

Computer Science Distilled
Learn the Art of Solving Computational Problems

计算机科学精粹

【巴西】沃德斯頓·費雷拉·菲爾多 著
蔣楠 译

- 摒弃传统计算机科学教材的枯燥方式
- 将抽象理论具体化、复杂问题简单化
- 传授实现高效程序设计至关重要的基础知识



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

TURING

图灵程序设计丛书

计算机科学精粹



【巴西】沃德斯頓·費雷拉·菲爾多 / 著

蔣楠 / 譯

Computer Science Distilled

Learn the Art of Solving Computational Problems

人民邮电出版社

北 京

图书在版编目 (C I P) 数据

计算机科学精粹 / (巴西) 沃德斯頓·费雷拉·菲尔
多著 ; 蒋楠译. — 北京 : 人民邮电出版社, 2019. 1
(图灵程序设计丛书)
ISBN 978-7-115-49919-6

I. ①计… II. ①沃… ②蒋… III. ①计算机科学
IV. ①TP3

中国版本图书馆CIP数据核字 (2018) 第249565号

内 容 提 要

本书以浅显易懂的语言、简明扼要的形式介绍计算机科学领域的重要知识
点, 较少涉及学术概念, 着力将抽象理论具体化、复杂问题简单化。主要内容
包括逻辑、计数等基本概念, 数据类型, 算法, 计算机体系结构, 程序设计,
等等。

本书既适合计算机专业技术人员, 也适合对计算机科学感兴趣的普通读者。

-
- ◆ 著 [巴西] 沃德斯頓·费雷拉·菲尔多
译 蒋楠
责任编辑 朱巍
责任印制 周昇亮
- ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件: 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京瑞禾彩色印刷有限公司印刷
- ◆ 开本: 880×1230 1/32
印张: 5.25
字数: 157千字 2019年1月第1版
印数: 1—4 000册 2019年1月北京第1次印刷
著作权合同登记号 图字: 01-2018-4176号

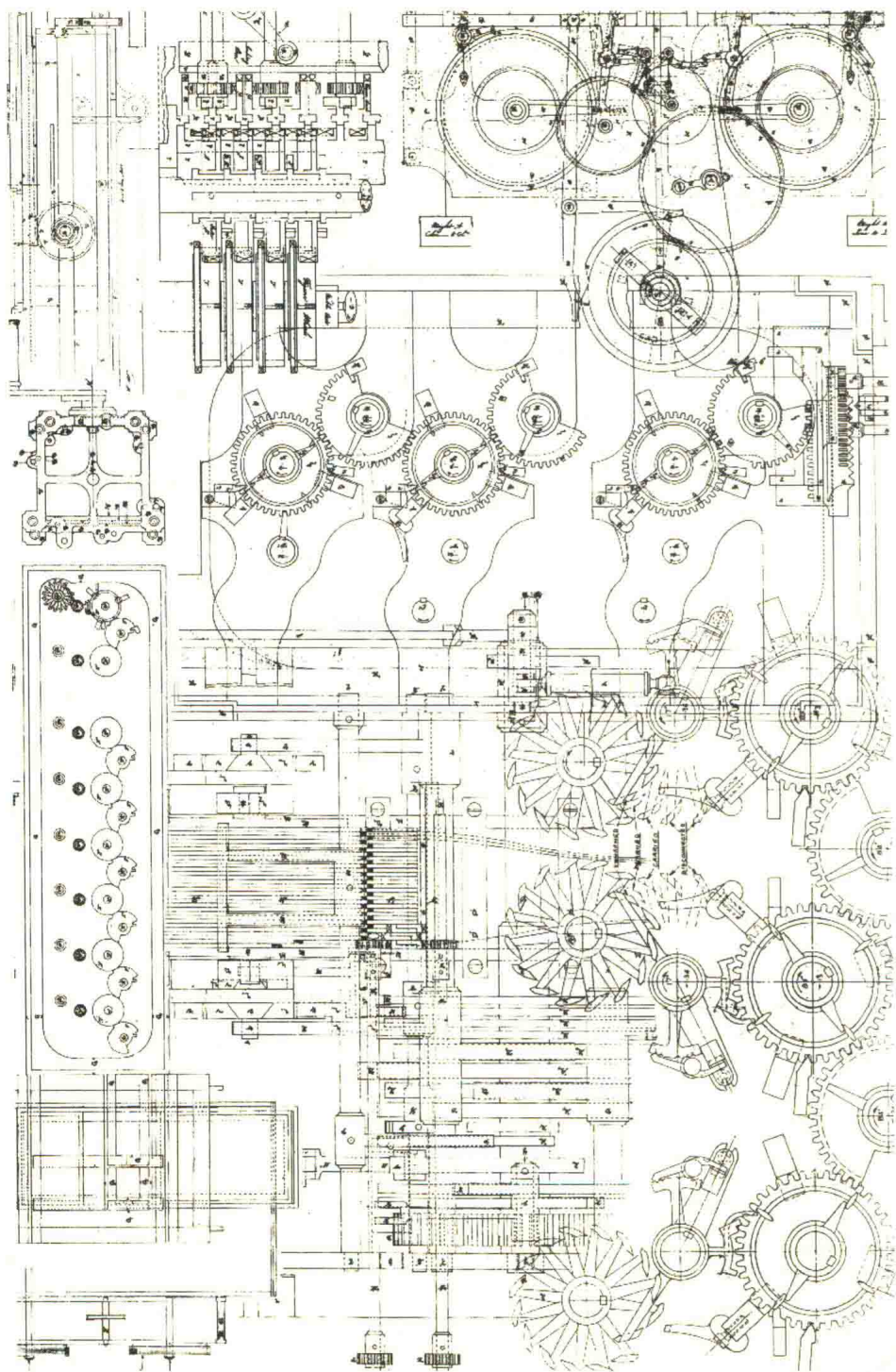


定价: 49.00元

读者服务热线: (010)51095186转600 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号



查尔斯·巴贝奇的分析机原理图

版权声明

Authorized translation from the English language edition, entitled *Computer Science Distilled: Learn the Art of Solving Computational Problems*, First Edition, by Wladston Ferreira Filho, Published by Code Energy LLC Copyright © 2017.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from the author.

CHINESE language edition published by Posts & Telecom Press, Copyright ©2019.

英文原版由 Code Energy LLC 出版，2017。

简体中文版由人民邮电出版社出版，2019。英文原版的翻译得到 Code Energy LLC 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——Code Energy LLC 的许可。

版权所有，未得书面许可，本书的任何部分和全部不得以任何形式重制。

我知道 2 加 2 等于 4，如果能证明这一点我会很高兴，但必须承认，如果能让 2 加 2 等于 5，那么我会更高兴。

——拜伦勋爵

取自 1813 年致未婚妻安娜贝拉的信函。
他们的女儿埃达·洛夫莱斯是第一位程序员。

译者序

本书英文版封面图片根据 1845 年的一份分析机原理图绘制。分析机是历史上第一种可编程计算机，也是其发明者查尔斯·巴贝奇得以在计算机史上留名的主要作品。

巴贝奇堪称“跨界”高手。他颇具数学天分，曾受聘担任剑桥大学卢卡斯教授；他被誉为计算机先驱，为差分机与分析机的发明耗尽一生心血；他还是当时知名的经济学家，曾撰写 19 世纪 30 年代最有影响力的经济学著作《论机械和制造业的经济》。

巴贝奇同样是一位理想主义者。在英国政府停止对差分机项目的资助后，巴贝奇依然锲而不舍，开始设计功能更为强大的分析机。它分为运算单元与存储单元，通过打孔卡进行输入，并使用与汇编语言类似的编程语言。分析机将运算、存储、I/O 功能相互分离，与如今的计算机有异曲同工之妙。

然而，在缺乏政府资助的情况下制造分析机，无异于纸上谈兵，最终留下的只有数千页设计手稿。郁郁不得志的巴贝奇于 1871 年去世，报纸甚至还在讣告中嘲笑了他的失败。

但巴贝奇并非没有知音。在都灵访问期间，他鼓励后来担任意大利首相的路易吉·梅纳布雷亚撰写一篇有关分析机的论文。这份以法语写就的论文于 1842 年出版，后来被著名诗人拜伦的女儿埃达·拜伦译为英文。埃达对巴贝奇的才华颇为仰慕，两人于 1833 年相识后一直保持联系。埃达继承了母亲的数学天分，是为数不多能深刻理解巴贝奇思想的人，她甚至变卖自己的珠宝以支持分析机的制造。

埃达并非简单地翻译梅纳布雷亚的论文，她还添加了几乎达到原文长度四倍的注记，详细描述了使用分析机计算伯努利数的方法，并设计了世界上第一个计算机程序。埃达由此被视为历史上第一位程序员。为纪念

她的贡献，美国国防部将 1980 年发布的一门编程语言命名为 Ada。

《计算机科学精粹》堪称一本“网红”书，出版之后好评如潮，许多亚马逊用户毫不吝啬地打出五星高分。本书适合各个层次、各种背景的读者阅读，无论是经验不足的新人，还是科班出身的老手，想必都能从中找到适合自己的内容。学习本非枯燥之事，但一本浅显易懂的入门图书的确有事半功倍之效。

在北美的院校中，某些考试允许携带 cheat sheet（中文可称为“备忘单”或“速查表”），学生可以将自己认为重要的公式或知识点写在上面。从某种意义上说，《计算机科学精粹》就是这样一本具有 cheat sheet 性质的书。与图灵推出的《算法图解》类似，本书以浅显易懂的语言梳理了计算机科学领域的重要知识点，着力将抽象的理论具体化、复杂的问题简单化。当然，亦想抛砖引玉，希望在唤起读者对计算机科学的兴趣后，能深入阅读其他资料。

非常感谢北京图灵文化发展有限公司的朱巍老师给予译者的信任，以及李冰编辑为本书付梓所做的辛勤努力。虽然译者尽力而为，但水平有限，疏漏之处在所难免。恳请读者不吝赐教，提出宝贵的意见和建议。译者的联系方式：milesjiang314@gmail.com。

蒋楠

2018 年 8 月于温哥华

朋友是我们为自己选择的亲人。本书献给我的朋友 Rômulo、Léo、Moto 与 Chris，他们不断敦促我“完成这部该死的作品”。

前言

每个人都应该学习计算机编程，因为它教会你如何思考。

——史蒂夫·乔布斯

计算机以前所未有的力量改变了世界，一门新学科随之兴起，这就是计算机科学。它揭示了如何利用计算机解决问题，帮助我们充分发挥计算机的潜能。在这门学科的指引下，我们取得了难以置信的成就。

计算机科学无处不在，但学校传授的仍然是枯燥的理论，不少程序员甚至从未研究过它。然而，计算机科学对于实现高效的程序设计至关重要。我的一些朋友很难聘用到优秀的程序员——计算机虽然功能强大，可以驾驭它的人却不多。

我希望通过这本书推动读者高效地使用计算机。本书将以简明扼要的形式介绍计算机科学的知识，尽量少涉及学术概念。但愿计算机科学能在读者心中扎根，并提高读者编写代码的水平。



图1 “计算机问题”（取自 <https://xkcd.com/>）

目标读者

对于希望采用高效方法解决问题的读者，本书将是不二之选。编程经验并非必需，如果读者曾经写过代码，也了解 `for` 与 `while` 这样的基本编程语句，阅读本书将不会遇到任何障碍。不熟悉计算机科学的读者可以通过 Codecademy^① 进行学习，其在线课程提供一周的免费试听服务。而对具备计算机科学经验的读者来说，本书能有效地巩固所学知识。

计算机科学并非只和学者有关

这是一部关于计算思维的作品，适合所有人阅读。读者将学习如何把问题转换为可计算的系统，并在日常生活中应用计算思维：预取和缓存能简化打包过程，而并行有助于提高烹饪速度。另外，读者的代码会变得很棒！

愿原力与你同在。^② 😊

① 许多平台都提供不错的在线课程，涵盖 Web 开发、数据处理、人工智能、深度学习等多个领域。lynda.com、Udacity 都是北美较为知名的在线教育平台。——译者注

② 《星球大战》中绝地武士在分别时表示“再见”的祝福语，后引申为现实世界中粉丝之间的祝福语。——译者注

目 录

第 1 章 预备知识	1
1.1 想法	1
1.1.1 流程图	2
1.1.2 伪代码	3
1.1.3 数学模型	4
1.2 逻辑	5
1.2.1 运算符	6
1.2.2 布尔代数	8
1.2.3 真值表	9
1.2.4 逻辑在计算中的应用	12
1.3 计数	13
1.3.1 乘法	13
1.3.2 排列	14
1.3.3 具有相同项的排列	15
1.3.4 组合	16
1.3.5 求和	17
1.4 概率	19
1.4.1 对结果计数	19
1.4.2 独立事件	20
1.4.3 互斥事件	20
1.4.4 对立事件	21
1.4.5 赌徒谬误	21
1.4.6 高级概率	21
1.5 小结	22

第2章 复杂度	23
2.1 时间计算	25
2.2 大O符号	28
2.3 指数	29
2.4 内存计算	30
2.5 小结	31
第3章 策略	33
3.1 迭代	33
3.2 递归	36
3.3 蛮力法	38
3.4 回溯法	40
3.5 启发法	43
3.5.1 贪心法	43
3.5.2 利用贪心法求解电网问题	45
3.6 分治法	46
3.6.1 利用分治法求解排序问题	46
3.6.2 利用分治法求解最佳交易问题	49
3.6.3 利用分治法求解背包问题	50
3.7 动态规划	51
3.7.1 利用记忆化求解斐波那契数	52
3.7.2 利用记忆化求解背包问题	52
3.7.3 利用自底向上法求解最佳交易问题	53
3.8 分支定界法	54
3.8.1 上界与下界	55
3.8.2 背包问题中的上界与下界	56
3.9 小结	58
第4章 数据	59
4.1 抽象数据类型	60
4.2 常见抽象	62
4.2.1 基本数据类型	62

4.2.2	栈	62
4.2.3	队列	63
4.2.4	优先队列	63
4.2.5	列表	64
4.2.6	排序列表	64
4.2.7	映射	65
4.2.8	集合	65
4.3	数据结构	65
4.3.1	数组	66
4.3.2	链表	67
4.3.3	双向链表	68
4.3.4	数组与链表的比较	68
4.3.5	树	69
4.3.6	二叉查找树	70
4.3.7	二叉堆	73
4.3.8	图	74
4.3.9	散列表	74
4.4	小结	75
第 5 章	算法	77
5.1	排序	77
5.2	搜索	79
5.3	图	80
5.3.1	图的搜索	80
5.3.2	图着色	83
5.3.3	寻路	83
5.3.4	PageRank	86
5.4	运筹学	86
5.4.1	线性最优化问题	87
5.4.2	网络流问题	88
5.5	小结	89

第 6 章 数据库	91
6.1 关系数据库	92
6.1.1 关系	92
6.1.2 模式迁移	95
6.1.3 SQL	95
6.1.4 索引	97
6.1.5 事务	99
6.2 非关系数据库	99
6.2.1 文档存储	100
6.2.2 键值对存储	101
6.2.3 图数据库	102
6.2.4 大数据	103
6.2.5 SQL 与 NoSQL 的比较	103
6.3 分布式数据库	104
6.3.1 单主机复制	104
6.3.2 多主机复制	105
6.3.3 分片	105
6.3.4 数据一致性	107
6.4 地理数据库	107
6.5 序列化格式	108
6.6 小结	109
第 7 章 计算机	111
7.1 体系结构	111
7.1.1 存储器	112
7.1.2 CPU	114
7.2 编译器	118
7.2.1 操作系统	121
7.2.2 编译优化	121
7.2.3 脚本语言	122
7.2.4 反汇编与逆向工程	123
7.2.5 开源软件	124

7.3 存储器层次结构	125
7.3.1 处理器与存储器之间的鸿沟	125
7.3.2 时间局部性与空间局部性	126
7.3.3 一级缓存	127
7.3.4 二级缓存	127
7.3.5 第一级存储器与第二级存储器	128
7.3.6 外部存储器与第三级存储器	130
7.3.7 存储技术的发展趋势	130
7.4 小结	131
第 8 章 程序设计	133
8.1 语言学	133
8.1.1 值	134
8.1.2 表达式	134
8.1.3 语句	135
8.2 变量	136
8.2.1 变量类型	136
8.2.2 变量作用域	137
8.3 范式	138
8.3.1 命令式编程	138
8.3.2 声明式编程	140
8.3.3 逻辑编程	144
8.4 小结	145
附录	147
结语	151
后记	152

第1章

预备知识

计算机科学并非一门研究机器的学科，如同天文学并非研究望远镜一样。从本质上讲，数学与计算机科学具有统一性。

——艾兹赫尔·戴克斯特拉^①

计算机只能处理分解成块的问题，因此我们需要具备一定的数学知识。不过无须紧张，数学并非高深莫测，编写优秀的代码也很少要用到复杂的方程。这一章将介绍求解问题所需的基本知识，包括：

💡 采用流程图与伪代码对想法进行建模

✓ 根据逻辑判断对错

100 对事物进行计数

🔴 安全地计算概率

掌握这些知识后，我们就能将自己的想法转换为可供计算机执行的解决方案。

1.1 想法

面对复杂的问题时，请让大脑保持最佳状态，将所有重要内容写下来。我们的大脑很容易被各种事实和想法所淹没，“好记性不如烂笔头”在众多组织方法中占有重要地位。为此，这一节将讨论几种实现方法。首先介绍用于表示进程的流程图，然后利用伪代码编写可供实际编程使用的进程，并尝试通过数学工具对一个简单的问题进行建模。

^① 艾兹赫尔·戴克斯特拉 (Edsger Dijkstra, 1930—2002)，荷兰计算机科学家，其贡献涵盖编译器、操作系统、分布式系统、软件工程、编程语言、图论等多个领域，数据结构中的最短路径算法——戴克斯特拉算法就是以他的名字命名的。1972年，戴克斯特拉因在编程语言方面的贡献而获得图灵奖。——译者注

1.1.1 流程图

在讨论相互之间的协作过程时，维基人^①创建了一种随着讨论的进行而更新的流程图。对所提出的内容了然于心有助于讨论。

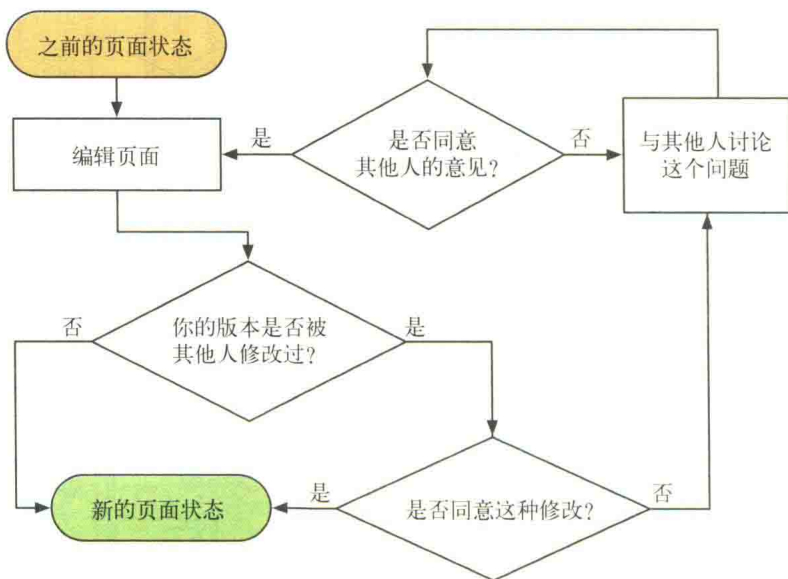


图 1-1 维基百科编辑流程

与上面的编辑流程类似，计算机代码本质上是一种进程。程序员通常使用流程图来编写计算进程。为便于他人理解，绘制流程图时应遵循以下原则^②：

- 将状态步骤和指令步骤置于矩形框内；
- 将决策步骤（对给定的条件进行判断）置于菱形框内；
- 不要将指令步骤与决策步骤放在一起；
- 使用箭头连接各个顺序步骤；
- 标明进程的开始和结束。

① 维基人指在维基百科上编写条目的贡献者，女性维基人有时候也被称为“薇姬人”。——译者注

② 国际标准化组织甚至专门制定了一项称为 UML（统一建模语言）的标准来定义软件系统图的绘制。