INTRODUCTION TO
COMBINATORIAL OPTIMIZATION
Second Edition

组合优化导论

（第二版）

越民义 李荣珩 著

科 学 出 版 社

华罗庚-吴文俊数学出版基金资助项目

组合优化导论

（第二版）

越民义　李荣珩　著

科学出版社

内 容 简 介

这是一本介绍组合优化这门学科的书，本书可看成三个部分，第一部分包括第 1 章、第 2 章和第 3 章，通过排序问题中较典型的例子介绍什么是组合优化中的可解问题. 第二部分即第 5 章，是启发式算法方面的，这主要是韩继业教授的工作. 第三部分由第 4 章、第 6 章和第 7 章组成，是近似算法方面的，其中第 4 章主要叙述装箱问题的一些经典结果，包括了作者在这方面的工作；第 6 章是关于 Steiner 比猜想的进展报告；第 7 章介绍 Coffman 等提出的多重算法. 后两章的结果都是作者给出的.

本书适合高等院校数学、管理、信息处理等有关专业的学生、教师和研究人员阅读.

图书在版编目（CIP）数据

组合优化导论/越民义，李荣珩著. —2 版. —北京：科学出版社，2014. 5

ISBN 978-7-03-040540-1

Ⅰ.①组… Ⅱ.①越… ②李… Ⅲ.①组合数学 Ⅳ.①O157

中国版本图书馆 CIP 数据核字(2014) 第 089899 号

责任编辑：徐园园　赵彦超／责任校对：桂伟利
责任印制：赵德静／封面设计：耕者设计工作室

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

双青印刷厂 印刷
科学出版社发行　各地新华书店经销
*
2001 年 12 月浙江科学技术出版社第一版
2014 年 5 月第　二　版　开本：720 × 1000 1/16
2014 年 5 月第一次印刷　印张：15 1/2
字数：296 000

定价：78.00 元

(如有印装质量问题，我社负责调换)

理工大学 藏書 图書馆

第二版前言

严格说来, 本书的内容与书名不是很相称. “导论”一词, 顾名思义, 是要将读者引入一个新的领域, 使其通过本书可以对“组合优化”这门学科有一较为全面的了解. 但本书在这一点上仍有欠缺之处. 相对说来, 组合优化是一门较为新颖的学科. 它要研究的主要对象是哪些? 它用来解决问题的主要工具是哪些? 一个问题得到了解决指的是什么? (图灵机虽然在概念上可以用来将问题分类, 但却无法用于处理具体问题, 例如, 是否可用它来决定“求圆周率”是否属于 P 类)? 限于作者的水平, 我们所能做的是: 从文献中寻找一些已经公认解决了的问题, 让读者明白, 这一类的问题就是组合优化问题. 所谓解决, 就是指对于一个给定的问题, 找到一个算法, 当问题的有关参数已经给定之后, 就可根据算法得出问题的解. 此外, 本书中还列举一些求近似解的问题, 它们主要来自本书几位作者的工作. 因此本书的内容是很局限的. 好在 Korte 与 Vygen 所著的《组合最优化: 理论与算法》(第四版) 的中文译本已于近日在科学出版社发行. 该书是一本全面地、系统性地介绍组合优化的书. 两书内容没有覆盖重复之处.

本书第一版于 2001 年由浙江科学技术出版社出版. 本书出版后就没有重印过, 其间, 有读者来信欲购此书而不可得. 我曾与出版社的周伟元同志联系过, 希望售我一些, 可以赠送读者. 周随即寄来几本, 并告诉我, 这是书库中仅存的. 未能满足某些读者的需求, 我深感内疚. 现在科学出版社希望出版此书, 我自然感到高兴.

不言而喻, 第一版问世已经十几年, 书中内容自应修改. 前面三章, 除 2.4 节外, 其余均未作更改, 至于 2.4 节, 由于第一版所根据的文章唐国春教授等发现有错, 需要改正. 现书中所载即出自他们之手. 第 4 章大部分是新的. 定理 4.6.1 的结论改进到 $\mathrm{FFD}(L) \leqslant \dfrac{11}{9}\mathrm{OPT}(L)+\dfrac{6}{9}$, 这是最终的结果, 是由 Dósa 和李荣珩等得出的. 对于 $\mathrm{FFD}(L)$ 的估计是组合优化理论中一个得到深入研究的问题. 在 1973 年, 首先由 D. S. Johnson 经过复杂的计算证明了 $\mathrm{FFD}(L) \leqslant \dfrac{11}{9}\mathrm{OPT}(L)+4$. 据说全文长达 100 多页 (未发表), 其后经 Baker, Yue 等将式中的 4 改成 1. 中间不少人曾在这个问题上作过努力, 现在得到完全解决, 这是难能可贵的. 第 4 章中还载有徐开红同志的一个很好的成果, 她考虑的问题是: 设 L 中的物体 G_i 皆满足 $G_i \leqslant a < 1$, 其中 a 是给定的常数. 第 6 章中关于 Steiner 比猜想的证明, 主要部分是全新的. 作者原稿中存在不少计算的错误, 经李荣珩教授仔细审查, 做了不少

改正. 本书所载皆经笔者看过, 若有谬误与不妥之处, 自应由作者负责. 若发现谬误, 祈来函指正, 使之得以改正, 当真诚感谢.

越民义

2013 年 11 月

此为试读,需要完整PDF请访问: www.ertongbook.com

第一版前言

本书所讨论的是组合优化中的几个著名问题：即排序问题、装箱问题、Steiner 树问题. 对于这些问题的讨论, 一般分为以下三个方面：① 计算复杂性的研究, 也就是论证所讨论的问题是否具有 NP 完备性. ② 寻求有效算法. ③关于近似算法的研究. 限于作者的研究工作范围, 本书只讨论有效算法和近似算法. 本书第 1 章是排序问题的概述, 第 2 章和第 3 章是排序问题的有效算法, 第 5 章是排序问题的启发式算法. 第 3 章原是韩继业教授和作者在 20 世纪 80 年代共同工作时的一些结果, 其中一部分是整理国外一些作者的工作, 一些启发式算法则是韩继业教授的研究成果. 1988 年作者在奥地利 Graz 大学工作时, 曾将这三章作为教材使用了一段时间. 当时我们计划将此部分内容扩写成书, 但由于各自的兴趣很快转移到别的问题, 写书的计划无法完成. 时至今日, 这些材料所涉及的理论和方法尚未在已出版的书中出现, 而且这个领域的研究还在不断深入地发展, 应用范围也在不断地扩大, 因此, 花费一些精力, 将上述排序问题的材料加以整理, 写入本书, 是很有必要的. 第 4 章的 4.6 节和 4.7 节, 第 6 章和第 7 章介绍作者近年来的一些研究成果. 主要是处理几个著名的组合优化问题的近似算法. 也就是说, 用一些简单易行的算法对这些问题求解, 得到的解答一般不是最优的, 必须对与这样得出的解答相应的目标函数的值和真正的最优解相应的值之差作出估计, 以判定这些方法的优劣. 由于绝大多数组合优化问题是 NP 完备的, 因此对近似算法的研究看来是一个很有前途的方向. 本书各章内容基本上是相互独立的, 读者可根据自己的需要阅读有关的章节.

越民义

2001 年 8 月

目　　录

此为试读，需要完整PDF请访问：www.ertongbook.com

第 1 章　概　　述

组合优化 (Combinatorial Optimization) 是运筹学中最重要的分支之一, 在这个领域内存在着大量在生产实际中急需解决而又极其困难的问题. 组合优化 (或称组合规划、离散优化) 是指研究决策变量只取某些离散数值 (通常为有限个) 的优化问题的一门学科. 设施选址 (Facility Location) 问题、装箱 (Bin-Packing) 问题、排序 (Sequencing 或 Scheduling) 问题、网络流 (Network Flow) 问题、旅行售货商 (Traveling Salesman) 问题、集合的覆盖 (Set Covering) 问题、可满足性 (Satisfiability) 问题、背包 (Knapsack) 问题等是组合优化中著名问题的 "代表". 本书主要讨论的是排序问题、装箱问题及 Steiner 树问题.

1.1　组合优化问题的算法

1.1.1　算法

组合优化中的实际问题都是大规模的, 不存在求解的公式, 因而需要设计出求解的程序. 求一个问题的解的程序称为算法. 本书讨论的算法大致可分为三个方面: 多项式时间算法、启发式算法和近似算法.

1. *多项式时间算法* (Polynomial Time Algorithm)

我们说一个组合优化问题是多项式时间问题或 P- 问题, 是指存在多项式时间算法求得这个问题的任何实例的最优解. 本书中关于排序理论中已经解决了的问题, 即已经找到了求解这类问题的有效方法 —— 多项式时间算法, 是我们从 20 世纪 80 年代开始, 对文献中有关这方面的研究成果加以挑选、整理所得到的. 其中包括单台机器和两台机器方面的成果, 以及对同顺序的三台机器的某些特殊情况的处理. 对同顺序的三台机器的某些特殊情况的研究工作主要是由韩继业教授完成的. 同顺序的三台机器的排序问题已在 1976 年被 Garey 等证明是 NP 完备的.

2. *启发式算法* (Heuristic Algorithm)

组合优化问题大都没有找到能求最优解的多项式时间算法, 因而启发式算法是解决排序理论中某些典型问题的常用方法. 这类算法, 一般说来, 其计算量是指数的, 但在实际应用场合却显得相当有效. 本书这部分工作主要也是由韩继业教授完成的.

3. 近似算法 (Approximation Algorithm)

在没有有效方法求得最优解的情况下, 近似算法是目前解决组合优化中大量实际问题的主要方法. 本书这方面的内容可分为两个部分：一部分是讨论排序问题, 所介绍的内容主要是由 R. L. Graham 在 20 世纪 70 年代和 80 年代做出的一些结果; 另一部分所介绍的则是作者近年来在排序问题、装箱问题、Steiner 树方面所得到的一些结果. 本书所讨论的几个问题都是组合优化近似算法中的著名问题. 由于近似算法是组合优化理论的一个重要组成部分, 因此这里的介绍或许会对一些感兴趣的读者提供帮助.

1.1.2 算法的评估

实际上, 组合优化中的绝大部分问题都是 NP 完备问题. 对于这种问题, 从计算量的角度看, 不大可能存在有效的求解方法 (所谓的多项式时间算法), 除非有 NP=P. 但这类问题在实际中常会遇到, 对于这类问题, 虽然一般不可能求到最优解, 但有希望找到某种方法求出它的近似解. 近似算法理论所研究的就是要对于一种给定的算法 (一般是简单易行的), 估计出由它所产生的解与问题的真正的最优解所对应的目标函数的值之间的差异程度. 因而在某种给定的标准之下, 它可以评估几种给定的算法之间的优劣. 通常评估算法优劣的标准有两个：一个标准是计算时间, 另一个标准是性能比.

本书介绍的数学理论和方法, 无论是从文献中收集来的, 还是作者自己提出的, 我们都尽量给出严格的论证.

1.2 排序问题的记号和模型描述

排序理论 (时间表理论) 是组合优化的一个重要组成部分, 近年来, 这方面无论是在决定性的情况还是在随机性的情况, 都得到了迅速的发展. 本书所讨论的只限于决定性的情况, 虽然这一领域的工作只是在 20 世纪 50 年代初期才开始, 但有关的文献在数量上却多得惊人. 由于它的应用范围逐渐扩大, 新的问题不断出现, 因而从事这一领域研究的人与日俱增, 其内容也越来越丰富, 应用也越来越广泛. 本书将对排序理论的几个典型的问题及算法作些介绍. 生产实际中存在大量的、不同类型的排序问题, 为了叙述简便, 人们使用了一些简便的描述排序问题的记号和方法. 我们在这里先做一个简略的说明, 下面所使用的记法来自 Lawler 等 (1993) 和唐国春等 (2003).

1.2.1 排序问题的记号

设有 n 个给定的工件 (Job)$J_1, \cdots, J_n$ 要在 m 台机器 $M_1, \cdots, M_m$ 上加工, 每

一机器 M_i 在任一时刻只能处理一个工件. 这种加工要受到某些给定条件的限制, 例如, 工件的发放时间 (即开始送到指定的机器上加工的时间), 必须交货的时间 (即工件必须加工完毕的最后期限), 工件之间由于工艺上的需要必须满足的先后顺序, 又如一些工件在某一台机器上加工, 要求这项加工必须一次完成 (即不能中断), 而在某些情况下则允许在加工过程中可以中断, 转为加工别的工件, 之后再恢复对原来工件的加工; 某一工件随其加工时间的长短需要付出一定的费用, 这项费用一般假设是时间的非降函数, 等等. 现在的问题是要求一种关于工件的排序, 使得某种给定的指标达到最优. 本书中所考虑的工件数据如下.

对每一给定的工件 J_j, 以下一些数据需要详细给出说明:

(1) 加工时间 p_{ij}, 即 J_j 在 M_i 上加工所需的时间, 若 $m=1$, 则令 $p_{ij}=p_j$.

(2) 工件准备就绪时间 (Release Time)r_j.

(3) 应交货时间 d_j, 即根据某种要求, 规定 J_j 必须加工完毕的时间, 我们总假定 $d_j \geqslant \sum\limits_{i=1}^{m} p_{ij}$.

(4) 权 w_j, 表示 J_j 的相对重要性.

(5) 价格函数 f_j, 是用来测量当工件 J_j 在时刻 t 完成时所需投入的费用 $f_j(t)$.

这里, 一般假定 p_{ij}, r_j, d_j, w_j 为整数.

1.2.2 排序问题的模型描述

1. 机器环境 (Machine Environment)

机器环境常见的 7 种类型是: 单台机 1 种、平行机 3 种和串联机 3 种. 平行机 (Parallel Machine) 是指所有机器的功能是一样的, 每一个工件只有一道工序且可以在任何一台机器上进行. 串联机是指一个工件有多道工序, 不同的工序需在不同的机器上进行.

(1) 我们用 m 表示机器台数, 单台机是指 $m=1$.

(2) P 表示所给 m 台机器 $M_1, \cdots, M_m$ 是同一型号且是平行的同型机 (Identical Parallel Machine), 即 $p_{ij}=p_j, j=1,2,\cdots,n$, 每一个工件 J_j 的加工只有一道工序 O_j, 它可以在任一机器 M_i 上完成.

(3) Q 表示 m 台机器 $M_1, \cdots, M_m$ 是速度一致且平行的同类机 (Uniform Parallel Machine), 它与 P 不同的是每一台机器 M_i 有自己的速度 q_i, 且对工件 J_j 满足 $p_{ij}=\dfrac{p_j}{q_i}$.

(4) R 表示 m 台机器 $M_1, \cdots, M_m$ 是速度不一致但平行的非同类机 (Unrelated Parallel Machine), 它与 Q 不同的是每一台 M_i 的速度 q_{ij} 是与工件 J_j 相关联的, 满足 $p_{ij}=\dfrac{p_j}{q_{ij}}$.

(5) O 表示所给的问题为自由作业型串联机 (Open Shop), 这里每一个工件 J_j 的加工有 m 道工序 $O_{1j}, \cdots, O_{mj}$, 其中 O_{ij} 表示 J_j 必须在 M_i 加工 p_{ij} 个单位时间, 但这 m 道工序的次序是可以任意选定的.

(6) F 表示所给的问题为流水作业型串联机 (Flow Shop), 这里每一个工件 J_j 的加工是一串工作 $O_{1j}, \cdots, O_{mj}$, 其中 O_{ij} 表示 J_j 的第 i 道工序必须在 M_i 加工 p_{ij} 个单位时间, 即所有工件在 m 台机器上的加工次序是固定且一致的.

(7) J 表示所给的问题为异序作业型串联机 (Job Shop), 这里每一个工件 J_j 的加工是一串工作 $O_{j(1),j}, \cdots, O_{j(m_j),j}$, 其中 m_j 是工件 J 所要加工的工序的总道数, $O_{j(i),j}(i = 1, 2, \cdots, m_j)$ 表示 J_j 的第 i 道工序必须在机器 $M_{j(i)}$ 上加工 $p_{j(i),j}$ 个单位时间, 且 $j(i) \neq j(i+1), i = 1, 2, \cdots, m_j - 1$, 即相邻两道工序不能在同一台机器上完成, 但不相邻的两道工序可以在同一台机器上完成.

2. **工件的特性** (Job Characteristics)

(1) pmtn(Preemption) 表示允许中断, 即任何一道工序的加工过程容许中断, 在以后的时间再继续加工.

(2) online(Semi-Online) 表示在线 (半在线) 排序.

(3) prec(Precedence) 表示工件之间有特定的先后加工顺序.

(4) r_j(Release Time) 表示工件到达或准备就绪时间, 即 J_j 可以开始在机器上进行加工的时间.

3. **最优准则或目标函数**

对排序问题, 通常我们做如下约定:

约定 1.2.1 (i) 机器一旦开始工作, 它就不能停下来, 除非已经无工件可以用来加工.

(ii) 任何一个工件 J_j 一旦在机器 M_i 上加工, 则在 p_{ij} 这段时间内, 这项加工一直继续, 除了在允许中断的情况, 且这种情况必须加以说明.

对于一组工件给定的一个排序, 我们有下面的记号:

C_j: 表示工件 J_j 的完工时间;

$L_j = C_j - d_j$: 表示工件 J_j 的延迟时间, 延迟时间可以为正也可以为负, 为负表示提前完工;

$T_j = \max\{0, C_j - d_j\}$: 表示工件的误工时间, 误工时间不能为负.

函数 $\sum\limits_{j=1}^{n} C_j, \sum\limits_{j=1}^{n} w_j C_j, C_{\max} = \max\{C_j | 1 \leqslant j \leqslant n\}, \sum\limits_{j=1}^{n} T_j, \sum\limits_{j=1}^{n} w_j T_j$ 是最常见的几个目标函数. 本书中其他最优准则或目标函数将根据具体问题给出相应的说明.

在没有特别说明的情况下本书将采用三参数 $\alpha|\beta|\gamma$ 来表示所给的问题类型, 这

里的 α, β 和 γ 是三个参数组, 分别表示机器环境、工件特征和目标函数. α 用数字表示机器台数, P, Q 和 R 分别表示三类平行机 (Parallel Machine): 同型机、同类机和非同类机. F, O 和 J 分别表示流水作业、自由作业和异序作业这三类串联机 (Shop Machine). 参数 β 表示工件特性, 这里有 r_j 表示工件的就绪时间可以不相同, 否则表示所有 r_j 为零; pmtn 表示加工允许中断, prec 表示工件之间有加工优先序. γ 表示最优准则或目标函数, 比如 $C_{\max}$ 表示最大完工时间 (Makespan), $\sum\limits_j C_j$ 表示总完工时间, $\sum\limits_j w_j C_j$ 表示带权总完工时间. 例如: $F2|\text{pmtn}|C_{\max}$ 或 $F2|\text{pmtn}|\text{makespan}$ 表示所讨论的模型是两台机器的流水作业, 允许中断, 工件之间无任何先后加工顺序, 工件无准备时间, 目的是要寻求一种排序, 使得整个加工过程所耗的时间最短. $1|\text{prec}, r_j|\sum\limits_j w_j T_j$ 表示所讨论的模型是单台机, 工件有优先序和准备时间, 目标是使带权总误工时间最小.

本书中, 对于排序 (Sequencing) 和时间表 (Scheduling) 这两个术语, 我们将不加区分地使用.

第 2 章　一台机器上的排序

本章将讲述关于在一台机器上排序的结果, 所讲的模型都是已有方法解决的, 有的比较简单, 有的则相当困难, 我们假定机器是从 $t=0$ 这一时刻开始工作的.

2.1　$1||\sum_{j=1}^{n} a_jC_j$

设有 n 个工件 $J_j(j=1,2,\cdots,n)$ 需要在一台机器 M 上加工, J_j 在 M 上的加工时间为 p_j, 设对于所有的 j, J_j 的发放时间 (即工前准备时间)$r_j=0$, 又设 a_j 为 J_j 在其完工之前每单位时间的损失, C_j 为 J_j 的完工时间, 显然, $C_1,\cdots,C_n$ 的值取决于所给的排序, 我们的目的是要寻求一种排序, 使总的损失 $\sum\limits_{j=1}^{m} a_jC_j$ 为最小.

2.1.1　算法

将 $J_1,\cdots,J_n$ 排成 $J_{i_1},\cdots,J_{i_n}$, 使得

$$\frac{p_{i_1}}{a_{i_1}} \leqslant \frac{p_{i_2}}{a_{i_2}} \leqslant \cdots \leqslant \frac{p_{i_n}}{a_{i_n}},$$

则排序 $J_{i_1} \to J_{i_2} \to \cdots \to J_{i_n}$ 即为最优.

例 2.1.1　设有关的数据由表 2.1.1 给出.

表 2.1.1　数据

J_j	J_1	J_2	J_3	J_4
p_j	4	6	7	9
a_j	0.5	1	0.8	1.2
p_j/a_j	8	6	8.75	7.5

由表 2.1.1 可知最优排序为 (J_2,J_4,J_1,J_3).

2.1.2　最优性证明

引理 2.1.1　令 $I=(i_1,i_2,\cdots,i_n)$ 为任一给定的 $(1,2,\cdots,n)$ 的排序, 若存在 r, 使得

$$\frac{p_{i_r}}{a_{i_r}} > \frac{p_{i_{r+1}}}{a_{i_{r+1}}}, \tag{2.1.1}$$

此为试读,需要完整PDF请访问: www.ertongbook.com