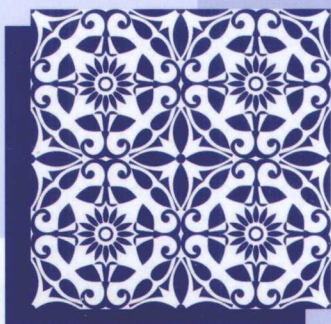


# 新编 计算机科学概论

刘芝 蔡敏 等编著



**C**omputer Science  
An Overview



机械工业出版社  
China Machine Press

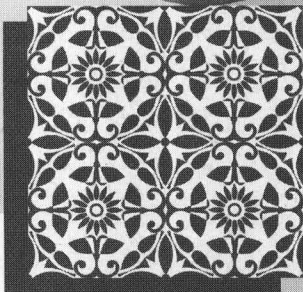
013027783

重点大学计算机教材

TP3-43  
608

# 新编 计算机科学概论

刘艺 蔡敏 等编著



Computer Science  
An Overview

TP3-43  
608



北航 C1637089



机械工业出版社  
China Machine Press

887750810

图书在版编目 ( CIP ) 数据

新编计算机科学概论 / 刘艺等编著. —北京: 机械工业出版社, 2013.3  
(重点大学计算机教材)

ISBN 978-7-111-40695-2

I. 新… II. 刘… III. 计算机科学 - 高等学校 - 教材 IV. TP3

中国版本图书馆 CIP 数据核字 (2013) 第 022390 号

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

本书以计算机科学与技术的发展历程为主线, 从基本概念入手, 分别沿硬件和软件两条线逐步深入。全书覆盖了计算机专业的所有专业课程知识领域, 并根据非计算机专业需要着重介绍基本原理, 为学生建立正确的基本概念。本书内容全面、完整, 注重学以致用, 并介绍了本领域的最新发展。

本书既适合作为高等院校计算机及相关专业基础课的教材, 又可作为计算机基础入门读物和广大计算机爱好者的参考资料。



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: 余洁

三河市杨庄长鸣印刷装订厂印刷

2013 年 4 月第 1 版第 1 次印刷

185mm × 260mm • 17 印张

标准书号: ISBN 978-7-111-40695-2

定价: 33.00 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: hzsj@hzbook.com



# 前 言

欢迎学习计算机科学概论课程。这将是一次美妙和激动人心的探索，可能为你今后从事的充满挑战和令人兴奋的 IT 职业奠定基础。众所周知，计算机在我们的日常生活中扮演了一个重要的角色，而且在未来也将一样。

计算机科学是一个充满了挑战和发展机遇的年轻学科，而计算机科学概论课程则是这门学科的重要基础。随着计算机在各行各业的广泛应用，很多非计算机专业也把计算机科学概论课程列为公共基础课之一。

既然是作为基础课的教材，那么本书所假定的读者可以既没有计算机应用经验，也不具备太多的计算机知识。即使是对计算机一无所知的人，也能通过学习本书而获取大量计算机科学的基本知识；如果读者已有一定的计算机应用经验那就更好，他将在本书中发现很多有用的理论知识，以提高自己的专业水平。

作为 IT 专业基础课教材，本书力求做到：知识体系完整，覆盖面广，内容翔实，文风严谨，深入浅出，并符合国内高校的教学实践需要。同时，本书紧跟时代，还介绍了一些计算机科学的最新发展和应用，如移动操作系统、云计算、物联网等。

最重要的一点，本教材在吸收国内同类课程教学改革经验的基础上，遵循 CC2005 课程体系，大胆参考了 ACM 推荐的 CS0、CS1、CS2 课程设计，与国际 IT 教材接轨，使其从广度上覆盖了计算机科学的主要领域。

## 本书结构

本书是为计算机概论 / 导论课程编写的。该课程将为计算机及其相关专业的本科生勾画出计算机科学体系的框架，为有志于 IT 行业的学生奠定计算机科学知识的基础，架设进一步深入学习专业理论的桥梁。本书强调的是概念而不是数学模型和技术细节，并通过大量的图片、表格和演示增强读者对阅读的理解和知识的掌握，通过范例讲解概念和相关的模型，通过关键字索引、小结和习题帮助读者掌握自己的学习效果。

本书分为 11 章，各章的主要内容如下：

- **第 0 章 绪论** 本章从整体上介绍计算机科学的概念和内涵。概要介绍计算机的发展历史，并分析计算机对现代社会的影响。
- **第 1 章 数据的表示与编码** 数据是计算机处理的基本元素，本章讨论数据的表示和运算，以及不同数据类型的编码和存储。
- **第 2 章 计算机体系结构** 本章重点讨论计算机的体系结构。先从系统的角度介绍了计算机系统的层次结构，然后从体系结构的角度分析了系统硬件的结构并重点讨论了处理器结构，最后介绍了计算机软件系统的结构。
- **第 3 章 计算机硬件组成** 本章讨论计算机硬件系统组成，讲述计算机主要硬件组成部分的原理、结构、实现和发展趋势，并讨论当前主流计算机的常用硬件。
- **第 4 章 数据结构与算法** 本章讲解了数据结构和算法的概念，以及算法的描述方法和基本结构，并讨论了几种常见的数据结构和常用算法。

- **第5章 操作系统** 本章讨论计算机的操作系统，阐述操作系统的功能、基本组成和运行原理。
- **第6章 程序设计语言** 本章讲述了计算机程序设计语言的发展和分类，介绍程序设计语言的类型和基本知识，讨论了程序编译生成的过程和原理。
- **第7章 数据库技术与应用** 本章讨论运用数据库方式管理大量数据的方法。着重阐述数据库方法的基本概念、原理，结合关系数据库介绍了 SQL 基本情况，介绍了数据挖掘技术和数据仓库的基本原理。
- **第8章 软件工程** 本章讨论规模化软件开发的方法。包括软件过程、需求分析、软件工程方法、软件质量管理、软件项目管理等内容。
- **第9章 计算机网络** 本章从计算机网络的形成与发展开始，依次讲述了计算机网络体系结构、组建网络和 Internet 基本技术及应用。
- **第10章 信息系统安全** 本章从计算机网络信息系统的安全威胁入手，依次讲述了信息安全的核心——数据加密、最常见的安全威胁——计算机病毒、最常见的网络信息安全技术——防火墙和入侵检测。

尽管本书包含了以上章节内容，但实际的教学进度和授课内容可以灵活确定，因为这要取决于课堂教学的安排或读者实际技能及对所讨论问题的熟悉程度。教学时数建议安排在 40~60 课时之间。

## 本书特色

本书特别有助于入门者学习。本书在讲述理论知识的同时，注重理论发展的历史背景、思考方法；在注重知识的系统性、完整性的同时，关注知识发展的动态性、渐进性；在关注知识传授有效性的同时，重点培养和引导读者进行思考、分析、探索问题。

### 1. 概念和知识面

贯穿本书，我们始终强调概念要比数学模型更重要，对概念的理解必然左右对模型的理解。同时，我们还特别注意开阅读者的知识面，使读者能够站在比较宽的层面上了解计算机科学。

### 2. 图文并茂

本书图文并茂，全书有大量精心设计的图片，这些图片可以增进读者对文字的理解。

### 3. 教辅

本书为教师提供 PPT 课件以及相应的教辅材料（如习题参考答案）。教师可以在机械工业出版社华章分社的网址上找到并下载它们：<http://www.hzbook.com/>。

### 4. 小结和习题

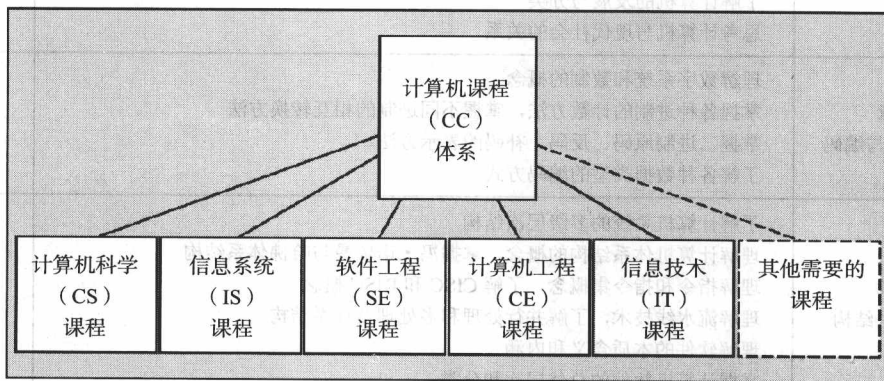
每一章的结尾都包括本章小结和本章习题。本章小结包括了对本章所有关键内容和知识点的简明概括，是复习时的参考。本章习题包括了复习题和练习题。

- 复习题：测试本章中所有要点和概念，帮助学生复习巩固重点内容。
- 练习题：通过课后练习题，检查学生能否运用掌握的概念和知识独立思考、解决问题。

## CC2005 课程体系

从 1990 年开始，美国电气电子工程师学会计算机分会（the Computer Society of the Institute for Electrical and Electronic Engineers，简称 IEEE-CS）和计算机协会（Association for Computing Machinery，简称 ACM）就着手开发新的本科生计算机课程体系。1991 年他们联合推出了 Computing

Curricula 1991 (简称 CC1991), 当时仅限于 Computer Science 和 Computer Engineering 两个专业的课程。1998 年秋季开始, IEEE-CS 和 ACM 联合投入新的力量更新该课程体系, 并在 2001 年开发出 Computing Curricula 2001 (简称 CC2001), 并将该计算机课程体系扩大到 Computer Science、Computer Engineering、Software Engineering、Information Systems 等多个专业。在 CC2001 的实施过程中, 专家们发现计算机课程所涉及的学科专业和教学范围正在不断扩大, 而且在内容和教学方面的变化也日新月异。IEEE-CS 和 ACM 意识到 10 年一次的 Computing Curricula 修订已经难以满足要求, 于是他们联合国际信息处理联合会 (International Federation for Information Processing, 简称 IFIP)、英国计算机协会 (British Computer Society, 简称 BCS) 等更多的组织开发了 Computing Curricula 2005 (简称 CC2005), 使之成为开放的、可扩充的、适合多专业的、整合了计算机教学相关原则体系观点的课程体系指南。其结构参见下图。



为了进一步反映当代计算机科学技术的发展水平, 与国际主流计算机教育思想接轨, 我们通过多年来对 IEEE-CS 和 ACM 的 Computing Curricula 课程体系的跟踪研究编写了本教材。在教材的编写中, 借鉴了 CC2005 课程体系的最新研究成果, 同时吸取了国外同类教材的优秀经验, 目的是进一步推动教材和课程改革, 培养有竞争力的人才。

## 致谢

本书是在作者多年科研和教学基础上编写的, 主要参考了作者已发表的文章和著作以及教学中积累的资料。书中还用到了其他中外文教材、资料, 由于无法在此一一列举, 现谨对这些教材和资料的作者表示衷心的感谢。

参与本教材编写的老师有南京航空航天大学的吴英、海军 91329 部队的李宇; 海军指挥学院的陈庆元、秦华、黄冬梅; 中电熊猫集团的蔡青; 还有洪蕾、张颖等。本教材最后由江苏警官学院的蔡敏老师统稿。

一本书的出版离不开许多人的支持, 尤其是这本书。为此感谢我们的家人和朋友。

由于作者水平有限, 书中难免有疏漏和不妥之处, 恳请各位专家、同仁和读者不吝赐教, 并在此表示衷心感谢!

2012 年 12 月 25 日于南京

# 教学建议

教学章节	教学要求	课时
第 0 章 绪论	<p>理解计算机科学的内涵</p> <p>了解计算机科学涵盖的领域和发展历程</p> <p>了解计算机的发展与分类</p> <p>思考计算机与现代社会的关系</p>	2
第 1 章 数据的表示与编码	<p>理解数字系统和数制的概念</p> <p>掌握各种进制的计数方法, 掌握不同进制的相互转换方法</p> <p>掌握二进制原码、反码、补码的表示方法</p> <p>了解各种数据类型的编码方式</p>	4~6
第 2 章 计算机体系结构	<p>了解计算机系统的多级层次结构</p> <p>理解计算机体系结构的概念, 掌握冯·诺依曼与哈佛体系结构</p> <p>理解指令和指令集概念, 了解 CISC 和 RISC 概念</p> <p>理解流水线技术, 了解并行处理和多处理器体系结构</p> <p>理解软件的本质含义和内涵</p> <p>掌握计算机软件的总体层次和分类</p>	4~6
第 3 章 计算机硬件组成	<p>了解计算机系统的硬件组成</p> <p>了解 CPU 的原理和结构</p> <p>理解存储器系统原理、组成</p> <p>理解总线与接口的原理, 了解其基本类型</p> <p>了解输入输出设备的原理和分类</p>	4~6
第 4 章 数据结构与算法	<p>理解数据结构的概念和结构</p> <p>理解算法的概念和基本特性, 了解算法的度量方法</p> <p>理解线性数据结构, 掌握栈和队列、串和数组等典型线性数据结构</p> <p>了解非线性的数据结构, 了解树、二叉树和图等典型非线性数据结构</p> <p>理解排序、查找的概念, 了解排序、查找的典型算法并能够比较它们</p> <p>理解递归的概念, 了解递归的实际应用</p>	4~6
第 5 章 操作系统	<p>理解操作系统的基本概念、功能</p> <p>理解操作系统的组成要素, 了解操作系统的分类</p> <p>理解进程、死锁等概念</p> <p>理解存储管理的内涵, 建立虚拟存储的概念</p> <p>理解文件系统的作用, 理解虚拟文件系统</p> <p>理解设备管理的概念和缓冲技术</p> <p>了解典型的操作系统</p>	4~6
第 6 章 程序设计语言	<p>了解计算机语言的发展历程</p> <p>理解机器语言、汇编语言、高级语言的概念及其差异和特点</p> <p>建立虚拟机的概念, 了解高级语言</p> <p>了解程序设计语言和程序语句的基本类型</p> <p>理解程序编译生成的过程和原理</p>	4~6

(续)

教学章节	教学要求	课时
第 7 章 数据库技术与应用	理解数据库的基本概念, 了解数据库系统的组成 理解数据模型和数据库模型 了解数据字典 掌握关系数据库的模型、结构和约束 理解 SQL, 学会使用基本的 SQL 理解数据仓库的概念, 了解数据挖掘的内涵、基本过程和应用	4~6
第 8 章 软件工程	理解软件工程的内涵 掌握软件开发过程 掌握软件需求分析的任务和方法 了解典型的软件工程方法 了解软件质量的含义和软件质量管理方法 了解软件项目管理和团队等概念	4~6
第 9 章 计算机网络	掌握计算机网络的内涵、功能和分类 了解 OSI 参考模型和 TCP/IP 模型 了解网卡、中继器、网桥、路由器、交换机和网关等典型网络设备 了解 TCP/IP 等 Internet 相关协议和典型应用 了解云计算与物联网的概念与应用	4~6
第 10 章 信息系统安全	了解信息安全的问题根源和信息系统面临的安全威胁 了解网络攻击的常见形式 掌握计算机信息系统安全的目标、服务、安全机制和策略 理解密钥, 了解典型密码算法和加解密技术 了解计算机病毒的破坏行为、特征和分类 了解计算机病毒的结构模式和工作机理 掌握计算机病毒防治的一般方法 理解防火墙的基本概念和工作原理 了解入侵检测系统的工作原理、分类和结构及其发展趋势	2~4
总课时		40~60



# 目 录

前言	1.3.1 字符的编码 .....	34
教学建议	1.3.2 静态图像的编码 .....	37
	1.3.3 动态数据的编码 .....	39
第 0 章 绪论 .....	本章小结 .....	40
0.1 什么是计算机科学 .....	本章习题 .....	41
0.1.1 计算机科学的提出 .....	第 2 章 计算机体系结构 .....	44
0.1.2 计算机科学领域 .....	2.1 计算机系统的多级层次结构 .....	44
0.1.3 计算机与计算机模型 .....	2.2 计算机体系结构概述 .....	47
0.1.4 计算机科学中的经典问题 .....	2.2.1 计算机体系结构的基本概念 .....	47
0.2 计算机的历史 .....	2.2.2 相关概念 .....	49
0.2.1 计算机前传 .....	2.3 冯·诺依曼结构和哈佛结构 .....	50
0.2.2 电子计算机 .....	2.3.1 冯·诺依曼理论 .....	51
0.2.3 计算机的未来 .....	2.3.2 冯·诺依曼体系结构 .....	51
0.3 计算机与社会 .....	2.3.3 冯·诺依曼结构的演变 .....	53
0.3.1 计算机与生活 .....	2.3.4 哈佛结构 .....	54
0.3.2 有关计算机的伦理 .....	2.4 处理器体系结构 .....	54
0.3.3 与计算机有关的犯罪 .....	2.4.1 指令系统 .....	54
0.4 我国计算机科学的发展 .....	2.4.2 CISC 结构 .....	56
本章小结 .....	2.4.3 RISC 结构 .....	57
本章习题 .....	2.4.4 并行处理与流水线技术 .....	59
第 1 章 数据的表示与编码 .....	2.5 计算机软件系统 .....	62
1.1 数和数制 .....	2.5.1 软件的含义 .....	62
1.1.1 数字系统 .....	2.5.2 软件系统的组织 .....	62
1.1.2 计数与进制 .....	2.5.3 软件的社会形态 .....	64
1.1.3 二进制和位 .....	本章小结 .....	68
1.1.4 八进制和十六进制 .....	本章习题 .....	69
1.1.5 不同进制间的相互转换 .....	第 3 章 计算机硬件组成 .....	71
1.2 数值的表示与运算 .....	3.1 概述 .....	71
1.2.1 整数的表示 .....	3.2 处理器 .....	71
1.2.2 实数的表示 .....	3.2.1 运算器 .....	72
1.2.3 位的算术运算 .....	3.2.2 控制器 .....	73
1.3 非数值信息的编码 .....	3.2.3 摩尔定律与处理器的发展 .....	74

3.3 存储器 .....	75	5.3.3 顺序文件 .....	130
3.3.1 计算机的多级存储系统 .....	76	5.3.4 索引文件 .....	130
3.3.2 主存储器 and 高速缓存 .....	77	5.3.5 散列文件 .....	130
3.3.3 辅助存储器 .....	79	5.3.6 二进制文件和文本文件 .....	131
3.4 输入输出设备 .....	81	5.4 设备管理与驱动 .....	133
3.4.1 I/O 设备的分类 .....	81	5.5 典型操作系统 .....	135
3.4.2 常见输入设备 .....	82	5.5.1 Windows 系列 .....	135
3.4.3 常见输出设备 .....	83	5.5.2 Mac OS 系列 .....	136
3.5 总线与接口 .....	84	5.5.3 UNIX .....	136
3.5.1 总线 .....	84	5.5.4 Linux .....	137
3.5.2 接口 .....	86	5.5.5 移动操作系统 .....	138
本章小结 .....	90	本章小结 .....	139
本章习题 .....	90	本章习题 .....	140
第 4 章 数据结构与算法 .....	93	第 6 章 程序设计语言 .....	143
4.1 概述 .....	93	6.1 计算机语言的发展 .....	143
4.1.1 数据结构 .....	93	6.1.1 自然语言与形式语言 .....	144
4.1.2 算法 .....	95	6.1.2 机器语言与汇编语言 .....	145
4.2 线性结构 .....	99	6.1.3 高级语言 .....	146
4.2.1 线性表和串 .....	99	6.1.4 脚本语言 .....	147
4.2.2 栈和队列 .....	100	6.2 程序设计语言的范型 .....	147
4.2.3 数组 .....	101	6.3 程序设计语言的语法元素和功能划分 .....	149
4.3 非线性结构 .....	102	6.4 程序的生成和运行 .....	154
4.3.1 树 .....	102	6.4.1 编译和解释 .....	154
4.3.2 图 .....	104	6.4.2 程序设计环境 .....	157
4.4 基本算法 .....	105	本章小结 .....	158
4.4.1 排序 .....	106	本章习题 .....	159
4.4.2 查找 .....	108	第 7 章 数据库技术与应用 .....	162
4.5 递归 .....	109	7.1 数据库的基本概念 .....	162
本章小结 .....	110	7.1.1 什么是数据库 .....	162
本章习题 .....	111	7.1.2 数据模型 .....	165
第 5 章 操作系统 .....	114	7.1.3 数据字典 .....	167
5.1 操作系统概述 .....	114	7.1.4 数据库系统 .....	167
5.1.1 操作系统的概念与组成 .....	114	7.2 关系数据库及其运算 .....	168
5.1.2 操作系统的功能 .....	115	7.2.1 关系模型 .....	168
5.1.3 操作系统的分类 .....	116	7.2.2 结构化查询语言 .....	171
5.2 处理器管理和内存管理 .....	118	7.3 数据仓库与数据挖掘 .....	174
5.2.1 处理器管理与进程调度 .....	119	7.3.1 数据仓库 .....	174
5.2.2 内存管理 .....	124	7.3.2 数据挖掘 .....	175
5.3 文件系统 .....	126	本章小结 .....	179
5.3.1 文件的基本概念 .....	126	本章习题 .....	179
5.3.2 文件结构及存取方式 .....	126		

第 8 章 软件工程 .....	182	9.2.3 TCP/IP 模型 .....	215
8.1 从软件到软件工程 .....	182	9.3 组建网络 .....	217
8.1.1 再认识软件 .....	182	9.3.1 计算机网络拓扑 .....	218
8.1.2 软件工程的内涵 .....	183	9.3.2 网络的物理组成 .....	219
8.2 软件开发过程 .....	184	9.3.3 网络连接 .....	219
8.2.1 软件生命周期 .....	184	9.4 Internet 及其应用 .....	220
8.2.2 软件过程模型 .....	185	9.4.1 Internet 简介 .....	220
8.3 软件工程的方法与工具 .....	188	9.4.2 传输协议与地址 .....	221
8.3.1 结构化方法 .....	188	9.4.3 Internet 应用 .....	225
8.3.2 面向对象方法 .....	191	9.4.4 Internet 发展新趋势 .....	228
8.3.3 计算机辅助软件工程 .....	194	9.5 云计算与物联网 .....	229
8.4 软件需求管理 .....	195	9.5.1 云计算 .....	229
8.4.1 软件需求 .....	196	9.5.2 物联网 .....	231
8.4.2 需求分析 .....	196	本章小结 .....	232
8.4.3 需求的管理 .....	198	本章习题 .....	233
8.5 软件质量管理 .....	199	第 10 章 信息系统安全 .....	236
8.5.1 软件质量与评价 .....	199	10.1 信息安全概述 .....	236
8.5.2 软件评审与测试 .....	200	10.1.1 安全威胁 .....	236
8.6 软件项目管理 .....	202	10.1.2 对信息系统安全的理解 .....	237
8.6.1 项目管理内容 .....	202	10.1.3 如何才能安全 .....	238
8.6.2 项目管理过程 .....	203	10.2 数据加密 .....	241
8.6.3 团队 .....	205	10.2.1 密码学基础 .....	241
8.6.4 管理者 .....	205	10.2.2 加密技术 .....	242
本章小结 .....	206	10.3 计算机病毒 .....	244
本章习题 .....	206	10.3.1 计算机病毒概述 .....	244
第 9 章 计算机网络 .....	209	10.3.2 计算机病毒的作用机理 .....	246
9.1 什么是计算机网络 .....	209	10.3.3 计算机病毒防治 .....	248
9.1.1 计算机网络的历史 .....	209	10.4 防火墙与入侵检测技术 .....	249
9.1.2 计算机网络的含义 .....	210	10.4.1 防火墙 .....	249
9.1.3 计算机网络的分类 .....	211	10.4.2 入侵检测技术 .....	251
9.2 计算机网络体系结构 .....	213	本章小结 .....	256
9.2.1 概述 .....	213	本章习题 .....	257
9.2.2 OSI 参考模型 .....	214	参考文献 .....	259

# 第0章 绪论

本章介绍计算机科学的概念，讨论计算机科学的内涵，简要回顾计算机的发展历史，并分析计算机对现代社会的影响。

在本书中如无特殊说明，所说的计算机是指现代电子计算机。

计算机科学是在现代电子计算机发明以后，随着计算机的广泛应用和技术的发展逐渐形成的一门新兴学科。与传统领域相比，它的发展非常迅速，其影响几乎涉及我们社会的所有层面。

## 0.1 什么是计算机科学

计算机科学是研究计算机及其周围各种现象和规律的科学，亦即研究计算机系统结构、软件系统、人工智能以及计算本身的性质和问题的学科。计算机是一种由电能驱动，在一定控制下能够自动进行算术和逻辑运算的电子设备，通俗地说就是能够进行计算的机器。计算机处理的对象都是信息，因而也可以说，计算机科学是研究信息处理的科学。

### 0.1.1 计算机科学的提出

自计算机发明以来，曾经围绕着计算机科学能否独立地成为一门学科产生过许多争论。最早的计算机科学学位课程是由美国普度大学于1962年开设的，随后斯坦福大学也开设了同样的学位课程。但针对“计算机科学”这一名称在当时引起了激烈的争论。毕竟当时的计算机主要用于数值计算，因此大多数科学家认为使用计算机仅仅是编程问题，不需要做深刻的科学思考，没有必要设立学位。总而言之，当时很多人认为计算机从本质上说是一种职业而非学科。

20世纪七八十年代计算技术得到了迅猛的发展并开始渗透到大多数学科领域，但以往激烈的争论仍在继续。计算机科学能否作为一门学科，计算机科学是理科还是工科或者只是一门技术？针对激烈的争论，1985年ACM和IEEE-CS联手组成攻关组开始了对计算作为一门学科的存在性证明，经过近4年的工作ACM攻关组提交了计算作为一门学科（Computing as a Discipline）的报告，完成了这一任务。该报告的主要内容刊登在1989年1月的《ACM通讯》（Communications of the ACM）杂志上。

#### ACM：计算机协会

ACM（Association for Computing Machinery）是一个国际科学教育计算机组织，它致力于发展在高级艺术、最新科学、工程技术和应用领域中的信息技术。它强调在专业领域或在社会感兴趣的领域中培养、发展开放式的信息交换，推动高级专业技术和通用标准发展。



#### 知识扩展

ACM于1947年创立，是世界上第一个也一直是最大的科学教育计算机组织。其创立者和成员都是数学家和电子工程师，其中之一约翰·迈克利（John Mauchly）是ENIAC的发明家之一。成立这个组织的初衷是使计算机和新兴工业领域的科学家和技术人员能有一个交换信息、经验知识和创新思想的场合。经过几十年的发展，ACM的成员们为今天我们所所谓的“信息时代”做出了贡献。他们所取得的成就大部分出版在ACM印刷刊物上，并获得了ACM颁发的各种领域中的杰出贡献奖，如A. M. Turing（图灵）奖。

国际公认的计算机科学领域的最高荣誉是ACM设立的图灵奖，被誉为计算机科学的诺贝尔奖。它的获得者都是本领域最为出色的科学家和先驱。华人中首获图灵奖的是姚期智先生，他于



2000 年以其对计算理论做出的根本性的、意义重大的贡献而获得这一崇高荣誉。

### IEEE-CS: 电气电子工程师学会计算机分会

电气电子工程师学会 (Institute of Electrical and Electronics Engineers) 是一个国际性的电子技术与信息科学工程师的协会, 是世界上最大的专业技术组织之一, 拥有来自 170 多个国家的数十万会员。IEEE 设有 37 个专业分会 (Society) 和 3 个联合会 (Council), 37 个专业分会覆盖了电力、电子、信息等广泛领域, 其中计算机分会 (Computer Society, CS) 有 10 多万会员。

IEEE 是定位在“科学和教育, 并直接面向电子电气工程、通信、计算机工程、计算机科学理论和原理研究, 以及相关工程分支的艺术和科学的组织”。IEEE 承担着多个科学期刊和会议组织者的角色。IEEE 制定了全世界电子和电气还有计算机科学领域超过 900 个现行工业标准, 是主要的国际标准机构。

ACM 对计算机科学是这样定义的: 计算机科学 (计算学科) 是对描述和变换信息的算法过程的系统研究, 包括它的理论、分析、设计、有效性、实现和应用。他们认为, 全部计算科学的基本问题是“什么能够 (有效地) 自动进行”。如今“计算机科学”一词是一个非常广泛的概念, 在本书里, 我们将其定义为“与计算机相关的问题”。

计算机科学是一门实用性很强、发展极其迅速的面向广大社会的学科, 它建立在数学、电子学 (特别是微电子学)、磁学、光学、精密机械等多门学科的基础之上。但是, 它并不是简单地应用某些学科的知识, 而是经过高度综合形成一整套有关信息表示、变换、存储、处理、控制和利用的理论、方法和技术。它是一门包含各种各样与计算和信息处理相关的主题的系统学科, 从抽象的算法分析、形式化语法等, 到更具体的主题如编程语言、程序设计、软件和硬件等。作为一门学科, 它与数学、电子学、机械学和管理学等学科之间存在不同程度的交叉和覆盖。

## 0.1.2 计算机科学领域

计算机科学的分支领域包括: 数值和符号计算、体系结构、数据结构和算法、操作系统、程序设计、软件工程、数据库和信息检索、计算理论和人工智能等。

### 1. 数值和符号计算

本领域研究有效和精确地求解由数学模型所导出方程的一般方法。基本问题包括: 怎样才能按照给定精度很快地解出给定类型的方程; 怎样对方程进行符号运算, 如积分、微分和化简为最小项等; 怎样把这些问题的回答加入到有效、可靠、高质量的数学软件包中去。

### 2. 体系结构

本领域主要研究将硬件和软件组织成有效和可靠系统的方法。基本问题包括: 在一个机器中实现信息处理、存储和通信的最好方法是什么; 如何设计和控制大的计算系统并且使它们能够在有错误和故障的情况下完成预期的工作; 什么类型的体系结构能使许多处理单元有效地协同工作, 实现并行计算; 怎样测度计算机的性能。

### 3. 数据结构和算法

本领域主要研究一些特定类型的问题及相对应的数据结构和解决方法。基本问题包括: 对给定类型的问题, 最好的算法是什么; 它们要求多少存储空间和时间; 空间与时间的折中方案是什么; 存取数据最好的方法是什么; 最好算法的最坏情况是什么; 算法的运行按平均来说好到何种程度; 算法的一般化程度, 即什么类型的问题可以用类似的方法处理。

#### 4. 操作系统

本领域研究允许多种资源在程序执行中有效配合的控制机制。基本问题包括：在计算机系统运行的各级上可见的对象和允许的操作是什么；每一类资源允许有效使用的最小操作集是什么；怎样组织接口，使得用户只处理资源的抽象形式，而可以不管硬件的实际细节；对作业调度、内存管理、数据传输管理、外存资源管理、并发任务间的协调控制、可靠性和安全性的有效控制策略是什么；系统应该在什么功能上扩展；怎样组织分布式计算，使得许多由通信网络连接起来的自治机器能够参与同一计算。

#### 5. 程序设计

本领域研究执行算法的虚拟机的符号表达、算法和数据的符号表达以及从高级语言到机器码的有效翻译。基本问题包括：由一种语言给出的虚拟机的可能的组织（数据类型、运算、控制结构、引入新类型和运算的机制）是什么；这些抽象怎样在计算机上实现；用什么样的符号表达（语法）可以有效地指明计算机应该做什么。

#### 6. 软件工程

本领域研究满足技术要求、安全、可靠、可信的程序和大型软件系统的设计。基本问题包括：程序和系统开发背后的原理是什么；怎样证明程序或系统满足技术要求；怎样给定技术要求，使之不遗漏重要的情况，而且可以分析它的安全性；怎样使软件系统通过不同阶段不断改进；怎样使软件设计得易理解和易修改。

#### 7. 数据库和信息检索

本领域研究对大量持续的可分享的数据集合进行组织，使之能够有效地查询和刷新。基本问题包括：用什么样的模型化概念表示数据元和它们之间的关系；怎样把存储、定位、匹配和检索等基本操作组合成有效的事务处理；这些事务处理怎样与用户有效地交互；怎样把高级查询翻译成高性能的程序；什么样的机器结构能有效地检索和刷新；怎样保护数据，以抵制非法存取、泄露或破坏；怎样保护大型数据库不会由于同时刷新而导致不相容；当数据分散在许多机器上时，怎样使安全保护和访问性能二者得以兼顾；怎样索引和分类正文，以达到有效的检索。

#### 8. 人工智能和计算理论

本领域研究动物和人类（智能）行为模型。基本的问题包括：摹本的行为模型是什么，我们怎样建造机器来模拟；由规则赋值、推理、演绎和模式计算所描写的智能可以达到什么程度；由这些模型模拟行为的机器最终能达到什么性能；感知的数据应如何编码，使得类似的模式有类似的码字；驱动码怎样与感知码相联系；学习系统的体系结构如何，以及这些系统如何表示它们对外部世界的知识；怎样才能用有穷离散过程精确地逼近连续或无穷的过程；怎么处理逼近导致的误差等。



注意

读者在阅读上述内容及本章其他内容时可能会发现某些概念或术语比较生僻，这些概念和术语我们将在以后的章节逐一讲述。

### 0.1.3 计算机与计算机模型

阿兰·图灵在1937年首次提出了一个通用的计算设备的设想。他设想所有的计算都可能在一台特殊的机器上执行，这就是现在所说的图灵机。尽管图灵对这样一种机器进行了数学上的描述，但他更关注计算的哲学定义，而不是建造一台真实的机器。他将该模型建立在人们进行计算过程

的行为上，并将这些行为抽象到用于计算的机器模型中，这才真正地改变了世界。

通用图灵机是对现代计算机的首次描述，只要提供合适的程序该机器就能做任何运算。图灵机是计算机的功能抽象模型，它可以实现真实计算机的所有功能，但仅限于功能描述，对于性能描述需要采用其他方式。

基于通用图灵机建造的计算机都是在存储器中存储数据。在 1944~1945 年期间，冯·诺依曼指出，鉴于程序和数据在逻辑上是相同的，因此程序也能存储在计算机的存储器中。基于冯·诺依曼模型建造的计算机分为 4 个子系统：存储器、算术/逻辑单元、控制单元和输入/输出单元。冯·诺依曼模型中要求程序必须存储在内存中，这与早期只有数据才存储在存储器中的计算机结构完全不同。现代计算机的存储单元主要用来存储程序及其响应数据，这意味着数据和程序应该具有相同的格式，实际上它们都是以位模式（0 和 1 序列）存储在内存中的。

计算机科学的大部分研究是基于“图灵机模型”和“冯·诺依曼模型”的，它们是绝大多数实际机器的计算模型。作为此模型的开山鼻祖，邱奇-图灵论题（Church-Turing Thesis）表明，尽管在计算的时间、空间效率上可能有所差异，现有的各种计算设备在计算能力上是却等同的。尽管这个理论通常被认为是计算机科学的基础，可是科学家也研究其他种类的机器，如在实际层面上的并行计算机以及在理论层面上的概率计算机和量子计算机等。在这个意义上讲，计算机只是一种计算的工具。著名的计算机科学家 Dijkstra 有一句名言：“计算机科学之关注于计算机，并不甚于天文学之关注于望远镜。”

#### 0.1.4 计算机科学中的经典问题

在社会发展过程中人们提出过许多具有深远意义的科学问题，它们对计算机科学的一些分支领域的形成和发展产生了重要的作用。另外，在计算机科学的发展过程中，为了便于理解有关问题和概念，人们还给出了不少反映该学科某一方面本质特征的典型实例。计算机科学中典型问题的提出及研究不仅有助于我们深刻地理解该学科，而且还对该学科的发展有着十分重要的推动作用。下面分别对图论中有代表性的哥尼斯堡七桥问题、算法与算法复杂性领域中有代表性的梵天塔问题（又译为汉诺塔问题），以及并发和互斥领域中的哲学家共餐问题进行分析。计算机科学中其他的典型问题请读者参考有关资料。

##### 1. 哥尼斯堡七桥问题

18 世纪的东普鲁士有一座名叫哥尼斯堡（Konigsberg）的城堡，城中有一个岛。普雷格尔（Pregel）河的两条支流环绕其旁，并将整个城市分成北区、东区、南区 and 岛区 4 个区域。全城共有 7 座桥，这 7 座桥将 4 个城区连起来，如图 0-1 所示。人们常通过这 7 座桥到各城区游玩，于是产生了一个有趣的数学难题：寻找不重复地走遍这 7 座桥回到原出发点的路径。该问题就是著名的哥尼斯堡七桥问题。

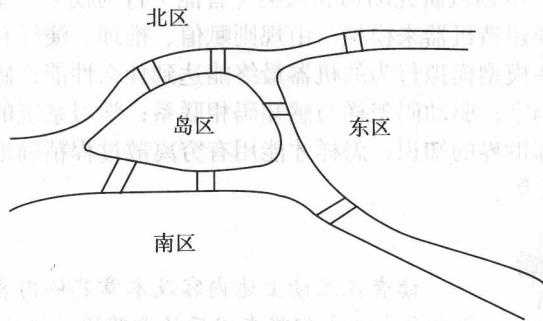


图 0-1 哥尼斯堡七桥示意图

1736 年大数学家列昂纳德·欧拉发表了关于哥尼斯堡七桥问题的论文——《与位置几何有关的一个问题的解》。他在文中指出，从一点出发不重复地走遍 7 座桥最后又回到原出发点是可能的。为了解决哥尼斯堡七桥问题，欧拉用 4 个字母 A、B、C、D 代表 4 个城区并用 7 条线表示 7 座桥，如图 0-2 所示。

在图 0-2 中, 只有 4 个点和 7 条线, 这样做是基于问题本质来考虑的。它抽象出问题最本质的东西, 忽视问题非本质的东西如桥的长度等, 从而将哥尼斯堡七桥问题抽象为一个数学问题, 即经过图中各边一次且仅一次的回路问题。欧拉在论文中论证了这样的回路是不存在的, 后来人们把存在这种回路的图称为欧拉图。

欧拉在论文中将问题进行了一般化处理, 即对给定的任意一个河道图与任意多座桥, 判定是否能每座桥恰好走过一次, 并用数学方法给出了 3 条判定的规则:

1) 如果通奇数座桥的地方不止两个, 满足要求的路线是找不到的。

2) 如果只有两个地方通奇数座桥, 可以从这两个地方之一出发, 找到所要求的路线。

3) 如果没有一个地方是通奇数座桥的, 则无论从哪里出发所要求的路线都能实现。

欧拉的论文为图论的形成奠定了基础, 今天图论已广泛地应用于运筹学、信息论、控制论等学科之中, 并已成为我们对现实问题进行抽象的一个强有力的数学工具。随着计算机科学的发展, 图论在计算机科学中的作用越来越大, 同时图论本身也得到了充分的发展。

在图论中还有一个很著名的哈密尔顿回路问题。该问题是 1859 年爱尔兰著名学者威廉·哈密尔顿爵士 (W. R. Hamilton) 提出的一个数学问题。其大意是在某个图  $G$  中能不能找到这样的路径, 即从一点出发, 不重复地走过所有结点, 最后又回到原出发点。哈密尔顿回路问题与欧拉回路问题看上去十分相似, 然而又是完全不同的两个问题。哈密尔顿回路问题是访问每个结点一次, 而欧拉回路问题是访问每条边一次。对图  $G$  是否存在, 前面已给出欧拉回路问题的充分必要条件, 而至今仍未找到满足哈密尔顿回路问题的充分必要条件。

## 2. 梵天塔问题

相传印度教的天神梵天在创造地球这一世界时, 建了一座神庙。神庙里竖立三根宝石柱子, 柱子由一个铜座支撑。梵天将 64 个直径大小不一的金盘子, 按照从大到小的顺序依次套放在第一根柱子上, 形成一座金塔, 如图 0-3 所示, 即所谓的梵天塔, 又称汉诺塔。天神让庙里的僧侣们将第一根柱子上的 64 个盘子, 借助第二根柱子, 全部移到第三根柱子上, 即将整个塔迁移。同时定下 3 条规则:

1) 每次只能移动一个盘子。

2) 盘子只能在三根柱子上来回移动, 不能放在他处。

3) 在移动过程中, 三根柱子上的盘子必须始终保持大盘在下, 小盘在上。

天神说, 当这 64 个盘子全部移到第三根柱子上后, 世界末日就要到了。这就是著名的梵天塔问题。

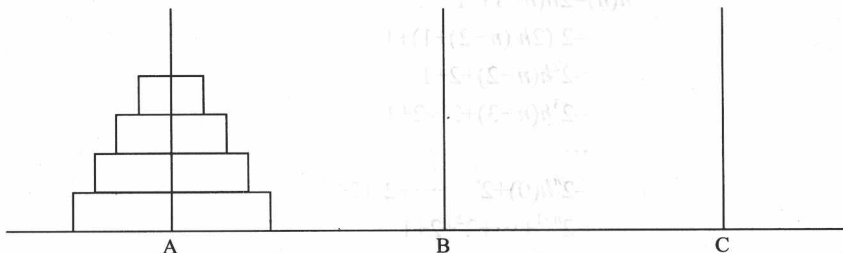


图 0-3 梵天塔问题示意图

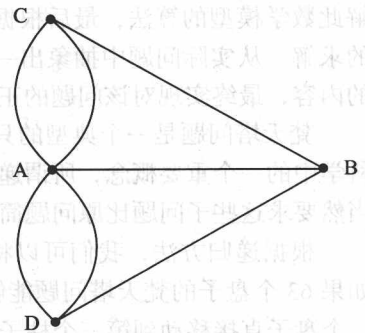


图 0-2 哥尼斯堡七桥问题的欧拉图



用计算机求解一个实际问题，首先要从这个实际问题中抽象出一个数学模型，然后设计一个解此数学模型的算法，最后根据算法编写程序，经过调试、编译、链接和运行，从而完成该问题的求解。从实际问题中抽象出一个数学模型的实质，其实就是用数学的方法抽取其主要的、本质的内容，最终实现对该问题的正确认识。

梵天塔问题是一个典型的只能用递归方法，而不能用其他方法来解决的问题。递归是计算机科学中的一个重要概念，所谓递归就是将一个较大的问题，归约为一个或多个子问题的求解方法。当然要求这些子问题比原问题简单一些，且在结构上与原问题相同。

根据递归方法，我们可以将 64 个盘子的梵天塔问题，转化为求解 63 个盘子的梵天塔问题。如果 63 个盘子的梵天塔问题能够解决，则可以先将 63 个盘子先移动到第二个柱子上，再将最后一个盘子直接移动到第三个柱子上，最后将 63 个盘子从第二个柱子移动到第三个柱子上，这样则可以解决 64 个盘子的梵天塔问题。依此类推，63 个盘子的梵天塔问题，可以转化为 62 个盘子的梵天塔问题。62 个盘子的梵天塔问题，又可以转化为 61 个盘子的梵天塔问题，直到 1 个盘子的梵天塔问题。再由 1 个盘子的梵天塔问题的求解求出 2 个盘子的梵天塔问题，直到解出 64 个盘子的梵天塔问题。

下面用 C 语言对该问题的求解算法进行描述：

```

hanoi(int n,char left,char middle,char right)
{
    if(n==1) move(1,one,_,three);
    else
    {
        hanoi(n-1,left,right,middle);
        move(1,left,_,right);
        hanoi(n-1,middle,left,right);
    }
}

```

代码中， $n$  表示  $n$  个盘子的梵天塔问题， $left$  表示第一个柱子， $middle$  表示第二个柱子， $right$  表示第三个柱子。函数  $hanoi(n-1,left,right,middle)$  表示  $n-1$  阶梵天塔，从第一个柱子借助第三个柱子先移到第二个柱子上，函数  $move(1,left,_,right)$  表示将第一个柱子上最后一个盘子，直接放到第三个柱子上。函数  $hanoi(n-1,middle,left,right)$  表示  $n-1$  个盘子，从第二个柱子借助第一个柱子移到第三个柱子上。

在以上 C 语言描述的算法基础上做适当扩充，就可以形成一个完整的程序。经过编译和链接后，计算机就可以执行这个程序，并严格地按照递归的方法将问题求解出来。现在的问题是，当  $n=64$  时，即有 64 个盘子时需要移动多少次盘子和要用多少时间。按照上面的算法， $n$  个盘子的梵天塔问题需要移动的盘子数是  $n-1$  个盘子的梵天塔问题需要移动的盘子数的 2 倍加 1。于是：

$$\begin{aligned}
 h(n) &= 2h(n-1)+1 \\
 &= 2(2h(n-2)+1)+1 \\
 &= 2^2h(n-2)+2+1 \\
 &= 2^3h(n-3)+2^2+2+1 \\
 &\dots \\
 &= 2^n h(0)+2^{n-1}+\dots+2^2+2+1 \\
 &= 2^{n-1}+\dots+2^2+2+1 \\
 &= 2^n-1
 \end{aligned}$$