



教育部大学计算机课程改革项目规划教材

国家精品课程主讲教材

# 大学计算机

## ——信息、计算与智能

□ 李波 主编

□ 赵英良 程向前 乔亚男 编



高等教育出版社  
HIGHER EDUCATION PRESS



教育部大学计算机课程改革项目规划教材  
国家精品课程主讲教材

# 大学计算机——信息、计算与智能

Daxue Jisuanji——Xinxi、Jisuan yu Zhineng

李 波 主编

赵英良 程向前 乔亚男 编



高等教育出版社·北京  
HIGHER EDUCATION PRESS BEIJING

## 内容提要

本书根据 2012 年教育部高等教育司教改项目“理工类专业大学计算机系列课程建设”(高教司 2012118)和 2009 年陕西普通高等学校教学改革研究重点攻关项目“计算思维为核心的计算机基础教学改革研究与实践项目”(11BG01)这两个项目的研究设想及成果编写而成。全书共 7 章,内容主要分为数学基础、信息、计算、智能、问题求解、算法与数据结构、计算思维,试图从数学描述、计算机科学理论、问题求解及构造等方面初步让学生了解和掌握计算机学科的核心内容,体会计算机思维。在实验方面,每章都精心设计了若干实验,并给出实验指导。学生利用这些工具和给出的参考实现可方便地完成这些实验,从而帮助学生理解和掌握书中的基本概念和原理。

本书适合作为高等学校大学计算机基础课程和计算机专业的计算机导论课程的教材,也可作为认识和了解计算机学科的参考书。

## 图书在版编目(CIP)数据

大学计算机:信息、计算与智能/李波主编.

—北京:高等教育出版社,2013.8

ISBN 978-7-04-038202-0

I. ①大… II. ①李… III. ①电子计算机-高等学校-教材 IV. ①TP3

中国版本图书馆 CIP 数据核字(2013)第 174523 号

|         |          |          |          |
|---------|----------|----------|----------|
| 策划编辑 刘茜 | 责任编辑 唐德凯 | 封面设计 张申申 | 版式设计 杜微言 |
| 插图绘制 尹莉 | 责任校对 杨凤玲 | 责任印制 刘思涵 |          |

|      |                   |      |                                                                   |
|------|-------------------|------|-------------------------------------------------------------------|
| 出版发行 | 高等教育出版社           | 咨询电话 | 400-810-0598                                                      |
| 社址   | 北京市西城区德外大街 4 号    | 网 址  | <a href="http://www.hep.edu.cn">http://www.hep.edu.cn</a>         |
| 邮政编码 | 100120            |      | <a href="http://www.hep.com.cn">http://www.hep.com.cn</a>         |
| 印 刷  | 山东鸿杰印务集团有限公司      | 网上订购 | <a href="http://www.landaco.com">http://www.landaco.com</a>       |
| 开 本  | 787mm×1092mm 1/16 |      | <a href="http://www.landaco.com.cn">http://www.landaco.com.cn</a> |
| 印 张  | 23.75             | 版 次  | 2013 年 8 月第 1 版                                                   |
| 字 数  | 530 千字            | 印 次  | 2013 年 8 月第 1 次印刷                                                 |
| 购书热线 | 010-58581118      | 定 价  | 33.00 元                                                           |

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换

版权所有 侵权必究

物 料 号 38202-00

# 序

图灵奖得主 Edsger Dijkstra 说过“我们所使用的工具影响着我们的思维方式和思维习惯，从而也将深刻地影响着我们的思维能力。”计算机和互联网为人们终生学习营造了广阔的空间和良好的学习环境，它们不仅提供了解决问题的有效方法和手段，而且提供了一种独特的处理问题的思维方式，这种思维方式正在对人们产生深刻的影响，并将进一步推动人类文明进步和科技发展。

科学发现三大支柱（理论科学、实验科学和计算科学）中，源于计算科学领域的计算思维正日益广泛地被人们重视和探究并运用。计算思维是指运用计算科学的基础概念去求解问题、设计系统和理解人类行为的一种方法，它包括了一系列广泛的计算机科学的思维概念、方法和工具。事实上，我们已经见证了计算思维对其他学科的影响。例如：计算生物学正在改变着生物学家的思考方式；计算博弈理论正在改变着经济学家的思考方式；纳米计算正在改变着化学家的思考方式；量子计算正在改变着物理学家的思考方式；等等。

大学教育不能局限于基本知识传授，更要培养学生的理性思维能力，对科学追求的精神，以及学生的高尚人格。很高兴见到由西安交通大学计算机教学实验中心老师们编写的这部书稿，对他们在教学研究上的不懈努力和开拓精神感到非常欣慰，他们一直在大学计算机课程改革道路上不断探索、实践、总结、优化，致力于提供更完美的学习内容，引导学生的认知及思维层面的潜力激发。

这本书从全新的视角对大学计算机知识进行了梳理和组织，体现了作者对于计算思维培养的深入理解和充分认识。内容主要分为数学基础、信息、计算、智能、问题求解、算法与数据结构和计算思维 7 个章节，试图从数学基础、计算机科学理论、问题求解和构造等方面初步让学生把握学科的核心内容，体验计算思维。

第 1 章有选择地集中阐述了计算机学科相关的重要数学理论，因为理论上，所有能被计算机处理的问题均可以用数学方法解决，它是计算学科中最根本的研究方法。作者用著名的哥尼斯堡七桥问题、四色猜想问题来解释问题求解的数学抽象，有趣而生动。看得出来，在融会贯通计算相关内容上狠下了一番功夫，见解独到，写得也通俗易懂。

第 2、3、4 章选取了“什么是信息”、“什么是计算”、“什么是智能”这三个计算机学科的基本问题作为核心内容进行阐述和讨论。其中，对计算本质的认识过程，包括计算手段应该器械化、计算过程应该形式化、计算的执行应该自动化。计算机的本质是“以简单的、有限的离散构造解决无限的问题”或“以简单的、有穷的离散构造解决无穷的问题”等观点为主线的精彩阐述，体现了作者在信息表示、计算理论、人工智能方面多年的积累，更多地向读者展现了关于信息、计算和智能方面的历史、人物以及发生在他们身上的经典故事，所以，阅读起来并

## II 序

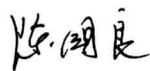
不会感到枯燥乏味。

第 5、6 章围绕计算机科学中的系统构造和问题求解，介绍计算机求解问题的一般方法和求解策略。例举“囚徒困境”、“汉诺塔问题”、“旅行商问题”等学科中经典问题的求解算法，不仅让关键问题一目了然，而且富有趣味。

第 7 章在列举计算生物学、仿生计算、计算社会科学等学科相互融合的案例的基础上，总结和引出了计算思维的概念。以 21 世纪中将影响社会发展的若干重要计算理论、模型为逻辑主线将内容贯穿，涉及的面非常广，内容也非常新，看得出作者对于当前的计算科学前沿是相当熟悉和了解的。

本书推荐给对计算机科学有兴趣的广大在校大学生。除此之外，对于那些对计算机科学有一定了解的读者来说，阅读本书后也有利于理解计算机学科的核心概念和原理，在学科知识的融会贯通上定有收获。

马克思说过，一门学科，只有运用了数学才算是成熟的学科。而我们认为，一门学科，一旦运用了计算机技术，它就成为了先进的学科。计算机科学不只是关乎计算机，就像音乐产业不只是关乎麦克风一样，计算思维是普适的、泛在的，是振兴大学计算教育的途径，必将促使科学与工程领域产生革命性的创新成果。希望本书的读者能把握学科发展，了解计算之美，真正地能像计算机科学家一样思考。



中国科学院院士  
中国科学技术大学教授  
深圳大学教授  
2013 年 6 月

# 前 言

美国 National Research Council Committee on Fundamentals of Computer Science 在 2004 发表了题为 “Computer Science: Reflections on the Field” 的重要报告，对计算机科学 (Computer Science) 进行了较为准确的定义。报告认为，“计算机科学是研究计算机以及它们能干什么的一门学科。它研究抽象计算机的能力与局限，真实计算机的构造与特征，以及用于求解问题的数不清的计算机应用。研究领域主要涉及：符号及其操作；多种抽象概念的创造和操作；算法；创造各种人工结构，尤其是不受物理定律限制的结构；利用并应对指数增长；探索计算能力的基本极限；关注与人类智能相关的复杂的、分析的、理性的活动。”

计算机基础教学是培养大学生综合素质和创新能力不可或缺的重要环节。在新形势下，计算机基础教学的内涵在快速提升和不断丰富，进一步推进计算机基础教学改革、适应计算机科学技术发展的新趋势，是国家创新战略对计算机教学提出的重大要求。九校联盟 (C9) 计算机基础课程研讨会上达成共识：要旗帜鲜明地把“计算思维能力的培养”作为计算机基础教学的核心任务。所谓计算思维就是运用计算机科学的基础概念进行问题求解、系统设计和人类行为理解。

2005 年，在庆祝 SCIENCE 创刊 125 周年之际，该刊在当年 7 月 1 日出版的专辑上公布了 125 个最具挑战性的科学问题。在今后 1/4 个世纪的时间里，人们将致力于研究解决这些问题。其中前 25 个被认为是最重要的问题。与计算机科学直接相关的有两个：第 19 个问题“什么是传统计算的极限？”和第 94 个问题“通过计算机进行学习的极限是什么？”2008 年卡内基·梅隆大学计算机系的周以真 (JEANNETTE M. WING) 教授在 Communication of ACM 期刊上发表了题为 “Five deep questions in computing” 一文，提出了计算机学科五个基本问题。这五个基本问题是：

- ① Does P equal NP?
- ② What is computable?
- ③ What is intelligence?
- ④ What is information?
- ⑤ (How) can we build complex systems simply?

周以真的问题 1 和 2 与 SCIENCE 提出的第 19 个问题基本相同，和什么是计算、计算的极限和计算的复杂度有关。SCIENCE 的第 94 个问题和周以真的第 3 个问题相关，涉及什么是智能。

这五个基本问题是计算机学科的基础，虽然许多科学家已经进行了深入的研究，也取得了丰硕的成果，但这五个问题仍旧是计算机科学家当前和未来面临的科学挑战，需要进一步探索，

## II 前言

一旦有了新的认知或突破必然会促进计算机科学在科学上的巨大进步和在技术上的巨大革新。

在上述论述和思想的启发下选择什么是信息、什么是计算、什么是智能这三个问题作为大学生入门的第一门计算机课程——大学计算机基础的核心内容，希望学生对计算机学科的基本问题和理论基础有初步的了解和正确的认识。故该课程的教学目的是从培养学生科学认知能力出发，让学生理解和建立“信息、计算、智能”这三大核心科学概念；围绕计算思维的精髓培养学生掌握以“合理抽象、高效实现”为特征的构造性过程的能力；让学生了解学科发展，展示计算之美。

为了达到上述目标，该课程不应该是程序设计课程，不应该是信息素养课程，更不应该是计算机硬软件知识介绍课程，也不应介绍面向对象的构造方法等偏向技术实现的内容。第一门课强调的是思维训练而不是具体实现（**Mental not Metal**）。因此，本书的大部分内容相对以往的大学计算机基础的课程内容有较大的变化。主要内容分为数学基础、信息、计算、智能、问题求解、算法与数据结构和计算思维共七章，试图从数学、理论、构造等方面初步让学生学习学科的核心内容，体会计算思维。作为一种大学计算机基础课程的创新尝试，本书试图为大学一年级学生回答这样一些与计算机科学有关的问题：

- 大学生为什么需要了解计算机科学？

大学一年级学生需要重视大学计算机基础在学习过程中的作用，理解“计算思维”能力对大学生素质养成的支柱性影响，并由此建立更加完备的计算机基础知识体系和符合大学生学科发展所需的学习内容。

- 哪些重要的数学理论与计算机科学有关？

本书第1章描述了科学抽象的方法和原则，包括数学抽象，形式化与计算描述，计算机科学与数学的关系，常用的数学描述方法（包括集合、巴科斯范式、正则表达式、有限状态机等），以及超文本标记语言和可扩展标记语言。

- 信息科学与计算机科学之间有何关联？

第2章从信息论的角度描述了计算机科学中的信息处理问题，阐述了信息量的度量、信息的数字化过程、信息的编码和表示，以及人类信息处理的发展史。

- 计算机科学的核心理论和经典问题有哪些？

第3章回顾了人类对计算本质的认识过程和计算机发展的历史，阐述了计算理论中若干重要的问题和基础概念（图灵机和停机问题、通用编程语言）、计算机体系结构和计算机学科中的典型问题。

- 计算机科学在人工智能研究的发展方面有哪些重要进展？

第4章对与计算机科学有关的“智能”一词进行仔细的辨析和定义，并列举了计算机科学中经典的图灵测试进行说明。对人工智能发展历史和当前的发展前沿使用了若干典型案例（用遗传算法解决旅行商问题、群体智能）进行说明。

- 使用计算机进行问题求解的一般过程是什么？

第5章阐述了问题求解的基础和问题求解的一般过程中所包含的明确问题、理解问题、

寻找备选方案、选择最优方案、列出求解方案的指令和评价解决方案等六个步骤。

- 如何使用算法进行问题求解？

第6章阐述了算法的描述、时间复杂度和空间复杂度评价等算法基础概念，并列举了查找、排序、哈夫曼编码、最短路径、最小生成树等典型算方案例。

- 算法与数据结构之间有什么关系？

第6章阐述了二者之间的关系，并列举了线形表与数组、树与哈希表等常见的数据结构。

- 什么是计算思维？如何理解计算的本质？

第7章对计算思维和计算一词存在的中外学科定义上的微妙差别进行了详细的阐述，并举出案例进行说明。

- 计算机科学与其他学科发展的关系是什么？

第7章所描述的计算透镜成为计算机科学影响各学科门类的集中体现，同时列举了计算生物学、仿生计算、计算社会科学等计算机科学与其他学科相互融合的案例。

- 计算与社会发展之间存在哪些联系？

第7章所阐述的社会计算、普适计算、企业计算揭示了计算科学在社会发展方面所取得的大量成果与发展中的案例。

本书在实验方面采用了函数式程序设计语言 **Haskell** 作为描述和试验工具，也引入了 **Netlogo**、**JFLAP**、图灵机模拟器和 **CLab** 等工具。每章都精心设计了若干实验，并给出了全部 **Haskell** 参考程序。利用这些工具和参考程序可方便地完成这些实验，从而帮助学生理解和掌握书中的基本概念和原理。

在附录中给出了 **Haskell** 程序设计的基本知识和学习资源，可帮助学生理解书中的基本内容，提供进一步学习的扩展知识。本书不是程序设计语言的教材，不期望学生通过该课程的学习掌握程序设计，而把这个任务留给后续课程。“大学计算机基础”只是大学阶段众多课程中的一门，该课程的学习也只是培养计算思维的第一步，后续的课程设置及内容安排更要进一步地贯彻计算思维，通过一系列的课程和训练使大学生真正地能像计算机科学家一样思考。

本书的教学重点在前6章，主要通过讲授方式帮助学生掌握基本概念和原理。第7章可以采用“学生自我阅读—教师讲授重点案例—进行课堂讨论”的教学方式，以激发学生自身的创造力和思辨精神。实验设计得较多也较为丰富，教师可自行安排和参考。

本书由李波主编，乔亚男（第1章）、李波（第2—4章和附录）、赵英良（第5、6章）和程向前（第7章）提供了本书的初稿，由李波进行了取舍和修订，并添加了各章的实验。

在本书的编写过程中，学习和引用了中国科学技术大学陈国良院士、合肥工业大学李廉教授、西安交通大学邓建国教授和郑庆华教授、桂林电子科技大学董荣胜教授的学术思想，在此深表感谢。国防科技大学王挺教授和桂林电子科技大学董荣胜教授在百忙之中认真地对本书进行了审稿并提出了宝贵意见，国防科技大学的李曦副教授在 **Haskell** 引入计算机基础教学方面给予了精心的指导，在此一并表示衷心的感谢。西安交通大学计算机教学实验中心主任冯博琴教授十分关心本书的编写，多次提出建议和要求，在此深表感谢。

#### IV 前言

本书得到了2012年教育部高等教育司教改项目“理工类专业大学计算机系列课程建设”(高教司2012118)和2009年陕西普通高等学校教学改革研究重点攻关项目“计算思维为核心的计算机基础教学改革研究与实践项目”(11BG01)的支持,本书的主要思想也来自这两个项目。

由于编者水平有限,书中难免有缺点和错误,欢迎读者批评指正。

编 者

2013年4月

## 郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任；构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人进行严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话 (010) 58581897 58582371 58581879

反盗版举报传真 (010) 82086060

反盗版举报邮箱 dd@hep.com.cn

通信地址 北京市西城区德外大街4号 高等教育出版社法务部

邮政编码 100120

# 目 录

|                                   |    |
|-----------------------------------|----|
| 第 1 章 符号模型的数学描述 .....             | 1  |
| 1.1 科学抽象 .....                    | 1  |
| 1.1.1 科学抽象的意义 .....               | 2  |
| 1.1.2 数学抽象的基本原则 .....             | 3  |
| 1.1.3 七桥问题 .....                  | 5  |
| 1.1.4 四色猜想问题的抽象 .....             | 7  |
| 1.1.5 抽象的层次 .....                 | 8  |
| 1.2 形式化与计算描述 .....                | 9  |
| 1.2.1 形式化 .....                   | 9  |
| 1.2.2 计算科学与数学的关系 .....            | 12 |
| 1.2.3 计算描述的特点 .....               | 14 |
| 1.3 常用数学描述 .....                  | 16 |
| 1.3.1 集合 .....                    | 16 |
| 1.3.2 巴科斯范式 .....                 | 18 |
| 1.3.3 正则表达式 .....                 | 19 |
| 1.3.4 有限状态机 .....                 | 21 |
| 1.3.5 HTML .....                  | 23 |
| 1.3.6 XML .....                   | 25 |
| 1.4 有限状态机的应用举例 .....              | 27 |
| 1.5 实验 Haskell 使用入门 .....         | 30 |
| 1.5.1 什么是 Haskell .....           | 30 |
| 1.5.2 初次使用 Haskell .....          | 30 |
| 1.5.3 以编程方式使用 Haskell .....       | 32 |
| 1.5.4 实验安排建议 .....                | 33 |
| 1.6 进一步阅读材料 .....                 | 33 |
| 习题 .....                          | 34 |
| 第 2 章 什么是信息 .....                 | 35 |
| 2.1 信息的概念与定义 .....                | 35 |
| 2.1.1 热身问题 .....                  | 35 |
| 2.1.2 信息的定义及其物理解释 .....           | 36 |
| 2.2 信息量的数学表达 .....                | 40 |
| 2.2.1 信息量的数学表达——信息熵 .....         | 40 |
| 2.2.2 信息熵的应用 .....                | 43 |
| 2.2.3 信息概念的进一步讨论 .....            | 45 |
| 2.3 数字化 .....                     | 45 |
| 2.3.1 比特 .....                    | 45 |
| 2.3.2 信息表示的种类 .....               | 49 |
| 2.3.3 模拟信号的数字化 .....              | 50 |
| 2.4 信息表示方法 .....                  | 52 |
| 2.4.1 二进制 .....                   | 52 |
| 2.4.2 字符信息的编码 .....               | 55 |
| 2.4.3 图形图像的数据表示 .....             | 57 |
| 2.4.4 格雷码 .....                   | 58 |
| 2.5 信息处理 .....                    | 59 |
| 2.5.1 人类信息处理发展史 .....             | 59 |
| 2.5.2 信息处理过程 .....                | 60 |
| 2.6 哈夫曼编码 .....                   | 62 |
| 2.7 熵的计算及应用举例 .....               | 65 |
| 2.8 实验 .....                      | 67 |
| 实验一 香农实验 .....                    | 67 |
| 实验二 编写 Gray 码产生程序 .....           | 68 |
| 实验三 编写一个哈夫曼编解码的程序 .....           | 68 |
| 2.9 进一步阅读材料 .....                 | 69 |
| 习题 .....                          | 70 |
| 第 3 章 什么是计算 .....                 | 72 |
| 3.1 计算机的历史及人类对计算本质的认识<br>过程 ..... | 72 |

## IV 目录

|                         |            |                                |            |
|-------------------------|------------|--------------------------------|------------|
| 实验五 并行快速排序.....         | 291        | 7.3 社会计算.....                  | 317        |
| 6.7 进一步阅读材料.....        | 291        | 7.3.1 社会计算的研究内容.....           | 318        |
| 习题.....                 | 292        | 7.3.2 社会网络媒体.....              | 319        |
| <b>第 7 章 计算思维</b> ..... | <b>294</b> | 7.3.3 计算社会科学.....              | 323        |
| 7.1 计算思维导引.....         | 294        | 7.3.4 人本计算.....                | 325        |
| 7.1.1 科学思维.....         | 294        | 7.3.5 集体智能、威客、众包.....          | 327        |
| 7.1.2 计算思维的产生.....      | 297        | 7.4 普适计算.....                  | 328        |
| 7.1.3 计算思维的定义.....      | 299        | 7.5 普适计算应用举例.....              | 333        |
| 7.1.4 计算思维的特征和作用.....   | 301        | 7.6 实验.....                    | 337        |
| 7.1.5 计算思维与计算机科学.....   | 302        | 实验一 京都基因和基因组百科全书               |            |
| 7.1.6 计算思维培养.....       | 303        | (KEGG).....                    | 337        |
| 7.2 计算透镜.....           | 304        | 实验二 利用 CLab4 研究高级语言的机器         |            |
| 7.2.1 计算生物学.....        | 306        | 执行.....                        | 337        |
| 7.2.2 人类基因组计划.....      | 311        | 7.7 进一步阅读材料.....               | 339        |
| 7.2.3 霰弹枪测序.....        | 312        | 习题.....                        | 339        |
| 7.2.4 仿生计算.....         | 314        | <b>附录 Haskell 程序设计基础</b> ..... | <b>341</b> |
| 7.2.5 万维网和因特网.....      | 315        |                                |            |

# 第 1 章 符号模型的数学描述

数学方法是指解决数学问题的策略、途径和步骤，它是计算学科中最根本的研究方法。理论上，凡能被计算机处理的问题均可以转换为一个数学问题。换言之，所有能被计算机处理的问题均可以用数学方法解决。更进一步讲，凡能以离散数学为代表的构造性数学方法描述的问题，当该问题所涉及的论域为有穷，或虽为无穷但存在有穷表示时，这个问题也一定能用计算机来处理。

构造性数学方法具有 3 个基本特征：一是高度的抽象性和概括性；二是精确性，即逻辑的严密性及结论的确定性；三是应用的普遍性和可操作性。构造性数学方法的这些特征决定了它的普遍适用性和它在计算机科学研究中所具有的举足轻重的地位和作用。

## （1）为计算机科学与技术研究提供简洁精确的形式化语言

随着计算机科学技术研究不断向深层次发展，对于微观和宏观世界中存在的复杂的自然规律，需要抽象、准确、简洁地进行表述，而这只有用数学的形式化语言才能做到。数学模型就是运用数学的形式化语言，在观测和实验的基础上建立起来的，它有助于人们从本质上认识和把握客观世界。数学中众多的定理和公式就是典型的简洁而精确的形式化语言。

## （2）为科学技术研究提供数量分析及计算的方法

一门科学从定性分析发展到定量分析是成熟的标志，数学方法从中起了重要作用。计算机的问世更为科学的定量分析和理论计算提供了必要条件，使一些过去虽然能用数学语言描述，但仍然无法求解或不能及时求解的问题找到了解决的方法。如及时准确的天气预报、汛期水库水量调度等都是借助于精确的数值计算和理论分析进行的，其中数学和计算机都发挥了非常重要的作用——没有数学，找不到计算模型；没有高性能的计算机，就无法及时计算出结果。

## （3）提供逻辑推理的工具

数学的严密逻辑性使它成为建立一种理论体系的重要工具，公理化方法、形式化方法用数学方法研究推理过程，把逻辑推理形式加以公理化、符号化，为建立和发展科学的理论体系提供了有效的工具。

## 1.1 科学抽象

客观世界任何一个具体的事物都包含着各种各样的属性，包含着各个不同的方面、各种不同的关系。这些不同的属性、方面和关系彼此互相联系、制约，在一定条件下组成一个统一体，构成一个事物的整体。马克思说“具体之所以具体，因为它是许多规定的综合，因而是多样性的统一”。因此，离开了这种多样性的统一，就没有任何具体，也没有任何事物。任何一个事

物离开了它与周围的关系、联系，任何一个事物的属性离开了与同一事物其他属性的联系，就都变成不是具体的东西。

**抽象**是从许多事物中，舍弃个别的、非本质的属性，抽出共同的、本质的属性的过程。抽象，是具体的对立面，是被人们从对象的全部联系、全部关系中抽取出来、孤立出来的东西。它是许多关系中的一个关系，许多属性中的一个属性，许多方面中的一个方面。

科学的抽象方法是在认识或研究过程中，有意识、有目的地撇开研究对象的某些方面、因素和属性，并对对象的某一方面的关系、属性抽取出来的研究方法。

### 1.1.1 科学抽象的意义

康德把人类的认识过程分成感性认识、知性认识和理性认识3个阶段，其中知性认识的特点是假设演绎、分解拆零、条清缕晰、合理分解问题和充分展示矛盾产生的各种规定性，形成概念明晰且彼此分明，保证意义上清楚，字面上无误。

爱因斯坦认为，西方科学的发展是以两个伟大成就为基础的，一是希腊哲学家发明的形式逻辑体系（在欧几里得几何学中）；二是通过系统的实验发现有可能找出的因果关系。

科学性方法包括：抽象化，运用符号，通过建立模型、逻辑分析、推理、计算，进行推断、优化以及善于应用计算机进行数值实验等，把科学方法作为通用的语言，具有跨文化性、整体性、形式化、简约化及非功利性等特征，从而超越和挑战感性思维的局限，改变重经验描述而轻逻辑思辨的风气，形成以逻辑思维为基础的理性思维。合理地把握“有”与“无”，“虚”与“实”，“远”与“近”，“表”与“里”等各种关系，通过科学方法，使我们分析得快，知道得准，了解得多。引导人们抓住根本，集中要害，把知识树牢、树严，而不是只注意枝叶和花果。概念的科学抽象不是纯思维的产物。从形式上看，它是主观的、理想的、非原型的；从内容上看，它又是客观的、具体的。

抽象是科学最本质的特征之一。经典物理学中不难体会到诸如抽象性、精确性、严格性、广泛性等科学抽象的基本特征。当然，抽象并非只是自然科学所独有的性质，社会科学也都具有抽象的性质，只是科学的抽象有其自身的特点。

首先，是它具有客观实在的根基性（抽象性只是第二位的），如经典物理中的质点、刚体、轻绳、理想气体等。物理、化学等其他一些学科的抽象常常是借助于实验，但受实验制约，它不能只利用自身已有的概念、定理和公式，还需借助已知的相互关系，通过推理、计算而获得新的发现。以数学为例，全部数学概念，从初等的、原始的自然数、整数，最基本的点、线、面等图形，以及在此基础上形成的有理数、无理数、复数、函数、微积分、 $n$ 维空间、无穷空间等一系列高度抽象的概念，都来源于非常实在的现实“原型”和知觉经验。

其次，是一般性与普适性。因为科学抽象是舍弃了原型中那些只具有个别属性的东西，仅保留了量的关系、空间形式及公共的基本的物理特性。其忽略个别的性质（特殊性），就是限制性的、非本质的具体特征。在数学中甚至可以完全忽略其物质属性，如物理的、化学的，等等。比如最简单的自然数中的加、减、乘、除，对所有自然界中客观物体数量的累加或分配都

是适用的，而并不在意到底是什么东西的相加。

第三，是数理科学抽象具有一种高抽象性。因为它的抽象是一级一级地抽象和提高的，普适性越强，需要具有的抽象度就越高。以数学为例，数学的抽象所具有的特点是全抽象性和自抽象性。

### 1.1.2 数学抽象的基本原则

从纯数学的角度进行分析，数学抽象是一种构造性的活动。纯数学研究对象，即“数学抽象物”，正是通过这种思维活动得到建构的。

在严格的数学研究中，无论所涉及的对象是否具有明显的直观意义，都只能依据相应的定义去进行（演绎）推理，而不能求助于直观。从而，在这样的意义上，数学的抽象就是一种构造性的活动。数学的研究对象正是通过这种活动逻辑地得到“构造”的。

对于数学对象的“逻辑建构”可以做出进一步的分析。

#### 1. 理想化

理想化抽象就是通过对实际事物或一些客观现象进行比较和理想的概念化，并确立出一定的彼此关系。理想化抽象的例子很多，比如通常从几何角度讲的圆、直线都是理想化的，而实际生活中的圆、直线、三角形与理想情况相比较都是有误差的，都是近似的。

以希腊数学为例，它是从对雅典生活的具体事实所作的几何的和数值的刻画中抽象出来的，是从理想主义的层面来反映现实的。理想的圆、方、点等是无限精确的，极限化的东西，在客观中并不存在，这都是理想化的强调方式。多个质点用 $\Sigma$ 求和，通过可连续分割形成以简单求和来积分，用已知的理想状态（简单结果直接相加）来模拟，用简单运算叠加，凑成复杂运算，把一个大问题化成几个小问题来解，把研究对象作为特殊的具有特定属性的对象，并从其所处的个性环境中剥离出来。

相对于那些明显具有现实原型的数学对象而言，数学抽象是一个理想化的过程。例如在现实世界里找不到“没有大小的点”和“没有宽度的线”等。

科学的抽象增加了概念的广度、深度和清晰性。依靠抽象，达到用适当的公理，明确地把一些客观对象的共同本质都表达出来。随着科学中抽象程度的发展，需把某些客观对象的本质方面明确地表达出来。这种抽象可以在几个相继阶段中发生并可以把科学的不同分支联系起来。概念观念是判断的要素，是解释的工具，是逻辑的工具。它对概念的理解有两种方式：或是直接领会或是间接迂回地领会。

概念是确定的意义（经抽象共同性质达到整体性、稳定性、明晰性和普遍性，并消除含糊的歧义影响，使知识具有类比能力，使其标准化）。

科学概念在帮助人们认识未知事物时有3个方面的作用：鉴别的工具；补充的工具；把同类事物纳入同一种体系的工具。

一个科学概念的广度是指这个概念的适用场合的广泛性，以及所作抽象的确切性和相关性。由理性方法，提出一个简单的理想的模型。如质点，只要物体在运动中，其大小和形状对

运动状态影响不大,那么不论庞大的宇宙天体,还是微不足道的灰尘,都可以被看成一个质点,它的线度同轨道的大小相比较可忽略,同时还不管它的化学成分是什么——是固体,还是气体(如彗星)。因此,理想气体、点电荷等都是质点模型。把宇宙万物都看做质点,通过简化来突出本质,拨冗去繁排除形状和大小等非本质的属性,成为具有一定质量的单纯得不能再单纯的几何点。

概念的建立不仅受到严密性的制约,而且还受到应用目的和抽象背景起源的制约。当涉及转动问题时,质点简化失效,必须考虑物体的运动状态与物体大小、形状和质量分布的关系,同时若这些物体在运动时它的形状几乎不变,就可认为是刚体。

任何物体位置的变化都可以用平动+转动来模拟。

同时,抽象已经增加了对表示清晰性的要求。如果研究的对象是抽象的,那么对它一定要通过精确而抽象的描述来理解,而不是通过它的直观内容来理解。所谓从具体的特殊性到抽象的一般性,是指对事物经抽象化之后才把清晰性、可靠性、严密的精确性联系在一起,达到一种对概念的条清缕晰的认识。

### 2. 形式化

数学对象的“逻辑建构”还是一个“模式化”即“重新构造”的过程。由于数学对象的逻辑建构是借助于纯粹的数学语言得以完成的,因此相对于现实原型而言,通过数学抽象而形成的数学概念及概念体系(理论)就具有更为普遍的意义。它所反映的已不只是这一特定事物或现象的量性特性,而是一类事物在量的方面的共同特性。也正因为这样,数学的研究对象就应当被看成是一种(量化)模式。正如怀特海(White Head)所指出的:“数学就是对模式的研究”。

### 3. 精确化

由于数学对象的“逻辑建构”是借助于数学语言完成的,因此相对于一般的朴素概念或现实原型而言,这就是一个精确化的过程。例如,瞬时速度的概念只有借助于“导数”的概念才能得到严格的定义。

由于在简化阶段有唯一确定的解(只有这一种可能性),可达到一种不容置疑的真确性;而对各种可能都存在的情况,若不能判别是哪种可能性占具绝对优势,从而产生多重解,将丧失准确判断,容易造成歧路亡羊。

简化后,使之能够简单相加或相减,因此可以拆零和分解,可以标准化,流水线作业和批量生产,具有专业效益。

### 4. 自由化

由于运用数学语言进行“逻辑建构”,导致了与现实原型的“分离”,这就为思维的“自由创造”提供了极大的空间,从而创造出一个丰富多彩的“数学世界”。

### 5. 形式化

由于数学对象的“逻辑建构”,导致了数学以“模式”这种“纯形式”为直接对象进行研究。模式的建构标志着由特殊上升到了一般,以模式为直接对象去从事研究也就应当说是一种纯形式的研究。例如,就纯数学的研究而言,即使所涉及的概念和理论具有明显的直观意义(如欧几里得的几何理论),也不能求助于直观,而只能依据相应的定义去进行推理。在所谓的

“形式主义”观点中，数学证明是一种由明确给定的（变形）法则所唯一决定的形式过程。

在纯粹数学的研究中，应当借助于明确的定义去构造出相应的量化模式，并以此为直接对象从事纯形式的研究。也正因如此，作为数学抽象物的量化模式在概念意义上就应具有一定层次上的普遍性和概括性，在表述形式上则应具有无歧义的逻辑精确性和简洁性。现代的科学哲学研究突出强调了范例在整个科学活动中的重要作用，如著名科学家、哲学家库恩认为，科学传统（他称之为“范式”）主要体现在相应的范例之上。他认为：范式代表某些具体的科学成就事例与成功模式，涉及某些实际问题的阐释，科学家们通过认真体会和仿照这些表述进行自己的工作。显然，这也就是对于范式的直接肯定。

综上所述，由于数学的概念、理论、公式、定理、问题、方法等都应被看做模式。因此，上述的“模式建构形式化原则”就确实是数学抽象的基本准则，而数学就是对模式的研究。

另外，在计算机科学中抽象经常和建模（Modeling）对应，计算机最初用来对事物“建模”或者模拟事物。无论对象是核武器还是工资单，方法都是一样：创建一个和目标对象在某些重要方面相似的计算机系统，然后观察这个系统，根据对象的行为作出推断。这样做的核心概念是抽象：这里有一个理想的系统，常常由方程组来定义，它既像目标系统，又像计算机模型。建模取得了巨大的成功，现在它已经用来理解，甚至常常用以控制物质和概念上的系统，比如银河系、蛋白质、目录清单、飞行中的飞机等系统。建模的关键就是利用抽象思维寻找问题的简化表达，即将复杂的问题进行简化以达到计算机能求解的目的。从这种意义上讲，抽象首先是简化（Simplify），是处理及管理复杂度（Complexity）的工具。“抽象”这个词的拉丁文为 *abstractio*，它的原意是排除、抽出。中文的抽象，除了抽出事物的本质、共性的特征外，还具有“玄之又玄”的意味。这个意味在计算机科学中是不存在的。例如，在中文语境中很多人把凡是不能被人们的感官所直接把握的东西，也就是通常所说的“看不见，摸不着”的东西，叫做“抽象”；有的则把“抽象”作为孤立、片面、思想内容贫乏空洞的同义词。这些是“抽象”的引申和转义。

### 1.1.3 七桥问题

本节通过一个实例来理解一般问题求解的抽象过程。

1736 年 Euler 访问普鲁士（Prussia）的哥尼斯堡（Konigsberg）（今俄罗斯加里宁格勒，Kaliningrad Russia）。哥尼斯堡城中有一条名叫普雷格尔（Pregel）的河流，河中有两个岛，把城市分割成北区、岛区、南区和东区。连接两个岛屿和两岸有七座桥（图 1-1）。河中的岛上古老的哥尼斯堡大学，学生们经常沿河过桥散步。在散步的过程中，爱动脑筋的人们提出了一个问题：一个散步者能否一次走遍七座桥，而且

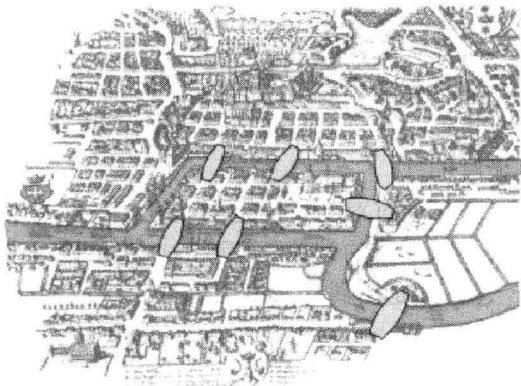


图 1-1 哥尼斯堡城中的七座桥