

玩儿转算法面试

liuyubobobo

慕课网《玩转算法面试》

讲师：liuyubobobo

版权所有，侵权必究

这个课程的目标是什么？

让大家在面对面试中的算法问题时，有一个合理的思考路径；

面对算法面试，~~讲稿~~不畏惧。

- 因为面试中的算法问题，通常并不“复杂”，远远不需要啃完一本《算法导论》

什么是算法面试？

讲师：luyanbo
慕课网《玩转算法面试》

版权所有，侵权必究

算法面试是什么？

让大家在面对面试中的算法问题时，有一个合理的思考路径；

• 不代表能够“正确”回答每一个算法问题，但是合理的思考方向其实更重要，

也是正确完成算法面试问题的前提

• 算法面试优秀不意味着技术面试优秀

• 技术面试优秀不意味着能够拿到Offer

什么是给出合理的思考路径

讲师：liuyuboobobo

慕课网《玩转算法面试》

版权所有，侵权必究

什么是给出合理的思考路径

算法面试的目的不是给出一个“正确”答案，
而是展示给面试官你思考问题的方式。

“正确”本身是一个相对概念

算法面试不是高考。

把这个过程看作是和面试官一起探讨一个问题的解决方案。

对于问题的细节和应用环境，可以和面试官沟通。

这种沟通本身很重要，它暗示着你思考问题的方式。

我们需要对一组数据进行排序

讲师：liuyubobobo

慕课网《玩转算法面试》

版权所有，侵权必究

对一组数据进行排序

快速排序算法 $O(n \log n)$



对一组数据进行排序

这组数据有什么样的特征？

- 有没有可能包含有大量重复的元素？
- 如果有这种可能的话，三路快排是更好地选择。

对一组数据进行排序

这组数据有什么样的特征？

- 是否大部分数据距离它正确的位置很近？是否近乎有序？
- 如果是这样的话，插入排序是更好地选择。

对一组数据进行排序

这组数据有什么样的特征？

- 是否数据的取值范围非常有限？比如对学生生成绩排序。
- 如果是这样的话，计数排序是更好地选择。

对一组数据进行排序

对排序有什么额外的要求？

- 是否需要稳定排序？
- 如果是的话，归并排序是更好地选择。

对一组数据进行排序

数据的存储状况是怎样的？

- 是否是使用链表存储的？
- 如果是的话，归并排序是更好地选择。

对一组数据进行排序

数据的存储状况是怎样的？

- 数据的大小是否可以装载在内存里？
- 数据量很大，或者内存很小，不足以装载在内存里，需要使用外排序算法。

对一组数据进行排序

- 有没有可能包含有大量重复的元素？
- 是否大部分数据距离它正确的位置很近？是否近乎有序？
- 是否数据的取值范围非常有限？比如对学生成绩排序。
- 是否需要稳定排序？
- 是否是使用链表存储的？
- 数据的大小是否可以装载在内存里？

什么是“正确”的回答——一个算法问题

正确还包含对问题的独到见解；优化；代码规范；容错性；

等等等等.....

什么是“正确”的回答——面试一个算法问题

这个课程不仅仅是给出解决算法问题的代码

什么是“正确”的回答——一个算法问题

如果是非常难的问题，对你的竞争对手来说，也是难的。

关键在于你所表达出的解决问题的思路。

甚至通过表达解题思路的方向，得出结论：这个问题的解决方案，
应该在哪一个领域，我可以通过查阅或者进一步学习解决问题。

慕课网《玩转算法面试》
讲师：liuyubobobo
版权所有，侵权必究

算法面试优秀不意味着技术面试优秀

讲师：liuyuxiaoBobo

慕课网《玩转算法面试》

算法面试优秀不意味着技术面试优秀

算法面试只是技术面试的一部分。

根据你的简历和应聘职位的不同，势必要考察其他技术方面。

算法面试优秀不意味着技术面试优秀

项目经历 和 项目中遇到的实际问题

你遇到的印象最深的bug是什么？

面向对象

设计模式

网络相关； 安全相关； 内存相关； 并发相关； ...

系统设计； scalability

讲师：liuyubobobo

技术面试优秀不意味着能够拿到Offer

讲师：
liuyanbobobo

慕课网《玩转算法面试》

技术面试优秀不意味着能够拿到Offer

技术面试只是面试的一部分。面试不仅仅是考察你的技术水平，还是了解你的过去以及形成的思考行为方式。

关于过去：参与项目至关重要

讲师：liuyabobo

慕课网《玩转算法面试》

版权所有，侵权必究

技术面试优秀不意味着能够拿到Offer

项目经历

- 工作人士
- 研究生
- 本科生
 - 毕业设计
 - 其他课程设计

讲师：liuyubobobo
版权所有，侵权必究

技术面试优秀不意味着能够拿到Offer

如何找到项目？

- 实习
- 参与实战课程学习
- 慕课网
- Coursera
- ...

技术面试优秀不意味着能够拿到Offer

创建自己的项目

- 自己做小应用：计划表；备忘录；播放器...
- 自己解决小问题：爬虫；数据分析；词频统计...
- “不是项目”的项目：一本优秀的技术书籍的代码整理等...
- 分享：自己的技术博客；github等等

技术面试优秀不意味着能够拿到Offer

通过过去了解你的思考行为方式

- 遇到的最大的挑战?
- 犯过的错误?
- 遭遇的失败?
- 最享受的工作内容?
- 遇到冲突的处理方式?
- 做的最与众不同的事儿?

技术面试优秀不意味着能够拿到Offer

准备好合适的问题问面试官

- 整个小组的大概运行模式是怎样的?
- 整个项目的后续规划是如何的?
- 这个产品中的某个问题是如何解决的?
- 为什么会选择某些技术? 标准?
- 我对某个技术很感兴趣, 在你的小组中我会有怎样的机会深入这种技术?

算法面试仍然是非常重要的部分

讲师：[liuyanbobobo](#)

慕课网《玩转算法面试》

让我们一起玩儿转算法面试：）

讲师：liuyao**bobo**

版权所有，侵权必究

慕课网《玩转算法面试》

慕课网《玩转算法面试》
讲师：liuyubobobo
版权所有，侵权必究

如何准备算法面试

慕课网《玩转算法面试》

讲师：liuyuxiangobobo

版权所有，侵权必究

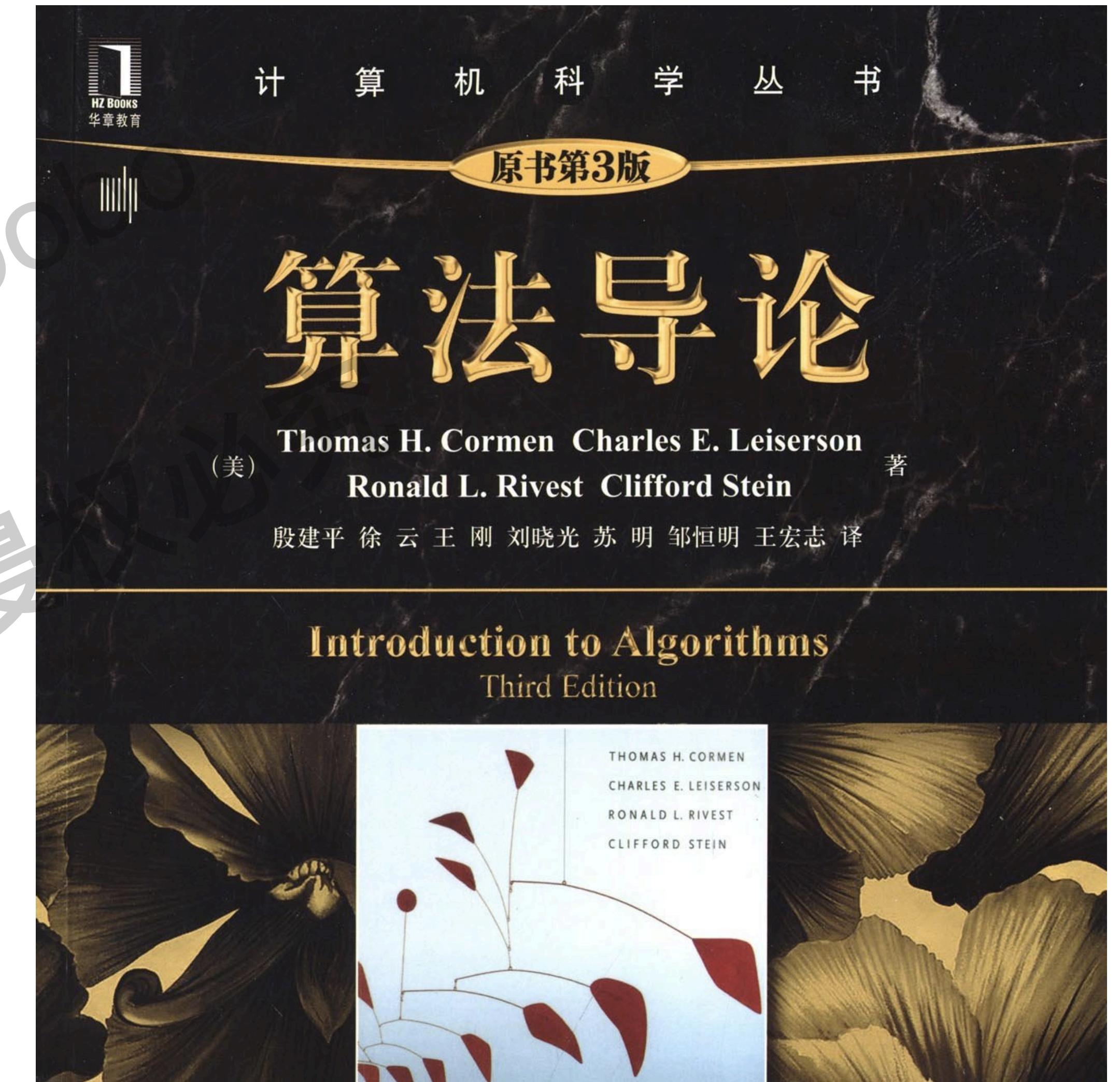
准备面试 和 准备算法面试 是两个概念

算法面试，只是面试中的一个环节。

算法面试并没有那么难

远远不需要啃完一本《算法导论》

- 强调理论证明



算法面试并没有那么难

Our analysis uses indicator random variables (see Section 5.2). We define

$$X_{ij} = \mathbb{I}\{z_i \text{ is compared to } z_j\},$$

where we are considering whether the comparison takes place at any time during the execution of the algorithm, not just during one iteration or one call of PARTITION. Since each pair is compared at most once, we can easily characterize the total number of comparisons performed by the algorithm:

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}.$$

Taking expectations of both sides, and then using linearity of expectation and Lemma 5.1, we obtain

$$\begin{aligned} E[X] &= E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}\right] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n E[X_{ij}] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \Pr\{z_i \text{ is compared to } z_j\}. \end{aligned} \tag{7.2}$$

$$\begin{aligned} \Pr\{z_i \text{ is compared to } z_j\} &= \Pr\{z_i \text{ or } z_j \text{ is first pivot chosen from } Z_{ij}\} \\ &= \Pr\{z_i \text{ is first pivot chosen from } Z_{ij}\} \\ &\quad + \Pr\{z_j \text{ is first pivot chosen from } Z_{ij}\} \\ &= \frac{1}{j-i+1} + \frac{1}{j-i+1} \\ &= \frac{2}{j-i+1}. \end{aligned} \tag{7.3}$$

The second line follows because the two events are mutually exclusive. Combining equations (7.2) and (7.3), we get that

$$E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{2}{j-i+1}.$$

We can evaluate this sum using a change of variables ($k = j - i$) and the bound on the harmonic series in equation (A.7):

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{2}{j-i+1} \\ &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\ &< \sum_{i=1}^{n-1} \sum_{k=1}^n \frac{2}{k} \\ &= \sum_{i=1}^{n-1} O(\lg n) \\ &= O(n \lg n). \end{aligned} \tag{7.4}$$

Thus we conclude that, using RANDOMIZED-PARTITION, the expected running time of quicksort is $O(n \lg n)$ when element values are distinct.

算法面试并没有那么难

远远不需要啃完一本《算法导论》

- 强调理论证明

高级数据结构和算法面试提及的概率很低

- 红黑树
- B-Tree
- 斐波那契堆
- 计算几何
- 数论
- FFT

算法面试并没有那么难

远远不需要到达信息学竞赛的水平



算法面试的准备范围

不要轻视基础算法和数据结构，而只关注“有意思”的题目

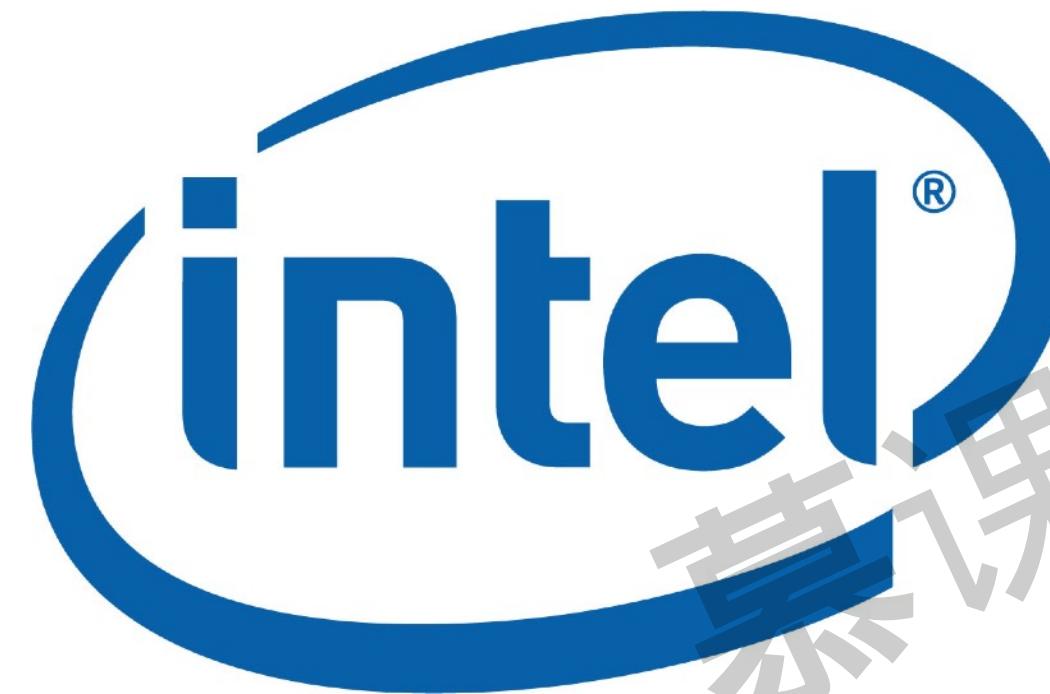
- 各种排序算法
- 基础数据结构和算法的实现：如堆、二叉树、图...
- 基础数据结构的使用：如链表、栈、队列、哈希表、图、Trie、并查集...
- 基础算法：深度优先、广度优先、二分查找、递归...
- 基本算法思想：递归、分治、回溯搜索、贪心、动态规划...

算法面试的准备范围

- 各种排序算法
- 基础数据结构和算法的实现：如堆、二叉树、图...



算法面试基础问题的例子



初始序列为1 8 6 2 5 4 7 3的一组数采用堆排序，当建堆(小根堆)完毕时，堆所对应的二叉树中序遍历序列为：()

- A. 8 3 2 5 1 6 4 7
- B. 3 2 8 5 1 4 6 7
- C. 3 8 2 5 1 6 7 4
- D. 8 2 3 5 1 4 7 6

算法面试基础问题的例子



对一个含有20个元素的有序数组做二分查找，数组起始下标为1，则查找A[2]的比较序列的下标为·（）

- A. 9、5、4、2
- B. 10、5、3、2
- C. 9、6、2
- D. 20、10、5、3、2

算法面试基础问题的例子



一组记录排序码为 (5、11、7、2、3、17)，则利用堆排序方法建立的初始堆为 ()

- A. (11、5、7、2、3、17)
- B. (11、5、7、2、17、3)
- C. (17、11、7、2、3、5)
- D. (17、11、7、5、3、2)
- E. (17、7、11、3、5、2)
- F. (17、7、11、3、2、5)

算法面试基础问题的例子



在图采用邻接表存储时，求最小生成树的Prim算法的时间复杂度为()

- O(n)
- O(n+e)
- O(n^2)
- O(n^3)

算法面试的准备范围

不要轻视基础算法和数据结构，而只关注“有意思”的题目

- 各种排序算法
- 基础数据结构和算法的实现：如堆、二叉树、图...
- 基础数据结构的使用：如链表、栈、队列、哈希表、图、Trie、并查集...
- 基础算法：深度优先、广度优先、二分查找、递归...
- 基本算法思想：递归、分治、回溯搜索、贪心、动态规划...

选择合适的OJ

选择合适的OJ

online judge

在线判题系统

慕课网「玩转算法进阶」

选择合适的OJ

不要选择过于偏向程序设计竞赛的OJ



选择合适的OJ



LeetCode

Online Portal for IT Interview

真实的面试问题

<http://www.leetcode.com>

选择合适的OJ



HackerRank™

Revolutionizing Tech Recruiting

对问题的分类很详细

<http://www.hackerrank.com>

在学习和实践做题之间，要掌握平衡

讲师：liuyanbobobo

慕课网《玩转算法面试》

解决算法面试问题的整体思路

讲师：liuyanbobobo

版权所有，侵权必究

慕课网《玩转算法面试》

注意题目中的条件

给定一个有序数组...

有一些题目中的条件本质是暗示：

- 设计一个 $O(n \log n)$ 的算法
- 无需考虑额外的空间
- 数据规模大概是10000

当没有思路的时候

自己给自己几个简单的测试用例，试验一下

不要忽视暴力解法。暴力解法通常是思考的起点。

不要忽视暴力方法

amazon



Adobe®



yelp

Bloomberg

LeetCode 3 Longest Substring Without Repeating Characters

在一个字符串中寻找没有重复字母的最长子串

如"abcabcbb", 则结果为"abc"

如"bbbbbb", 则结果为"b"

不要忽视暴力法

在一个字符串中寻找没有重复字母的最长子串

- 对于字符串s的子串 $s[i...j]$
- 使用 $O(n^2)$ 的算法遍历 i,j , 可以得到所有的子串 $s[i...j]$
- 使用 $O(\text{length}(s[i...j]))$ 的算法判断 $s[i...j]$ 中是否含有重复字母

复杂度 $O(n^3)$, 对于 $n=100$ 的数据, 可行

优化算法

遍历常见的算法思路

遍历常见的数据结构

空间和时间的交换 (哈希表)

预处理信息 (排序)

在瓶颈处寻找答案: $O(n \log n) + O(n^2); O(n^3)$

实际编写问题

极端条件的判断

- 数组为空? 字符串为空? 数量为0? 指针为NULL?

变量名

模块化, 复用性

对于基本问题，白板编程

讲师：luyuanbobo

慕课网《玩转算法面试》

使用语言



《玩转算法面试》

版权所有，侵权必究
liuyubobobo

使用语言





Swift



慕课网《玩转算法面试》
讲师：liuyubooboo
LeetCode
Online Portal for IT Interview
版权所有，侵权必究

慕课网《玩转算法面试》

讲师：liuyubobobo

版权所有，侵权必究
liuyubobobo