和睦相处的一些人员对，找出你在宴会中打算邀请的最大熟人子集使得他们彼此都能够和睦相处。

在历史长河中，人们总结得出：在概念上，计算就是在给定的输入上用有限个步骤得到输出的过程。他们认为，“计算”就是“在演草纸上根据一定规则写出一些数字的人”。

20 世纪中叶，科学上的一项重要突破就是对“计算”的精确定义。根据这个定义，人们马上清楚了计算其实就存在于各种物理系统和数学系统中——图灵机（Turing machine），λ 演算（Lambda calculus），细胞自动机（cellular automata），指针机（pointer machine），弹跳桌球（bouncing billiards ball），康韦生命游戏（Conway's game of life），等等。出人意料的是，这些形式的计算都是等价的，也就是说，其中一个模型能够实现的所有计算也能在其他模型上完成（参见第 1 章）。在这种认识的基础上，人们马上发明了能够执行所有程序的硬件，这就是标准的通用电子计算机。在接下来的几十年中，计算机迅速被社会接纳，这使得计算融入了现代生活的方方面面，也使得计算问题在设计、计划、工程、科学发现等人类活动中变得越来越重要，于是计算机算法（亦即，求解计算问题的各种方法）变得无处不在。

然而，计算并不“仅仅”是一种用于实践的工具，它也是一个主要的科学概念。事实上，科学家现在将许多自然现象都视为一种计算过程，这些计算过程实际上是细胞自动机的推广。例如，对生物繁衍过程的理解曾经导致了自体复制计算机的发现。再比如，物理学家薛定谔（Schroedinger）曾在他的书［Sch44］中预言细胞中肯定存在类似于 DNA 的物质，后来沃森（Watson）和克里克（Crick）发现了这种物质，克里克坦承他们的研究正是受到了薛定谔的工作的启发。如今，各种计算模型已经成为生物学和神经科学中许多领域研究的基础。电子电动力学（QED）等几种物理理论的本征刻画非常类似于计算的定义，这种现象甚至还促使一些科学家相信整个宇宙就是一台超级计算机（参见 Lloyd［Llo06］）。更有趣的是，这样的物理理论在过去的十年中已经被用于设计量子计算机（参见第 10 章）。

可计算性与计算复杂性

研究者在成功地给出计算的概念之后，开始研究什么样的问题是可计算的。他们证明了几个有趣的问题本质上都不是可计算的，也就是说，没有计算机能够求解这些问题而不在任何输入上陷入无限循环（即不停机）。可计算性理论是一个优美的专题，但本书不会用太多篇幅讨论它。我们仅在第 1 章中简要地讨论可计算性理论，感兴趣的读者可以参阅标准的教科书 [Sip96，HMU01，Koz97，Rog87]。事实上，本书的焦点是计算复杂性理论，它关注计算的效率，也就是量化地研究求解给定计算任务所需的计算资源的数量。下面将会对计算效率非形式化地进行量化。然后，再讨论复杂性研究中有关计算效率的几个问题。

计算效率的量化

我们用前面提到的三个计算任务来阐释计算效率的含义。首先，考虑两个整数的乘法。我们可以用两种不同的方法（或算法）来完成这项任务。第一种方法是累加算法；也就是说，为了计算 $a \cdot b$，只需用 $b-1$ 次加法将 b 个 a 进行累加。第二种方法是图 I-1 所示的小学列式算法。虽然累加算法比小学列式算法简单，但我们总觉得后者比前者更好。事实上，小学列式算法的效率更高。例如，计算 577 乘以 423 时，累加算法需要计算 422 次加法；但小学列式算法却只需将其中一个数分别与 3 个一位数相乘，再计算 3 次加法。

```
      5 7 7
      4 2 3
  ------------
    1 7 3 1
  1 1 5 4
2 3 0 8
  ------------
2 4 4 0 7 1
```

图 I-1　用 577 · 423 展示乘法的小学列式算法

算法效率的量化就是研究算法所执行的基本操作个数随着输入规模的增长是如何变化的。在整数乘法中，单个数位相加或相乘就是基本操作（在其他场合中，除法也可以是基本操作），而两个因数中的数位个数就是输入规模。在计算两个 n 位整数的乘法时（也就是说，两个因数都大于 10^{n-1} 且小于 10^n），小学列式算法执行的基本操作数不超过 $2n^2$，而累加算法执行的基本操作数至少是 $n10^{n-1}$。经过这样的分析，两种算法之间的巨大差异就显而易见了。即使计算两个 11 位整数的乘法，执行小学列式算法的便携式计算器也会快于执行累加算法的超级计算机。对于稍大一些的整数，小学五年级学生用笔和纸执行列式算法也会胜于执行累加算法的超级计算机。由此可见，算法的效率明显比运行算法所用的技术要重要得多。

更令人意外的是，借助傅里叶变换可以设计出更快的乘法算法。这个算法 40 年前才被人们发现，在这个算法中，两个 n 位整数的乘法仅用 $cn\log n\log\log n$ 个操作就可以完成，其中 c 是独立于 n 的绝对常数，参见第 16 章。我们将这样的算法称为 $O(n\log n\log\log n)$ 步的算法，参见后面的记号约定。随着 n 的增大，该算法执行的基本操作数将远小于 n^2。

对于线性方程组的求解，经典的高斯消元法用 $O(n^3)$ 个基本的算术操作就可以求解 n 个变量上 n 个方程构成的方程组（虽然这一算法用高斯的名字命名，但早在公元 1 世纪中国数学家就已经掌握了这种算法的某种形式）。20 世纪 60 年代末，斯特拉森（Strassen）找到了更加高效的算法，该算法大约只需要执行 $O(n^{2.81})$ 个操作就能求解这个问题。目

前，求解这一问题的最佳算法需要执行 $O(n^{2.376})$ 个操作，参见第 16 章。

宴会问题也有一段有趣的历史。同乘法的情况一样，宴会问题也存在显而易见的简单低效算法——从大到小地依次枚举 n 个人的每个子集，直到找到一个子集使得其中不含任何两个无法和睦相处的人。这个算法需要执行的计算步骤数可能与 n 个人的所有子集数一样多，亦即 2^n。这使得该算法根本无法用于实践，因为如果某人用这个算法来安排一个 70 人参加的宴会，即使她用超级计算机来进行处理，也需要提前一千年开始筹备。但出乎意料的是，人们至今仍没有为宴会问题找到效率显著更优的算法。事实上，正如第 2 章中我们将会看到的那样，我们有理由怀疑宴会问题不存在高效的算法，因为我们可以证明它等价于独立集这个计算问题，而独立集问题以及其他成千上万个计算问题都是 **NP**-完全问题。著名的 **P** ≠ **NP** 问题（参见第 2 章）是问：有没有哪个 **NP**-完全问题存在高效的算法？

证明高效算法的不存在性

我们已经看到，某些计算任务存在非平凡的算法使得其效率比几千年来人们一直使用的算法更高。一件特别有意义的事情是：证明某些组合任务的当前算法是最佳的。也就是说，这些计算任务不存在更有效的算法。例如，我们可以证明整数乘法的 $O(n\log n\log\log n)$ 步算法无法进一步改进，这就说明乘法在本质上确实比加法更难，因为加法存在 $O(n)$ 步算法。再比如，我们还可以证明，没有算法能用少于 $2^{n/10}$ 个计算步骤来求解宴会问题。证明这样的结论是计算复杂性的一个核心任务。

如何才能证明这种不存在性呢？求解计算任务的算法可能有无穷个！因此，我们只能用数学手段证明其中的每个算法都比已知的算法更低效。这种方法之所以可行，是因为计算本身也是一个精确的数学概念。事实上，一旦这样一个结果被证明，则它必然吻合于数学上的某个不可能性结果，例如，几何中其他公理无法推导出欧几里得平行公理、尺规作图无法三等分一个角等。这些结论都是数学上最有价值、最可靠和最出人意料的结果。

在复杂性理论中，我们很少能证明这种不存在性结果。但是，在能力弱于一般计算机的计算模型上，我们确实已经证得了一些重要的不存在性结果。本书第二部分将讨论这些结果。由于在一般的计算机上我们还缺少这样的好结果，因此复杂性理论在一般计算机上获得的重要结果指的是在不同复杂性问题之间建立的相互联系。人们在这方面获得了很多漂亮的结果，本书介绍了大量这样的结果。

关于计算效率的几个有趣问题

现在，我们概述关于计算复杂性的几个重要问题，本书后续章节将详细地讨论这些问题。附录 A 给出了相关的数学背景。

1. 生命科学、社会科学和运筹学等学科中的很多计算任务都通过搜索海量的解空间来找出问题的解。比如，前面已经提到的线性方程组的求解和宴会问题中找出最大的受邀者集合都属于这种情况。这种搜索通常称为穷举搜索，因为搜索过程穷举了所有的可能。这种穷举搜索能替换为更有效的算法吗？

正如我们将在第 2 章所见，这本质上就是著名的 **P** ≠ **NP** 问题，它是计算复杂性理论的核心问题。很多有趣的搜索问题都是 **NP**-完全问题，这意味着，如果 **P**≠**NP** 这个著名的猜想是正确的，则这样的搜索问题将不存在高效的算法；亦即，这种搜索问题本质上是难解的。

2. 用随机性（即硬币投掷）能加快算法的计算速度吗？

第 7 章介绍了随机计算，并为一些特定的计算任务给出了高效的概率型算法。但是，第 19 章和第 20 章给出了人们最近得到的一个出人意料的结果，该结果提供了很强的证据来表明随机性对提高计算效率而言作用非常有限，因为任意概率型算法都可以替换为一个效率相当的确定型算法（即不用投掷硬币的算法）。

3. 如果允许算法在小部分输入上出错，或者只要求算法求得问题的近似解，那么难解的问题会变得容易一些吗？

平均复杂性和近似算法的研究出现在本书的第 11 章、第 18 章、第 19 章和第 22 章。这几章还建立了上述问题、随机性的效能、数学证明的不同概念和纠错码理论之间的优美的相互联系。

4. 计算上的难解问题对实践有什么帮助呢？例如，我们能借助这些难解问题构造出牢不可破的密码协议吗（至少相对于大家认可的敌手而言）？

正如第 9 章所讲，安全的数字密码与 **P** ≠ **NP** 问题（参见第 2 章）和平均复杂性（参见第 18 章）密不可分。

5. 我们能利用物质的违背直觉的量子力学性质建造出更快的计算机吗？

第 10 章将介绍量子计算机这一备受瞩目的概念，它利用量子力学加速某些计算。彼得·肖尔（Peter Shor）已经证明，只要量子计算机被建造出来，则它能够高效地完成整数的因数分解（继而，现今的很多密码都将被攻破）。但是，要建造出量子计算机，目前还存在着很多令人生畏的障碍。

6. 只有人才能证明数学定理吗？换句话说，数学证明能被自动生成吗？能在不完整阅读数学证明的情况下验证数学证明的真伪吗？证明者和验证者通过对话来完成的交互式证明比标准的“静态”数学证明更有效力吗？

证明是数学上的核心概念。事实证明，它也是计算复杂性的核心概念。而且，计算复杂性已经对数学证明的含义赋予了新的解释。数学证明能否自动生成将取决于 **P** ≠ **NP** 问题的答案（参见第 2 章）。第 11 章研究概率可验证明（Probabilistic Checkable Proof）。概率可验证明是一种健壮的数学证明。要查验这种证明的真伪，只需概率地选取证明中的少数几个位置进行查验即可。相比之下，传统的证明则需要逐行阅读才能查验其真伪。类似地，第 8 章介绍了交互式证明的概念，并用它得出了一些出人意料的结果。最后，第 15 章研究了证明复杂性。它是复杂性的一个子领域，研究各种命题的最小证明长度。

历经近 40 年的发展，复杂性理论仍是一门年轻的科学，许多重要结果的发现还不到 20 年。上述这些问题还没有被完全解决。一个令人意外的转折是，复杂性理论被用于某些数学定理的证明中：它们提供的证据表明计算复杂性中某些问题是难解的，参见第 23 章。

我们引用希尔伯特 1900 年演讲中的另一些话作为结束[⊖]。

> （这种）不可能性的证明古人早已实现……［而］在后来的数学中，关于某些解的不可能性的问题发挥了重要作用……
>
> 在其他科学中，人们也经常遇到一些老问题，通过不可能性的证明，这些问

⊖ 引文的第二段和第三段与希尔伯特演讲中的顺序是颠倒的。——译者注

题被一种对科学来说是最满意、最有用的方式解决了……在构造永动机的努力失败以后，科学家在这种机器不可能存在的情况下研究了自然力之间必须遵循的关系；而这个反问题导致了能量守恒定律的发现……

也许正是这一值得注意的事实，加上其他的哲学因素，给人们以这样的信念……每个明确的数学问题都必然能被毋庸置疑地精确求解，或者是成功地对所给问题做出了回答，或者是证明了该问题解的不可能性，从而表明解答所给问题的一切努力都肯定要归于失败……这种信念……对于数学工作者是一种巨大的鼓舞。在我们中间，常常听到这样的呼唤：这里有一个数学问题，快找出它的答案！你能通过纯粹的数学推理找到答案，因为数学中不存在不可知的东西。

目 录

Computational Complexity, A Modern Approach