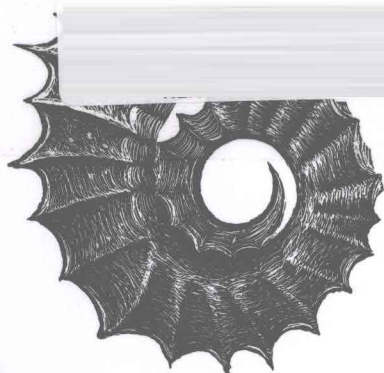


Mastering Algorithms with C

数据结构和算法领域最具特色的著作之一，公认的权威经典！

算法精解

C语言描述



O'REILLY®



机械工业出版社
China Machine Press



HZ BOOKS

Kyle Loudon 著

肖翔 陈舸 译

算法精解：C语言描述

Kyle Loudon 著

肖翔 陈舸 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

O'Reilly Media, Inc. 授权机械工业出版社出版

机械工业出版社

图书在版编目 (CIP) 数据

算法精解: C语言描述/ (美) 劳顿 (Loudon, K.) 著; 肖翔, 陈舸译.

— 北京: 机械工业出版社, 2012.8

(O'Reilly精品图书系列)

书名原文: Mastering Algorithms with C

ISBN 978-7-111-39426-6

I. 算… II. ①劳… ②肖… ③陈… III. ①算法语言—程序设计 ②C语言—程序设计
IV. TP312

中国版本图书馆CIP数据核字 (2012) 第187376号

北京市版权局著作权合同登记

图字: 01-2012-1676号

Copyright © 1999 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Machine Press, 2012. Authorized translation of the English edition, 1999 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由O'Reilly Media, Inc. 出版1999。

简体中文版由机械工业出版社出版 2012。英文原版的翻译得到O'Reilly Media, Inc.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——O'Reilly Media, Inc.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。

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

书 名/ 算法精解: C语言描述

书 号/ ISBN 978-7-111-39426-6

责任编辑/ 谢晓芳

封面设计/ Hanna Dyer, 张健

出版发行/ 机械工业出版社

地 址/ 北京市西城区百万庄大街22号 (邮政编码 100037)

印 刷/ 藁城市京瑞印刷有限公司

开 本/ 178毫米×233毫米 16开本 26.25印张

版 次/ 2012年9月第1版 2012年9月第1次印刷

定 价/ 79.00元 (册)

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

购书热线: (010)68326294; 88379649; 68995259

投稿热线: (010)88379604

读者信箱: hzsj@hzbook.com

O'Reilly Media, Inc. 介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

译者序

不管你是否承认，无论哪种软件开发项目，几乎所有的程序员、开发者在日常工作中都要同数据结构和算法打交道。当我们阅读源码，对大型软件项目进行层层抽丝剥离之后，呈现在我们面前的不再是复杂的层次结构和模式，而是回归到了程序的本质——数据结构和算法上来。有关数据结构和算法题材的经典书籍有很多，如《计算机程序设计艺术》、《算法导论》等，但本书绝不同于之前的相关书籍。当我们在学习数据结构和算法时，往往花费了大量的时间纠结于各种公式和理论证明上，学究气息过于浓厚而少了几分实践感。相关主题的书籍多是以伪码的形式来表达算法思路，缺乏同软件开发实践的结合。试问，当我们辛辛苦苦“学会”了一些算法和数据结构，但在实际编程中往往要么一写就错（伪码毕竟只是用来表达算法思路，但具体编程时会遇到很多实现上的细节问题），要么面对问题不知如何下手解决（因为没有实际的应用，不知道如何对问题建模），此时会不会觉得自己白学了？毕竟，我们学到的东西要懂得利用才能体现价值。将一个实际问题同我们学到的算法和数据结构相结合起来，这正是软件开发中的一项重要技能——抽象建模能力。

本书最大的特点是理论与实践相结合，这主要体现在以下几点上：

1. 书中第一部分从讲述相关的C语言基础知识和算法分析方法入手，在随后的章节中采用软件工程中的良好准则，结合作者的实践经验将基于接口的C程序设计理念贯穿全书。使得本书中的数据结构和算法实现能够以接口的形式充分得到复用。
2. 书中的代码实现主要以教学为目的，但也同样考虑到了实现效率的问题。对于实现方案的选择和取舍，都有详细的说明和解释。每章末尾的“问与答”将加深读者对相关章节的理解。

3. 除了对数据结构和算法本身的介绍，书中所有的应用举例都来自于真实的应用。这绝不是一般的练习题，而是算法和数据结构在真实世界中的应用。包括操作系统中的页帧管理、页面置换算法、表达式处理等等。通过实际的例子向读者展示了数据结构和算法的威力，有助于培养我们抽象建模的能力，从而更有效的将学到的知识用于解决实际问题中。

总的来说，这是一本对“打基础”很有帮助的书。书中对常见的数据结构如链表、栈、队列、集合、哈希表、树、堆、图都做了详细的分析并给出了具体的实现。算法方面除了最为常见的排序和检索外，还有数值计算、数据压缩、数据加密、几何计算等方面的主题。

当译者第一次看到本书的英文版时，立刻被本书独有的风格所吸引。再看出版时间，居然已超过了10年之久。令人奇怪的是，这样一本优秀的技术类图书在这么长的时间里居然从未引进到国内，无论是译作还是影印版都难觅其踪。因此我们鼓足勇气开始了本书的翻译。感谢机械工业出版社引进了本书，这才得以让本书的中译本得以和大家见面。由于时间仓促，且水平有限，在翻译过程中难免会出现一些错误，希望读者能批评指正。

译者

2012年8月3日

目录

前言	1
第1部分 预备知识	
第1章 概述	9
数据结构简介	10
算法简介	11
小酌软件工程	14
如何使用本书	15
第2章 指针操作	16
指针基础	17
存储空间分配	18
数据集合与指针的算术运算	20
作为函数参数的指针	22
泛型指针与类型转换	25
函数指针	28
问与答	28
相关主题	30
第3章 递归	31
基本递归	32
尾递归	35

问与答	37
相关主题	39
第4章 算法分析	40
最坏情况分析	41
O表示法	41
计算的复杂度	43
实例分析：插入排序	46
问与答	47
相关主题	48
 第2部分 数据结构	
第5章 链表	51
单链表介绍	52
单链表接口的定义	53
单链表的实现与分析	56
使用链表的例子：页帧管理	61
双向链表介绍	63
双向链表接口的定义	64
双向链表的实现与分析	67
循环链表介绍	73
循环链表接口的定义	74
循环链表的实现与分析	76
使用循环链表的例子：第二次机会页面置换法	79
问与答	82
相关主题	84
 第6章 栈和队列	85
栈的描述	86
栈的接口定义	87
栈的实现与分析	88

队列的描述	91
队列的接口定义	91
队列的实现与分析	93
队列示例：事件处理	95
问与答	96
相关主题	97
第7章 集合	98
集合介绍	99
集合的性质	100
集合接口的定义	102
集合抽象数据类型的实现和分析	105
Set示例：集合覆盖	112
问与答	116
相关主题	117
第8章 哈希表	119
链式哈希表的描述	121
链式哈希表的接口定义	124
链式哈希表的实现与分析	126
链式哈希表的例子：符号表	131
开地址哈希表的描述	133
开地址哈希函数的接口定义	136
开地址哈希表的实现与分析	138
问与答	144
相关主题	145
第9章 树	146
二叉树介绍	148
二叉树的接口定义	151
二叉树的实现与分析	155
二叉树示例：表达式处理	161
二叉搜索树介绍	165

二叉搜索树的接口定义	166
二叉搜索树的实现与分析	168
问与答	185
相关主题	187
第10章 堆和优先队列	188
堆的描述	189
堆的接口定义	190
堆的实现与分析	191
优先队列的描述	199
优先队列的接口定义	199
优先队列的实现与分析	201
优先队列的示例：包裹分拣	202
问与答	203
相关主题	205
第11章 图	206
图的描述	207
图的接口定义	214
图的实现与分析	217
关于图的应用举例：计算网络跳数	225
关于图的应用举例：拓扑排序	229
问与答	232
相关主题	234
 第3部分 算法	
第12章 排序和搜索	237
插入排序的描述	239
插入排序的接口定义	239
插入排序的实现与分析	240
快速排序的描述	242

快速排序的接口定义	243
快速排序的实现与分析	243
快速排序的例子：目录列表	247
归并排序的描述	249
归并排序的接口定义	249
归并排序的实现与分析	250
计数排序的描述	254
计数排序的接口定义	254
计数排序的实现与分析	254
基数排序的描述	257
基数排序的接口定义	257
基数排序的实现与分析	258
二分查找的描述	260
二分查找的接口定义	260
二分查找的实现与分析	261
二分查找的例子：拼写检查器	263
问与答	264
相关主题	266
第13章 数值计算	267
多项式插值法	268
多项式插值的接口定义	272
多项式插值的实现与分析	272
最小二乘估计法	274
最小二乘估计的接口定义	276
最小二乘估计的实现和分析	276
方程求解介绍	277
方程求解的接口定义	281
方程求解的实现与分析	282
问与答	283
相关主题	284

第14章 数据压缩	285
位操作的描述	288
位操作的接口定义	288
位操作的实现与分析	289
霍夫曼编码的描述	292
霍夫曼编码的接口定义	295
霍夫曼编码的分析与实现	296
霍夫曼编码的例子：网络优化	306
LZ77的描述	308
LZ77的接口定义	311
LZ77的实现与分析	312
问与答	321
相关主题	322
第15章 数据加密	324
DES算法介绍	326
DES的接口定义	334
DES算法的实现和分析	334
DES应用举例：分组加密模式	341
RSA算法介绍	344
RSA的接口定义	347
RSA算法的实现与分析	348
问与答	350
相关主题	352
第16章 图算法	354
最小生成树的描述	357
最小生成树的接口定义	358
最小生成树的实现与分析	359
最短路径的描述	363
最短路径的接口定义	364
最短路径的实现与分析	365

最短路径的例子：路由表	369
旅行商问题的描述	372
旅行商问题的接口定义	374
旅行商问题的实现与分析	374
问与答	377
相关主题	378
第17章 几何算法	380
测试线段是否相交	382
测试线段是否相交的标准方法	383
检测线段是否相交的接口定义	385
检测线段是否相交的实现与分析	385
凸包简介	387
Jarvis's March	387
凸包的接口定义	389
凸包的实现与分析	389
球面弧长	392
求解球面弧长的接口定义	395
求解球面弧长的实现和分析	395
球面弧长的应用举例：地球上两点之间的近似距离	396
问与答	398
相关主题	400

前言

当我第一次打算要写这本书时，我马上就想到应该交给O'Reilly来负责出版。O'Reilly是我联系的第一家出版商，也是我最希望合作的一家，因为他们出版的图书传统上都会刻有“实事求是”的烙印。可是对于一本有关数据结构和算法的技术类图书，通常人们对它的态度就不会是这样。通常当人们学习数据结构和算法时，都要花费很多时间去证明算法的正确性。结果就是，很多关于这个主题的书都带有很重的学究气息，而真正的细节（比如具体的实现和应用）则省略了。本书讲解数据结构和算法的工作原理和原因，它们在真实世界中的应用（包括许多例子），以及它们的具体实现过程。只有在必要的时候才会以严谨的数学方式解释说明。

自然，我很高兴O'Reilly出版社看到了本书在这个主题上所涵盖内容的价值。前言部分包含我认为你会喜欢这本书的一些理由。前言也包含本书中的一些代码规范，书中定义的一些约定以及对参与了本书创作过程的工作人员的衷心感谢。

本书结构

本书分为3部分。第1部分由介绍性的内容组成，它们对于读者理解接下来的章节有很大帮助。第2部分介绍了计算机科学中最基本的几种数据结构。第3部分介绍用于解决常见问题的各种算法。接下来的内容将以更详尽的方式来描述和说明每一部分的主题，以及每个部分的各章的提要。

第1部分

第1部分包括第1~4章。第1章介绍数据结构和算法的概念，并给出使用数据结构和算法

的原因。这一章也介绍一些与软件工程相关的思想和主题，这些思想会贯穿全书。第2章讨论一系列与指针相关的主题。指针在本书中大量出现，因此这一章可作为对指针的回顾和复习。第3章介绍递归，这是一种在许多数据结构和算法中都会用到的技术。第4章介绍算法的分析方法。这一章中的内容用来分析本书中介绍的算法。

第2部分

第2部分包括第5~11章。第5章介绍多种形式的链表，包括单链表、双向链表以及循环链表。第6章介绍栈和队列，它们分别按照后进先出以及先进先出的方式来存储和返回数据。第7章介绍集合以及用来描述集合这种数据结构的数学知识。第8章介绍链式和开地址哈希表，还包含如何选择一个好的哈希函数以及如何解决哈希碰撞的内容。第9章介绍二叉树和AVL树，也包含多种树的周游算法。第10章介绍堆和优先级队列，这种数据结构有助于快速确定集合中的最大或最小元素。第11章介绍图，以及在许多图算法中常用到的两种重要算法：广度优先查找和深度优先查找。

第3部分

第3部分包含第12~17章。第12章涵盖多种排序算法，包括插入排序、快速排序、归并排序、计数排序以及基数排序，这一章也介绍二分查找。第13章涵盖数值计算方法，包括多项式插值算法、最小二乘估计以及用来解方程的牛顿迭代法。第14章介绍关于数据压缩的算法，包括霍夫曼编码以及LZ77。第15章讨论用于数据加密的DES和RSA算法。第16章涵盖图算法，包括用于确定最小生成树的Prim算法、确定最短路径的Dijkstra算法，以及解决旅行商问题的算法。第17章介绍几何算法，包括检测两条线段是否相交、凸包计算以及计算球面上两点间的弧长。

本书主要特点

这里列出的一些特点使我相信本书能够成为一本涵盖数据结构和算法方面的独特技术类图书。

每一章的格式保持一致

每一章（不包括第1章中的介绍性内容）都遵循一致的格式。根据读者当前的需要，这种格式使得本书大部分内容都可以以教材或参考资料的形式来阅读。

明确的主题和应用

每一章（除了第1章外）都以一个概要性的介绍开始，接着是一系列明确的主题以及与它们相关的真实应用介绍。

对每个操作、算法和例子的分析

本书针对抽象数据类型的每个操作、第12~17章的每种算法以及每个示例都提供详细的分析。关于算法分析的技术在第4章中介绍。

真实示例，不只是简单的练习

本书中所有的示例都来自真实的应用，不只是一般的练习题。这样的示例更令人兴奋，也能够从中学到更多的知识。

具体实现都采用正式的代码

所有的实现都用C语言编写而不是用伪代码来表示。这种方式的好处在于当实现许多数据结构和算法时，有很多细节问题是伪代码所不能解决的。

问与答引人深思

每一章的末尾（除了第1章外）都会有一系列的问题和对应的回答。这些问与答着重强调了这一章中的重要思想，而且还能使读者接触到一些附加的主题。

列出了相关的主题以进一步研究

每一章的最后一节（除了第1章外）都会列出一系列与本章主题相关的内容，以便读者进一步探讨。每一个相关主题都以一段简短的介绍来呈现。

大量的交叉引用和标注

交叉引用和标注将主题在一处提及，然后在别处介绍。因而能够方便地定位其他信息。

精心组织主题和相关应用

某一章中的数据结构和算法会采用其他章节介绍过的数据结构和算法来实现。因而，这可以作为如何利用其他数据结构和算法的例子。所有的相关依赖都通过交叉引用和标注的形式精心标记。

除了涵盖最重要的主题之外还有更多其他的内容

本书涵盖计算机科学领域中最基本的数据结构和算法。本书还介绍其他几个并不常出现在数据结构和算法类图书中的主题。其中包括数值计算方法、数据压缩、数据加密以及几何算法。

关于本书中的代码

本书中的所有实现都采用C语言编写。选择C语言是因为它仍然是当今使用最广泛的通用编程语言。C语言也是最好的编程语言之一，不仅用来探索数据结构和算法的细节，同时还能够使用在较高层次上。关于本书中的代码，有些事情需要注意。

所有的代码以教学目的为优先

代码注重运行效率上的考虑，但所有代码的主要目的是以清晰的形式阐述相关的主题。

所有的代码都在4种平台上做过完整的测试

用于测试的平台包括：HP-UX 10.20、SunOs 5.6、Red Hat Linux 5.1以及DOS/Windows NT/95/98。可以从附赠光盘的readme文件中获取更多的信息。

头文件中记录了所有的公共接口

每个实现都包含一个头文件以归档公有接口。本书给出了代码头文件。但是，只包含函数原型的头文件并没有给出。（比如，示例12-1包含sort.h，但并没有给出头文件，因为它只包含各个排序函数的原型而已。）

静态函数当做私有函数来使用

静态函数拥有文件作用域，只在定义它的文件内有效，因此把静态函数作为私有函数使用。某个特定于数据结构或算法实现的函数都不属于公有接口的一部分。

命名规则适用于全书所有的代码

常量全部以大写字母来定义。数据类型和全局变量以大写字母开头。局部变量以小写字母开头。抽象数据类型的操作以类型名开头，采用小写字母形式，紧跟着是一条下划线，然后是操作名的小写形式。

所有的代码都包含大量注释

注释可以帮助开发者不用读太多代码就能够理解代码的逻辑。当需要在正文中联系代码和解释时这非常有用。

用typedef重命名的结构体名就是它本身

结构体的名称总是在typedef之后的别名上再加一条下划线。对于有自我引用行为的结构体（比如第5章中的链表），命名结构体自身是很有必要的。为了保持一致性，本书所有的结构体都采用这种命名方式。

所有void函数都显式包含return语句

尽管这不是必需的，但这有助于快速定位函数的返回位置，而不必费力地去匹配的大括号。

约定

各阶段的计算机用户都可以识别出本书中使用的大部分约定。但是，这里还需要再做一点说明。

粗体斜体 (***Bold italic***)

非内部数学函数和数学变量以这种字体显示。