

# 代码面试&求职问路专题讲座



#### 左程云

华中科技大学本科,计算机科学与技术专业 芝加哥大学硕士,计算机科学专业 曾经在IBM、百度工作 刷题5年的算法热爱者 著有《程序员代码面试指南》,即将出版发行 QQ: 303233886 欢迎骚扰,我估计你们问的题我都刷过...



#### 笔试过程

考试内容为基础中的基础,看好书即可拿下,因为一般来讲卷面分数过35就可以拿到一个面试资格。

研发岗位笔试试卷题型分配为算法与数据结构部分(30%~40%),网络、数据库、操作系统、系统设计和语言特性等部分(60%~70%)

前端工程师的试卷几乎都是语言特性的题目(html/CSS/JavaScript/JQuery/Bootstrap等)

具体基于某一种编程语言的笔试试卷同上

算法岗位、大数据岗位的试卷算法与数据结构比例激增



#### 面试过程

- 1,介绍自己
  - 记住:
  - 1,干净利索,半分钟到一分钟的时长;
  - 2, 千万犯奇葩
- 2,介绍自己最得意的项目

#### 记住:

- 1,不要随便拿课程设计糊弄。
- 2,项目要么有业务上的亮点,要么是体现开源的特点。
- 3,最优的项目介绍是什么样子的?买个关子,接下来揭晓。



#### 面试过程

- 3、由难到易的几道面试题
  - 记住:
  - 1,题目模糊不清的地方迅速问,这可能是面试官在考察沟通能力
  - 2, 先把暴力解法与面试官进行简单讨论, 再思考和实现优化解
  - 3,与面试官谈论的基础是时间复杂度和空间复杂度
- 4,有什么想问的...

#### 记住:

- 1,这不是放松的时刻,其实这个阶段非常重要
- 2,接刚才的关子,最好项目是紧贴面试官日常工作的
- 3, 千万要自信, 并且对公司展现出热情
- 5,坚持住的最后一面

#### 记住:

重点考察的是表达能力以及能否融入团队,有面试题也是简单题

面试过程

面试第一原则: 让面试官喜欢你, 因为他们在选同事

special offer是个传说



#### 算法和数据结构题目

其他任何类型的题目见过就不可能错,这里要推荐一下牛客网:

- 1,网站上试卷很多。
- 2, 针对每道题目都有用户讨论。

为什么几乎所有月薪过万的公司都爱考算法与数据结构的题目?

- 1,真正代表coding能力
- 2, 因为特别不好准备, 所以一定程度上可以验出聪明人
- 3,人员淘汰率可控
- 4,如果你不是特别聪明,至少你可以选择勤奋

#### 刷题推荐的语言

选择哪种语言刷题,可能会面临这种语言的小面试题

推荐使用的语言:

C++, Java

不推荐使用的语言:

C, PHP与Python等脚本语言



#### 刷题过程中会遇到的困难

- 1, 教材内容不足以应付面试, 从知识到逻辑的练习都远远不够;
- 2,资料难收集,各种书,各种网站,各种博客,各种论坛,练哪个?
- 3,理解算法与数据结构的题目的门槛较高,过程虐心易放弃。
- 4,根本不知道自己到底刷到了什么程度,无法持久地进步。
- 5,因为时间关系,大多数同学根本做不到题目之间的关系总结。

总结起来一个字: 散



## 题目讲解即将开始!

本次讲座的ppt以及题目的所有实现代码请关注牛客网专版!

开场题目

折纸问题

请把纸条竖着放在桌子上,然后从纸条的下边向上方对折,压出折痕后再展开。此时有1条折痕,突起的方向指向纸条的背面,这条折痕叫做"下"折痕;突起的方向指向纸条正面的折痕叫做"上"折痕。如果每次都从下边向上方对折,对折N次。请从上到下打印所有折痕的方向。

开场题目

折纸问题

- 1,产生第i+1层折痕的过程,就是在每一条i层折痕的左右两边,依次插入"上"折痕和"下"折痕的过程;
- 2, 所有折痕的组织结构是一棵满二叉树, 在这棵满二叉树中, 头节点为 "下"折痕, 每一棵左子树的头节点都为"上"折痕, 每一棵右子树的头节点都 为"下"折痕;
- 3,从上到下打印所有折痕的过程,就是按照右,中,左的方式遍历整棵二 叉树的过程;



如何实现转圈打印矩阵?

如何实现顺时针旋转矩阵?

如何实现"之"字形打印矩阵?

- 1,不要去纠结具体位置该如何变化,看的宏观一点
- 2,转圈打印和旋转的问题用分圈结构
- 3, "之"字形打印用划分对角线的方式



根据最长回文子序列处理字符串

给定一个字符串str和它的一个最长回文子序列strLPS,返回字符串str在任意位置添加最少字符后,整体都是回文串的其中一种结果。

#### 例如:

str="AB1C2DE34F3GHJ21KL";

strLPS="1234321";

返回: "ABLK1C2DEJHG3F4F3GHJED2C1KLBA"



- 1, 依次找到strLPS中的每层"洋葱圈"
- 2,根据每层洋葱圈在str中找到每层的外部,左侧为left,右侧为right
- 3,将l+r(逆序)拷贝到左侧未设值的部分,将r+l(逆序)拷贝到右侧
- 4, 直到strLPS被剥完



#### 字符串旋转问题

给定一个字符串str,和一个非负的整数i代表str中的位置,包括i位置在内的左侧部分想移到右边来,i位置右边的位置想移到左边来,请写函数实现。例如:

str = "ABCDEFGH"; i = 4; 调整结果为: "FGHABCDE"

要求: 时间复杂度为O(N), 额外空间复杂度为O(1)。

#### 解法一:

先左部分逆序, 再右部分逆序, 然后整体逆序

#### 解法二:

- 1,根据i位置将str分为左右两个部分。如果左部分小,左部分就和右部分的右侧交换;如果右部分小,右部分就和左部分的左侧交换;
- 2,交换之后会产生新的左右部分,继续按照这种方式交换,直到左部分和右部分等长,交换之后就停止。

解法二的意义何在?



如果数组长度为N,判断是否数组中所有的数字都只出现一次。按下列两个 要求分别实现这一功能:

要求一: 时间复杂度为O(N)

要求二:额外空间复杂度必须为O(1)的情况下,请做到时间复杂度尽量低



#### 要求一的解法:

- 1,在遍历str字符的过程中,用map记录某一种字符是否出现过, 检查所有字符都不重复出现即可。
- 2, map的实现可以是固定长度的数组,也可以是哈希表。

#### 要求二的解法:

- 1, 先排序, 排序后相同的字符会贴在一起。
- 2,再遍历一次str进行检查即可。

关键在于对各种排序方法的深度理解。



定义新类型的二叉树节点如下:

```
class Node{
    public int value;
    public Node left;
    public Node right;
    public Node parent;
}
```

给定一个新类型二叉树节点node,它属于某一棵搜索二叉树,请找到node的后继节点。



- 1,如果node有右孩子,node的后继节点就是它右子树的最左节点
- 2,如果node没有右孩子,开始向上寻找
- 3, 当node成为其父节点p的左孩子时,返回p;
- 4,向上寻找的过程中,遇到p==null的情况,返回null即可;



#### 深度隐藏了算法原型

#### "小和"问题

有一个数组为{1,3,5,2,4}

1的左边小于等于1的值和为0;

3的左边小于等于3的值和为1;

5的左边小于等于5的值和为4;

2的左边小于等于2的值和为1;

4的左边小于等于4的值和为6;

小和=0+1+4+1+6=12。

给定一个数组arr,实现一个函数统计arr的小和。



## 深度隐藏了算法原型

- 1,对数组做归并排序
- 2,合并组的过程中,将小和累计出来

不申请任何额外变量,交换两个数的值

在数组中打印出现次数大于一半的数。如果不存在, 什么也不打印。

要求: 时间复杂度O(N),额外空间复杂度O(1)



```
arr = ...
int cand = 0;
int times = 0;
for (int i = 0; i != arr.length; i++) {
    if (times == 0) {
       cand = arr[i];
       times = 1;
    } else if (arr[i] == cand) {
       times++;
    } else {
       times--;
...最后查看cand是否真的是...
```

#### 上一道题目的扩展问题:

你有所有人的选票,每一张选票上都写着一个候选人的名字。如果有某一个候选人得票数超过总票数的一半,这次投票才有效,那个候选人会当选;如果没有任何一个候选人的得票数超过总票数的一半,这次投票无效,要重新投;你做为验票员有一个机器,可以判断任意两张选票是否投的是同一个人,但你自己不能看任何一张选票上选了谁;你如何设计验票流程,确定这次的投票有效;并在投票有效的情况下,把印有当选人名字的其中一张选票返回即可。



判断字符串的旋转词

对字符串的旋转操作描述如下:

例如: str = "123456", str的所有旋转词为: "123456", "234561",

"345612", "456123", "561234", "612345"。

给定两个字符串str1和str2,实现判断str1是否是str2的旋转词。

- 1, 生成str2+str2的字符串str2Double
- 2, 判断str1是否是str2Double的子串



大数据与空间限制

在包含20亿个32位整数的文件中,找到出现次数最多的数,内存限制2G,请提出解决方法

#### 大数据与空间限制

#### 解答:

- 1,了解哈希函数(散列函数)的概念和基本功能
- 2,对于哈希表一条记录占用多少空间有准确的估计
- 3,根据步骤2的估计,将大文件哈希成若干个小文件,然后求出每个小文件的No.1
- 4, 比较各个No.1, 找出最多的那个

关键:讨论一定要定量

#### 系统设计题

由若干服务器组成的集群一起提供服务,共同承担访问压力,请设计集群的 增减策略,使得当集群增加或减少机器时,数据移动的代价尽量小。

解答:一致性哈希算法



#### 刷题过程中不要刷的题

- 1, 先刷经典的、算法原型简单还有大数据的题目
- 2, 再什么都刷, 因为散。
- 3,在刷题的初期和中期(刷过的题目不足200道视为初中),不要去触碰以贪心算法作为最优解的题目,不要去触碰ACM重量级题目,尤其是与图有关的题目。

#### 举例:

- 1,AB城市道路攻陷问题
- 2, 拼接所有字符串产生字典顺序最小的大字符串



## 以下是求职问路环节