期末复习

岳镝

2025 年 6 月 6 日

遗留问题 1: 装箱

给定 n 件物品的重量 $\{w_j\}_{j=1}^n$，限制每个箱子的容量为 B，要求用最少的箱子装下所有物品

近似算法：依次处理每一件物品 [Delorme-Iori-Martello, 2016 survey]

- First-Fit (FF): 装进第一个能装得下的箱子，若不存在则新开一个，$\alpha = 1.7$
- Next-Fit (NF): 装进当前箱子，装不下则新开一个，$\alpha = 2$
- Best-Fit (BF): 装进剩余空间最小的可行箱子，若不存在则新开一个，$\alpha = 1.7$
- First-Fit Decreasing (FFD): 先按递减排序，再 FF，$\alpha = 1.5$
- Best-Fit Decreasing (BFD): 先按递减排序，再 BF，$\alpha = 1.5$

下界：装箱问题不存在 $\alpha < 1.5$ 的多项式近似算法，除非 $\mathrm{P} = \mathrm{NP}$ (习题 10.4)

遗留问题 2：顶点覆盖

考虑最小顶点覆盖问题的以下近似算法：

- 任取一条未覆盖的边，把这条边度数较大的一个端点加入顶点覆盖集

近似比？

算法设计与分析

学什么？

- 算法设计
  (1) 各种经典算法：分治、动态规划、贪心、回溯、线性规划、网络流...
  (4) 现代算法设计技术：近似算法、随机算法...
- 算法分析
  (2) 时间复杂度、空间复杂度、均摊分析...
- 一点基础的复杂性理论
  (3) 问题的计算复杂度、难解性...

数学基础

- 函数的阶
- 求和的方法
- 递推方程求解方法:
  - 迭代法
  - 递归树
  - 尝试法
  - 主定理

分治

- 重要技术
  - 芯片测试：两两分组，每轮淘汰至少一半
  - 多项式在单位根上求值：按奇偶次项划分（这在数学上对应着快速傅里叶变换 (FFT) 的深刻背景，和考试可能关系不大）
- 重要结果
  - 选第 k 小：$O(n)$ (联系第 8 章复习)
  - 方幂 a^n：$O(\log n)$
  - 最近点对距离：$O(n \log n)$

动态规划

基本格式

- 列递推方程
- 确定初始/边界条件
- 标记函数：考试如果要求求出最优解，必须写出标记函数
- 时间复杂度：通常可以先确定空间复杂度（状态的总数），然后看每一步递推的时间复杂度

贪心

- 几个经典的贪心算法正确性证明、时间复杂度
  - Huffman
  - Prim, Kruskal
  - Dijkstra
- 贪心算法的设计与分析
  - 对步骤/规模归纳
  - 交换论证
  - 模仿经典算法的设计/分析思路

回溯与分支限界

线性规划

- 问题的线性规划模型
- 线性规划标准形（怎样转化）
- 单纯形法
- 对偶线性规划（怎样转化）
- 整数线性规划

结合第 9, 10, 11 章复习

平摊分析

平摊分析的三种方法

- 聚集分析
- 记账法
- 势能法
  - 常见的势函数定义
  - 势函数 ≥ 0
  - 每步开销 $c_i \lesssim \Phi_{i-1} - \Phi_i + O(1)$

网络流

网络流的建模

- 拆点 (习题 7.7)
- 带需求的流通 (习题 7.21)
- 边上容量存在下界
- 最小费用流

网络流理论和算法

- 最大流-最小割定理
- 最大流的算法及其时间复杂度
  - 任意增广路径 (Ford-Fulkerson): $O(mf^*)$
  - 最短增广路径 (Edmonds-Karp): $O(nm^2)$
  - 分层辅助网络 + 前向增广路径 (Dinic): $O(n^3)$
  - ...
- 最小费用流算法
  - 负回路算法：保持流量，逐步减少费用
  - 最短路径算法：保持费用最小，逐步增大流量

问题的计算复杂度

分析技术

- 平凡下界
- 直接计数
- 决策树
  - 节点、边、路径、树的高度各自含义？
  - 估计树高：$\log$ 节点总数或叶节点数
- 根据算法执行构造最坏情况输入
- 归约

问题的计算复杂度

重要结果

- 检索：$\lfloor \log n \rfloor + 1$
- 排序：$\Omega(n \log n)$。哪些排序算法达到最坏情况最优？平均情况最优？
- 选最大：$n - 1$
- 选最大最小：$\lceil 3n/2 \rceil - 2$
- 选第二大：$n + \lceil \log n \rceil - 2$
- 选中位数：$3n/2 - 3/2$
- 选第 k 小：$n + \min\{k, n - k + 1\} - 2$
- 元素唯一性：$\Omega(n \log n)$
- 平面最近点对：$\Omega(n \log n)$

NP 完全性

NP 类定义（考虑判定问题）

- 多项式时间可验证
- 存在多项式时间非确定型判定算法

多项式时间变换与 NP 完全

- 多项式时间变换（Karp 归约）：定义、性质
- NP 完全：定义、性质
- Cook 归约（Turing 归约）

证明问题 Π 是 NP 完全

(1) $\Pi \in \mathrm{NP}$

(2) 某个已知的 NP 完全问题 $\Pi' \leqslant_p \Pi$

- 注意归约顺序
- 注意要使用 Karp 归约。实在不会证再用 Cook 归约也行

几个经典的 NP 完全问题：看书

近似算法

主要技术

- 设计近似算法，分析近似比
- 针对给定的算法构造紧实例
- 近似的困难性：$\mathrm{P} \neq \mathrm{NP}$ 假设下，利用归约可以证明
  - 一般的货郎问题不可近似
  - 装箱问题不存在 $\alpha < 3/2$-近似算法

书上的几个例子

- 顶点覆盖：2-近似
- 多机调度：$3/2$-近似
- Metric TSP: $3/2$-近似
- 背包问题：存在 PTAS
  - PTAS: 多项式时间近似方案
  - FPTAS: 完全多项式时间近似方案（定义看书）

随机算法

概率论基础：复习概率论

- Union bound
- Markov's Inequality
- Chebyshev's Inequality
- Chernoff/Hoeffding Bounds

随机算法的分类

- 拉斯维加斯型随机算法
  - 例子：随机快速排序、随机选择
- 蒙特卡洛型随机算法
  - 单侧错误与双侧错误
  - 例子：主元素测试、串相等测试、素数测试

减小错误概率的方法 (Amplification)

- 单侧错误：多次独立重复
- 双侧错误：多次独立重复 + Majority vote

例题 1 — 网络流

给定有向无圈图 (DAG) $G=(V,E)$。称 G 中若干条有向路径组成的集合 $\mathcal{P}=\{P_1,P_2,\ldots,P_m\}$ 是 G 的一个（顶点不交）路径覆盖，若满足任何顶点 $v\in V$ 恰好属于 $\mathcal{P}$ 中一条路径。

设计一个多项式时间算法，求 G 的最小路径覆盖。

- 提示：构造一个图 G'，使得 G 的最小路径覆盖恰为 $n-\text{Max-Flow}(G')$

例题 2 — NP 完全性

证明下列问题是 NP 完全的

(1) Π_1: 给定非负序列，能否划分成和相等的两个子序列？

(2) Π_2: 给定长度为偶数的非负序列，能否划分成和相等，长度也相等的两个子序列？

(3) Π_3: 给定长度为偶数的非负递增子序列，能否划分成和相等的两个子序列，且满足 a_{2i-1} 和 a_{2i} 属于不同的子序列 $(\forall\ 1 \leq i \leq n/2)$？

例题 3 — NP 完全性

证明支配集问题是 NP 完全的：给定无向图 $G = (V, E)$ 和正整数 $K \leq |V|$，是否存在子集 $V' \subseteq V$ 使得 $|V'| \leq K$ 且 $V \setminus V'$ 中的每个顶点都至少与 V 中一个顶点相邻？

- 提示：顶点覆盖 $\leqslant_p$ 支配集
- 怎样用顶点“替换”边？

例题 4 — 随机算法

对 d-正则图 $G=(V,E)$，考虑支配集构造算法

1. $V' \leftarrow \emptyset$
2. for $i=1,2,\ldots,k$ do
3. 随机采样 $v \in V$
4. $V' \leftarrow V' \cup \{v\}$

(1) 证明当 $k=\frac{2n\ln n}{d+1}$ 时，V' 是支配集的概率至少 $1-1/n$

(2) 构造一个拉斯维加斯型随机算法，返回 G 的一个支配集

例题 5 — 近似算法

考虑支配集问题的优化形式——最小支配集

(1) 写出最小支配集问题的整数线性规划模型

- 提示：用 x_v 表示顶点 v 是否属于支配集

例题 5 — 近似算法

考虑支配集问题的优化形式——最小支配集

(1) 写出最小支配集问题的整数线性规划模型

- 提示：用 x_v 表示顶点 v 是否属于支配集

$$
\begin{aligned}
& \text{minimize} \sum_{v \in V} x_v, \\
\text{s.t.} \quad & x_v + \sum_{u:\,(u,v) \in E} x_u \geq 1, \quad \forall v \in V \\
& x_v \in \{0,1\}
\end{aligned}
$$

例题 5 — 近似算法

考虑支配集问题的优化形式——最小支配集

(2) 将约束放松为 $x_v \in [0,1]$，得到线性规划问题

$$
\begin{aligned}
& \text{minimize} \sum_{v \in V} x_v, \\
\text{s.t.} \quad & x_v + \sum_{u\colon (u,v) \in E} x_u \geq 1, \quad \forall v \in V \\
& x_v \in [0,1]
\end{aligned}
$$

在多项式时间可求出最优解 $\{x_v^*\}_{v \in V}$。
假设图 G 的最大度是 Δ，令 $V' = \{v\colon x_v^* \geq \frac{1}{\Delta+1}\}$。证明 V' 是一个 $(\Delta+1)$-近似解

例题 5 — 近似算法

考虑支配集问题的优化形式——最小支配集

(3) 考虑任意图 G。对每个顶点 $v \in V$，令随机变量 $M_v^1, M_v^2, \ldots, M_v^\alpha \sim_{i.i.d.} B(1, x_v^*)$。令

$$V' = \{v \colon \exists\ 1 \leq i \leq \alpha,\ \text{s.t. } M_v^i = 1\}$$

证明：存在整数 $\alpha = O(\log n)$，使得 V' 以概率 $1 - 1/n$ 是 G 的一个支配集。进而证明 V' 以常数概率构成一个 $O(\log n)$-近似解

例题 6 — 近似算法

(1) 考虑求最小支配集的以下近似算法

- 每次找一对未被支配的相邻顶点，将它们加入 V'，直到所有顶点都被支配

构造实例说明该算法的近似比是 $\Omega(n)$

(2) 设计一个近似比为 $o(n)$ 的多项式时间确定性算法

- 提示：最小支配集可归约到最小集合覆盖。参见王千烨同学近似算法 2 回课 PPT