

Lecture 0: 课程简介

编写人: 吴一航 yhwu_is@zju.edu.cn

0.1 声明

欢迎大家来到《高级数据结构与算法分析》(Advanced Data Structure & Algorithm Analysis, 简称 ADS) 课程! 这是计算机学院的同学不可或缺的专业基础课程, 将会介绍一些重要但相对于上个学期数据结构基础课程更加复杂的数据结构, 以及基本的算法设计思想与应用。这门课的部分章节可以视为理论计算机科学的入门, 之前的学长学姐们也或多或少都与你们吐槽过这门课的难度, 但相信在各位的努力之下, 这些都不是问题。

本讲义初稿编写于 2023-2024 学年春夏学期, 目前正在积极修订中。编写过程中主要参考浙江大学《高级数据结构与算法分析》课程组 PPT, 以及后面列出的参考书。注意, 本讲义可能存在错误, 请仔细甄别, 若发现错误可以通过邮件等方式联系编写人。

0.2 为什么学这门课

一个经典的等式可以概括:

$$\text{Programming} = \text{Data Structures} + \text{Algorithms}$$

算法是解决问题的核心, 而数据结构则可以通过有效管理数据, 减少算法中对输入数据或中间数据的搜索、插入、删除、合并等操作的时间。

0.3 这门课学什么

这门课主要包含以下主题:

1. 高级数据结构

- 平衡搜索树: AVL 树, Splay 树, B+ 树, 红黑树
- 堆: 左式堆, 斜堆, 二项堆 (project 中还涉及斐波那契堆)
- 应用: 倒排索引 (搜索引擎算法)

- 时间复杂度分析方法：摊还分析

2. 算法分析

- 基本的算法设计思想：回溯法，分治法（含主定理），动态规划，贪心算法
- 难解问题的算法设计理论与方法：NP 问题，近似算法，局部搜索，随机算法
- 其他主题：并行算法，外部排序

0.4 有什么参考书

1. 《数据结构与算法分析：C 语言描述》(*Data Structures and Algorithm Analyses in C*), [美] Mark Allen Weiss

同数据结构基础教材。数据结构（AVL 树、Splay 树、左式堆、斜堆和二项堆等）以及部分的算法分析内容（回溯、分治和动态规划等）基本按照本教材思路进行。教材保持老外写书略啰嗦但比较清晰的特点，比 PPT 详细，如果只看 PPT 无法理解可以配合教材阅读。

2. 《算法导论》(*Introduction to Algorithms*) [美] Thomas H. Cormen, Charles E. Leiserson. Ronald L. Rivest, Clifford Stein

令人闻风丧胆的鸿篇巨著，MIT 的算法课教材，有的地方也将其根据作者姓氏首字母简称 CLRS。在这门课中，摊还分析、红黑树、B+ 树、贪心算法、NP 和随机算法等章节基本参考本教材，除此之外，作业中也有难题源于此书，project 中关于斐波那契堆的内容也可以参考。相信很多同学对这本书早有耳闻且有敬畏之心，个人不是非常推荐初学者阅读这本书，因为其中存在不少跳步或者写作不够清晰之处，并且内容太多可能让人无法在一学期内把握核心想法，更适合当个字典挑选部分与课程内容相关的章节、例子等阅读，至少比课程组的 PPT 详细许多。

3. 算法设计 (*Algorithm Design*) [美] Jon Kleinberg, Éva Tardos

同样是算法类图书的经典之作，Jon Kleinberg 和 Éva Tardos 是康奈尔大学的教授，其中 Jon Kleinberg 不到四十岁就拿到了美国三院（美国科学院、工程院、艺术与科学院）院士，Éva Tardos 也是算法博弈论领域的开拓者之一，因此都是顶级大牛。在这门课中，近似算法、局部搜索等章节基本参考本教材，并且也有部分习题出自本教材。我个人非常推荐阅读这本教材，虽然偶尔有点啰嗦，但写作非常清楚且令人感到舒适，其中的例子和习题也都非常经典，也会介绍一些算法设计的思想。

除上述教材外，我个人也推荐基础比较薄弱的同学参考哥伦比亚大学教授（此前是斯坦福大学教授）Tim Roughgarden 的《算法详解》，这本书总共四卷（每卷都特别薄，所以可以看得很快），覆盖了 FDS 和 ADS 的较多内容。Tim 是前面提到的 Éva Tardos 的学生，同样也是算法博弈论领域的开拓者之一。这套书的特点就是相当简单易懂（绝对的入门书籍），作者已经把算法的设计思想、流程等给你嚼烂了喂嘴里了，所以读起来特别亲切（特别是动态规划部分，读者阅读后一定会有所体会）。特别是学 ADS 之前也可以用这本书再回顾一下 FDS 的内容，也会有所收获。

0.5 杂谈

可以说这门课是理论计算机方向的入门课程，但显然完全不是一门合格的入门课程。第一课程设计并不合理，在短短一学期内讲清楚学清楚这么多的内容简直是天方夜谭，特别是算法分析部分，很多章节都可以单独扩展成一个学期的课程；第二，课程组授课老师水平参差不齐，很多老师自己也没有专门学过这些理论，并不是做这些方向的，因此讲课也基本只能按照 PPT 来，但很遗憾，课程组的 PPT 非常简洁，只是列举一些例子和解法，没有逻辑关系的说明，也没有什么算法设计的思想，让人很难学明白，甚至不知道自己在学什么，只是看到几个例子，因此学生大部分也只能靠自学，或者不怎么学，总而言之学不明白；第三，这门课教考略分离，这是大课程组难以避免的现象（当然 ADS 至少比数据库、计算机网络等课程要良心一点），因为期末考试是每个老师负责一个小部分，很多老师也不知道自己上课希望给同学们传达什么重要的东西，所以考试无法与授课非常贴合，并且老师出题风格很随机，有的可能就喜欢出一些钻牛角尖的概念，有的喜欢挑难题。

总而言之，我相信大部分同学都认为这门课是梦魇，是计院最难学明白的专业课（之一），但我希望同学们认清一个事实，**这不完全是你们的问题，整个课程安排和课程组都存在非常大的问题**。如果同学们愿意跳出课程设计的种种缺陷，多花时间看一些前面推荐的教材还有一些学长写的讲义 / 笔记，整理一下这些算法背后的思想等，我认为会收获很多，并且会认识到这门课本质并不是非常难，而不是像现在一样学完之后压根不知道自己学了啥，只知道这些东西好像很难，以后再也不愿意碰理论计算机方向了。至于考试，除了部分老师抽风出的题之外，大部分题目应该也是不难的（特别是近年难度还下调了的情况）。

关于这门课的其他信息（如评分标准、其他参考资料和学习建议等），同学们可以参考我和其他学长学姐们共同建设的图灵课程指南的[高级数据结构与算法分析部分](#)。祝大家都能在这门课中学有所成，收获满意的成绩！