

## Google Onsite

- roomba, 扫地机器人, 给定几个api: `void clean()` 清空当前单元格, `turnLeft(k)` 向左转k次, `turnRight(k)` 向右转k次, `move()` 从当前方向移动到下一个单元格, which returns boolean value 让机器人去清理整个房间 Time complexity, linear in term of room space. 采用dfs的算法, 关键是如何表示每个单元格

- 一片楼房有高度, 在一个矩阵里。从正面和侧面分别都可以得到一个侧影的高度。在不影响正面侧面的高度的情况下, 最大的高度和  
思路: 比较下每一组高度的最小值, 和即为所求

follow up: improve time complexity 排序一下可以降维, use binary search

- > 给定一组char, 包含{.小括号 Eg. : `a(b(c){2}){2}d` will be decompressed as `abccbccd`. {}中间的是前面 ( ) 里的内容的重复次数。

思路: 用栈, 所有的往栈里压, 直到), 开始把上一个 ( 之后的全部弹出到栈外, 然后乘以倍数, 然后翻转, 再重新压到栈里, 然后直到最后, 全都弹出栈, 翻转, 就是结果了

- > LZ77压缩算法。原: ABRA KEDABRA KADABRA 压缩后: `ABRA KED|tag|len|offset|KA|tag|len|offset` Compress: `ABRA KED[0xFE 4 7] [0xFE 5 5]` (tag = `0xFE/254`, len, offset) 有压缩后的字符串, 还原原来的。时间复杂度 $O(N)$ 。tag就是一个标记, len是引用了多长, offset是这个字符串往前数多少个是同一个字符串。遇到`0xFE`说明遇到压缩的部分了, 第一个是压缩的长度, 第二个是压缩的offset, 比如第一个case, 从D开始往回offset到A, 压缩的string是ABRA

这个有一定的小trick, 一个是会有地址偏移, 一个是可以引用引用, 如例子所示。要考虑引用的地址偏移。

follow up: 原文中遇到tag怎么办. 我的答案`|tag|tag` 考官答案`|tag|0`, 压缩效率更好

- > 字符矩阵中, 只能上下左右, 找到完整的目标单词。

我首先说要一个visited数组记录, 然后妹子就问举个例子, 然后就想啊想, 想出来, 比如pop这样的, 需要visited

- 用链表模拟数字+1的操作
- lc 425 word square
- lc 299

follow up: 问有个字典的话怎么minimum guess 次数 楼主不太会, 讲了两个方向一个是guess 一次 filter 一次字典一个是随机guess 直到猜对非常多个字母再filter

- lc305
- longest common substring
- 给定两个查询字符串集合, 一个是比较短的字符串集合, 一个是比较长的字符串集合, 查询长的字符串集合里面包含短的字符串集合里面字符串的所有长字符串

用的是rabin karp方法, 然后, 因为是查找多个字符串, 所以加上了一个HashMap做优化

- > 给定两个球队, A和B, 最多进行 $2n - 1$ 场比赛, 当某个球队有限赢得n场比赛的时候, 比赛结束, 有一个玩家, 特别喜欢A球队, 他有m元钱, 总是买A球队赢, 如果买了x元钱A赢, A赢得到 $2x$ 元, A输得到0元, 返回一个矩阵, 使得最终如果A先赢n场, 那么始终剩余 $2m$ 元, B赢了始终剩余0元

问了下时间复杂度, 以及如何证明一定存在这样的方案

- 给一个BST, 和一个min以及max, 求在min和max之间的所有元素和

- 给定两个比较大的整数，用字符串表示，a 和 b。求a 和 b的差，返回字符串，假设a >= b

- 给N\*N的矩阵，implement两个function 1) query(x1, x2, y1, y2), (x1, y1), (x2, y2)分别代表左上和右下的坐标，要求返回矩阵sum? 2) update (x, y, val), 更新某个坐标的value hint: query数量远远大于update数量，且可能有重复出现的情况 类似lc308

- 很长的doubly-linked list，难以遍历，输入是array of nodes比如[node1, n2, n3, n4]，找出被这个array中的node划分的disjointed blocks有多少个？如果相连算一个，比如 n1 <-> n2 <-> n6 <-> n3 <-> n5 <-> n4，返回3

- 给一个dict.txt和readline() function， dict.txt中每行是一个英文单词，比如 a, an, the, tank, ten, bet, ant, cut，给一个char set 比如[a, t, n, e] 返回所有由这个set中char组成的最长单词，set里面每个字母只能用一次，这个返回就是["ant", "ten"]

- 还是txt file，这次每一行是[id, parent id, weight]，代表父子节点关系和edge weight，file有若干这样的行，问给两个id，求出weight sum?

- 设计一个公司员工管理系统，实现三个函数，在input stream里不断被调用（sequence is not guaranteed）

- void manages(String m1, String e1) { // 代表m1是e1的直接 manager，每个人都只能最多一个manager}

- void peer(String e1, String e2) { // 代表e1和e2是同级关系}

- boolean isManaging(String e1, String e2) { // 如果e1, e2有direct或者indirect的管理关系，返回true，otherwise false}

- deck of cards，给一个array比如【2, 1, 3, 4, 5, 4, 5, 6, 7, 8】，给一个参数k，比如k = 5，问这个array能不能被fully divided by k，并且每个divide后的subarray是连续整数，比如这个可以divide成【1, 2, 3, 4, 5】和【4, 5, 6, 7, 8】，返回是true or false

- 自己设计API，设计一个类似家族图谱式的东西，然后再给你两个人，问这两个人有没有血缘关系。

其实就是图的问题，找Strongly connected component(directed graph)，两个node是不在在一个component里面。

- > 设计一个股票系统。一个update，每次给一个价格和时间，一个correct，每次可以改某个时间的价格。然后还有三个api要实现，current()返回最近时间的价格。max()和min()分别返回最大值最小值。

一个股票的stream，每个时间点(timeline)都有股票的更新，实现两个function: 1.给一个timeline，删掉当前点的股票 2.给你一个timeline让你求出历史股票数据的最大值和最小值 我使用heap和hashmap做的

- 矩阵转换问题。给一个矩阵比如 a:1, b:2 b:2, c:3 然后转换成 a, b, c 1, 2, ..., 2, 3 第一行放key，第二行开始放每个key的value，如果没有这个key就放.或别的补上。

- 问题蛮绕的，实现一个grep类似的函数。比如 [A1+A2, B1+B2, C1], grep(A=A1), 就是返回包含A1的结果，比如 [A1, B1, C1], [A1, B2, C1]。比较tricky的是可以grep(A=A1, A2, B=B1)，这样就是包含(A1||A2)&B1。

- warm up给个string[] = a, b, c, d ... 输出[a], [b, c], [d, e, f, g]... power of 2的输出。

- 之后题目是给你一个有向图，一个origin，返回能从原点返回原点的有向图。比如 /-----\ A -> B -> C -> D (C可以返回A)。这个时候输入origin A，那返回 /-----\ A -> B -> C，把D去掉而已。

- lc 289

生命游戏, 这题跟leetcode不一样的地方是input是二维boolean数组。我首先new一个新matrix存放结果; 面试官要求降低空间复杂度, 不能new matrix; input是boolean不能用来像int那样利用一些冗余bit, 后来我搞了个滚动int数组, 保存当前行和下一行的count, 对于每个cell center给下面三格和右边一格 +1, 反过来也+1. 这样的好处是cell不会访问之前已经处理过的结果。面试官说that might work.

follow up 如果数据量特别大怎么办 如果分布在多个machine上 每个机器运算性能不一样怎么办 follow up: 1. 这个board能上下连起来, 也可以左右连起来, 如何找到下一阶段的状态。2. board太大, 可以多台机器处理, 那么如何切割board, 对于切割的一小块, 有什么需要注意的, 如何解决。(1. 连起来问题就是对边界取模; 2. 切割出来的小board要注意边界问题, 解决办法就是在他们外面再加一圈。然后切割的办法, 我先提出常规的按行和按列切割, 然后面试官说你这样切割, 一个machine就会有八个邻居, 跟八个邻居交流会很花时间, 如何解决? 然后给了一个列子提示了一下, 我就想到可以把多行弄成一行, 然后二分切割, 这样一个机器就只会有左右两个邻居。面试官说对的, 这就是他组里面在做的事儿, 我就顺便请教了一下, 然后就聊聊天, 结束了。)

- 给了一个数组求最小值的code 问其中某一行运行的期望运行次数 用的dp

- 给一个grid 其中有些格子被mark了 求整个区域里所有的正方形个数 (正方形不能覆盖被mark的格子) dp

- lc 340

Longest Substring with At Most K Distinct Characters

我一开始写的nlogk的 小哥说还有o(n)的 follow up 如果是stream怎么改 写两个function 一个accept(char c)一个getLength 但是accept是logk的 于是小哥要求o(1)的 然后写了个double linkedlist

- > 给两个string 判断是否是合法的encode的

follow up 如果给一个test string 和一串备选的string 判断是否存在一对合法的pair把所有的string转成一个pattern (比如 apple -> abbcdd, banana -> abcbcb) 然后string array建trie tree 搜索的时间复杂度就是o(k)了 (但我隐隐感觉 小哥原本准备的答案不是这样的。。。虽然他还是让我按trie tree写了。。。然后时间空间复杂度也能接受的样子)

- 有一个迷宫, 都是房间, 从一个房间进去, 目的地是另一个房间放有宝藏, 房间与房间之间的门要用钥匙打开。特定的要是才能打开门, 钥匙可能存在于某些房间里问你 能不能找到宝藏。 follow up: 如果迷宫的定义不是用tree而是用graph怎么解 (可能会走到重复的节点) 自己定义 input 和数据结构。

- calculator, (+ 1 2) 就是1+2, 总而言之就是operator在前, 2个operands在后, 由空格分开

- tilt maze

- 假设有个地图, 上面有N辆自行车, 现在我们有N个人 设计个算法, 使得每个人能以最短的距离取车。greedy有edge case

- longest increasing sequence in a matrix. increasing指的是相邻两个相差1。Leetcode 329

- Choose a random point from one single rectangle. Choose a random point from multiple rectangles, if there is no overlapping among them. Choose a random point from multiple rectangles, if there is overlapping among them.

我大概说了下思路，中心思想就是分割成一个个竖矩形。先把矩形分成左右edge，然后edge根据x coord排序，然后左边扫描线扫过去。事前可能还需要做下discretization，把所有坐标对应到不同的integer index; 扫描过程中可以运用线段树优化，使每次query和update控制在logN. 大家可以参考Atlantis这道POJ的题，是求重叠矩形的总面积，很经典的扫描线+线段树。

- House robber 2

- Given a vector of key, value pairs, choose top K (in term of value) and each key can only repeated r times.  $O(N\log(n))$  running time.  $O(N\log(k))$  running time.  $O(N)$  time.

- Given a positive number n, generate all valid strings which consist of n parenthesis pairs. For example, if  $n == 1$ , return  $\{ "()" \}$  For example, if  $n == 2$ , return  $\{ "()", "()" \}$

- If a string A can be derived from another string B, by shifting each character of string B same distance ('z' will be shifted back to 'a', which means, it works in a circular way). Given a vector of strings, return groups of strings, where each group of string can be derived by any other string within that group by character shifting.

- Given a vector of boxes, which is described by width, long and height, calculate the longest sequence of boxes where all boxes of this sequence can be placed inside each other. Return this longest sequence.

- 在一个巨大字符串流中找第一个unique char 我基本把我能想到的各种解法(one pass, two pass, hashmap, linkedlist)都写出来了。最后面试官还问我怎么用个 `HashMap<Character, Boolean>` 去做，我当时没做出来，比较tricky。后来面试官又跟我说了下做法，这样的话得two pass，没有我说的optimal solution one pass好

- lc 734, 737 737跟题目不一样的地方就是，原题是A->B, A->C能推出B->C，我被问的时候，面试官说不能推出B->C。所以这道题只能保证transitive, 没保证symmetric. 所以这个图是有向图，有向图不能用Union Find做，只能老老实实用DFS.

- merge N 个sorted element list的题，要写成Generic的形式

- lc 256 Paint House, 只不过他要输出path就是最少cost的具体方案。

- 把你有没有pass考试这条信息通过一群同学传给你妈妈，每个同学可以选择说实话或者谎话每个同学可以选择告诉另一些同学或者直接告诉你妈妈，然后最后你妈妈会告诉你最后告诉她的同学说了什么（即知道最后跟你妈妈说的同学是说了实话或者谎话）问最后有多少同学说了谎，撒谎的同学的概率。

- 一个数，如果是偶数就除以2，如果是奇数就乘以3加1. 直到出现循环重复。然后第一问就是给个数然后这个数到循环前要多少步。第二问好像是给个数看是否这个数会出现循环(记不太清了，但是问题很简单不难)，然后就是问了下如果优化。

- 3sum smaller

- leetcode 490 the maze

- leetcode 394 Decode String

- 已知英文字母的摩斯密码映射（一个map, key是摩斯密码，value是一个字母），现给一串摩斯密码，求能得到哪几种解码？

- Leetcode 375

- 在电话拨号盘上，按国际象棋的骑士走法/中国象棋的马走法，限定走的步数，问一共有多少种走法

骑士走L形。给定十个数字，layout和Cell Phone上数字一样。然后给定一个长度，按照骑士L形走，有多少种走法。DFS + Memoization

- 给定一个unsorted array，然后给一些ranges, 每个range有起点的index和终点的index，判断每个range是不是等差序列

- Leetcode 403

- Remove duplicates characters and keep relative order.

- 把一棵随机的树，变成一个string序列，使得这个序列能够reconstruct the random tree. random tree 就是说这棵树的任意节点可能有多个子节点

- 颜色那一题, "red" 可以写成 rgb(255, 0, 0), 可以写成"#FF0000", 也可以写成 "#F00" 的缩略形式, 其中给了一个dictionary 的map, 你可以用map<key, value>, key 是"red", value 就是"#FF0000". 其中涉及到把10进制数变成16进制数

- 滑动拼图游戏，向一个方向滑动，使得相同的数字合并，比如 2 和 2 可以合并成 4 [ 1, 1, 2, 2 2, 0, 3, 3 0, 0, 1, 2 3, 3, 1, 2

0 代表是空的不能合并 上面数组向左滑动就变成了 2, 4, 0, 0 2, 6, 0, 0 1, 2, 0, 0 6, 1, 2, 0

- 一堆list，比如[a, b, c], [e, f, g], [d, e, f]... 实现一个class，有hasNext(), next()两个方法。next()按这个顺序：现打印所有列的第一个，然后第二个；用一个queue装iterator实现了。然后让添加prev()和hasPrev(). 注意的是，prev之后的next是前一个prev的next，不是新的next；提出了两个解法。细说了stack的方法。

- 给一堆string，顺序重要，从中找一个target string。比如cat: ["abc", "at", "ef"] -> true; 最简单的方法就是把所有string merge，然后找substring是target的。别用那么大的空间，你可以两两合并。不能用额外的空间，我就想给了一个最基本的遍历解法。本来还想说kmp的优化

- 实现一个数据结构：get(int key), put(int key, int val, int ttl) 就是一般的hashmap，但是每个key有expire的时间，如果expire了就输出-1；很简单对吧。

- 然后就是一堆coin，让你找所有可能的sum。combination问题很简单

- 让设计google doc的comment功能。大概意思就是doc的某一行可以插入comment，comment有高度。如果跟doc并行的那一个comment的行被占据了（比如：之前插入的太高了），你就往下找available的空间，再把后面已经插入的comment往后推。然后我觉得是变型题lc715。然后和面试官讲了一下大体思路，然后他说你答对了，写代码把。这个比较难写，特别是中间在循环里计算往下推的距离。大家写下这道题把。我运气好的事，面试前一天才看过，所有印象比较深。中间问了一点优化，和实现delete的想法，都答出来了。

- 有一道密码门，设有二进制密码，事先不知道密码，也不知道密码有多少位。密码输入只有0和1两个按钮。系统内部设有一个buffer会存储输入，当输入达到和密码相同的位数时，系统会比较输入和密码是否匹配，如果匹配，则门开；否则没有任何反应，并清空buffer。问如何试出密码？

首先，在不知道密码是多少位的情况下，这个题是很难解的。所以这道题的基本思路就是假设码是K位，然后要么证明假设成立，进而解出密码；要么反证密码不是K位，排除这种可能性。

从K = 1开始尝试，输入0和1两个bit，如果门开了，那么密码在0和1之中；如果没开，则码不是1位的。K = 2，尝试所有2个bit的组合00,01,10,11（01在K=1的时候已经试过了，可以不用试。但为了简化讨论，假设可以不用去管以前的输入）。同样的，要么门开，要么密码不是位。K = 3，这里有一个tricky的地方，就是在尝试所有3位的密码组合之前，先要清空buffer在尝试K=1和K=2的时候，我们一共输入了10个bit，假如密码真的是3位，那么此时buffer还留有1个bit，需要输入任意两个bit的值去清空buffer。然后才可以开始尝试所有3位的组合000,001,010,011,100,101,110,111。。。

Follow up:当我们尝试到某一个K值，在输入某一串长度为K的密码后门开了，那么密码一是我们最后输入的那K个bit吗？

不一定。因为真正密码的长度可能为K, K+1, K+2, ... 假设把我们在整个尝试的过程中有输入的bit拼在一起构成一个n个bit的输入流，最坏情况下密码可能是我们在不断尝试过程中按过的所有键的总和。目前我们唯一能确定的是真正的密码一定是这个n个bit输入流的某个后缀且长度至少为K。另外因为门开了，所以此时buffer一定是空的。那么在门开之后，何找到正确密码呢？

首先假设密码就是K位，把最后输入的K个bit重复一次，如果门开，则确定密码；否则密码不K位。然后假设密码是K+1位，此时需要输入1位任意值去清空buffer，然后尝试长度为K+1的后缀。。。以此类推

- 给一个LinkedList，返回倒数第n个node
- 给一个complete tree，返回最后一个level的叶子（用二分的思路）剩下的时间讨论了下类似于unique tree的题

• 一个matrix，里面的每一格是白色或者黑色，已知黑色的block都是connected的，让求matrix里的黑色格子的上下左右四个边界。我当时说DFS遍历，过程中记录四个边界，她说有没有better solution,然后提示我说如果是一个一维的你会怎么做，然后我就想到了二分。。。这一轮有一个比较trick的follow up，我硬是想了快20分钟，就是在找左右边界的时候我用了两个function，面试官问我能否把这两个合成一个，然后中间提示我说这个board里面只有黑白两个颜色，让我往颜色的方面考虑，可惜直到最后我也没想出做法。

- string to integer
- excel sheet column title
- 输入常数N, 生成 $1 \sim 2^N$ , 然后头尾两两合并，直到最后只有一个数组，暴力解就可以了

- find the only one peek in integer array, 答: binary search
- List Of String, find out the first pair with common unique characters (ABCCC和CAB就满足条件，应为他们都有ABC)

One for loop + HashMap<String, Integer>, 每进来一个word单独处理一下就好了 印度人曰: HashMap 太浪费空间了，提升一下

答: 不要任何数据结构了， two for loop生比吧 印度人曰: 太浪费时间了，中和一下吧

答: 用个set吧, 只要set里面有了, 就跳出loop, 和前面的比一下, 找到了就返回, 时间是linear的 印度人曰: what are you talking about, I didnt get? 印度人曰: 再提升一下效率

答: 还是生比, 用一个int[256]把已处理完的存起来 时间到, 没空写follow up的代码了

- Predict Search Term - Trie problem, 返回List<String>, 要求用线性的时间完成, 类似于leetcode Word Search 2 follow up: time/space complexity analysis follow up: 每个单词有权重, 权重大的在前面 follow up: 尽可能减少排序的时间, 答: 建树的时候每个节点放一个PriorityQueue follow up: 如果权重可以更新的话, 怎样用logN的时间完成, 答: 实现自定义的priorityQueue, 可以randomAccess任意节点

- LeetCode 399 follow up: 如何只用线性的空间复杂度, 而时间复杂度仍然是constant; 答: 所有的点连到一个中心点上, 这个点相当于一个transit, 和压缩过路径的union find有点像

- lc 159变形。find the longest substring with at most two instances(a character provided) in it

- 数组里不相邻的数字相加可以得到的最大值

- 给一个平行于x轴的矩形, 要求随机生成一个矩形内的坐标点。follow up: 给一堆不相交的平行于x轴的矩形。要求随机生成一个坐标点在任意一个矩形内。落在某一个矩形内的概率和矩形面积成正比

- 判断两个query是不是同义。给一个list的同义词数组, 比如[[fast, quick], [rate, percentage]], 输入两个string queries, 根据同义词数组来判断这两个query是否同义

- 包装过的merge intervals。add bold tag 给一个string和一堆substrings, 如果substring在string里, 就把string里面的substring那部分wrap上<b></b>。比如给string: “abcde”, substrings: ["ab", "d"], 要求返回 “<b>ab</b>c<b>d</b>e”

- 一个游戏, 有个board, n\*n吧, 给定颜色, 至少三种颜色。然后用户可以用手指交换相邻颜色, 让水平或垂直方向如果三个cell颜色一致, 就消掉, 可以得分, 然后换新的颜色, 继续玩 要求, 初始化这个board, 初始化情况, 不允许有水平或者垂直方向三cell的颜色一致, 另外要随机选取颜色, 保证概率一致。自己写class, members什么的

- 设计一个iterator, 读取一个数字串, 比如1, 36, 7, 2, 8, 10, 9, next() 返回大于5的数, 所以: hasNext() return true, next() return 36, next() return 7, hasNext()return true, next()return 8, next() return 10, next() return 9, hasNext() return false. 我感觉类似lc上的那个peek iterator

- 给一个rand() 函数可返回[0,1) 的double, 求半径为1的圆内的点, 当然是概率一致。我开始没理解意思, 因为他说了圆外的点, 我就老纠结产生x, y看满不满足 $x^2+y^2 \leq 1$ 的条件, 如果不满足怎么办这个去了 后来他告诉我, rand函数产生半径, 还可以产生, 0-360度角度, 然后sin, cos求x和y, 好吧。这个确实想不到。

- lc上的serilize和deserilize原题, 要求设计一个serialize和deserialize任意string的方法。要写code

- 给一个word和很多很多doc, 返回包含这个word的list of doc, 我用trie tree做的, 面试官好像没意见

- domino。每个domino 2个数字组成, 比如12 或者65 就是一个domino, 每个box有5个 domino, 比如71 89 23 64 19 21 让你设计一个class的addBox, 如果原先出现过就返回

false, 否则添加新的box, 返回true 举个例子, 71 89 23 64 19 21 和12 91 32 89 46 71 其实是一个box 其实最后要override hashCode和equals两个method

- 然后写一个判断两个binary tree 是不是similar tree
- 有个tree, 只有最下面的节点有值, 判断两个Tree 最下面所有的节点的值和排序是否一致, 要求space小于N
- 先是same binary tree, 再把tree变成graph, 判断是不是same
- number of island 变种, 围棋盘, 求下一子能够把对方围住的棋子数, (存在多个可能性, 比如下在A点能围5个, 下在B点能围6个, 求最后总数)
- 一道数学题, 一道逆向的BFS读tree
- leetcode原题: binary tree中的最长连续数列
- 给一个list of string (lowercases only), 要求返回归类后的结果list of list of strings, 归类方式如下: 如果str1里的每一个字母和str2里的对应字母, ord值距离相同, 两个str1, str2为一类。比如: input为 ['a', 'bb', 'b', 'aa'] a, b 为一类 aa, bb 为一类
- 给一个list of int, 找到一个平衡点的index, such that: 这个点左边所有数的和等于这个点右边所有数的和。 follow up: 写成generator的方式 javascript
- 地里出现过。给一个list of (时间点(int), 候选人名(str)), 给定时间点t, 要求返回在时间点t之前的票数最高的候选人 follow up: 同样的输入, 要求返回list of票数前K高的候选人。 follow up: 输入不给时间点t, 给你一个list of候选人名, 要求返回是在哪个时间点统计出的。
- 给你一个API, 输入4个经纬度的值(度数), 表示地图上两个点构成的区域(长方形), 可以返回最多50个在这个区域之内的objects。多次用同一些input返回的结果不会改变。要求写出怎么用这个API又快又好地拿到地球上的所有objects。  
follow up: 设计一个存数据的方法, 满足用户查“离我最近的星巴克”之类的query。
- String找最短substring 使得substring中的字符只包含 'a' 'b' 'c' 一次, 其他字符不管。没有就返回 ""
- Base64 encoding. 给你一个0-63及其对应的可打印字符列表。写这个对应函数。再写encode一个文件的函数。规则是每6个bit encode成一个byte的可打印字符。如果最后剩余不够6个字符怎么处理
- 带权随机数。除了用tree map(java) map(c++) 搞lower\_bound还能怎么整呀?  
follow up: 是每随机出一个数 这个数的权降低1. 比如 1=4 2=2 3=3 4=1 随机到1后 1的权变成 3 1=3 2=2 3=3 4=1
- 象棋马 8\*8 给一个起始位置和一个step 每个move按 2+1的模式(一个方向走两格, 换个方向再走一格。一起算一个move)走。走step个move一个有几种走法。时空复杂度多少
- 设计两个函数, cyclecount(num, mod), cycleHistogram(low, high, mod). 大概意思是, cyclecount(num, mod)做 digits square sum 然后 mod 操作。例如 mod(12), mod(5) -> (square(1) + square(2)) mod 5 = 0 -> square(0) mod 5 = 0. Stop return 2。做完digits square sum 之后就取mod, 直到mod 和之前的有重复形成cycle。Return 形成 cycle 之前的大小。CycleHistogram(low, high, mod) 就会给一个 [low, high]. 然后return一个histogram里面储存 [low, high]的数目里面cycle 大小为 1, 2, 3, 4, 5。。的有几个。 Follow up: 优化cycle的 count (memorization)



- 给定 grid 的 width, height, 起点, 终点, 问起点到终点的路径有多少个. 从 (i,j) 只能向 (i, j+ {-1,0,1})移动。 follow up: 如果有一个list的点, 你必须经过这些点才到终点路径是多少?

- 国人大哥, 给两个curve, c1, c2 上面的arbitrary若干点, 问题1) 设计一个两个 curve similarity的函数 (open ended), 2) 给定一个距离函数, similarity的定义是 每个在c1上面的点可以align到若干c2上面的点, 同样, 每个在c2上面的点可以align到若干c1上面的点. 在这所有的alignment当中sum of distance最小的 为这两个的similarity。 这个方法的复杂度是多少。 3) 用dp写计算这similarity的代码。刚弄完2, 三刚刚写完recursive relation 还没代码, 就没时间了。铁挂了这一轮。 dynamic time warping

- 设计一个猜词游戏, 五个空, 一个dictionary都是五个词, 如何设计heuristics让第一个猜词更有效, 就是猜到正确的字母才会在空位显示。 例如 要猜的词是apple, 目前状态为 \_ \_ \_ \_ \_ 猜paper后变为 appe \_

- lc 329 longest increasing path in matrix.

- lc 406

- lc 503

- lc 552

- valid palindrom 只考虑字符有效 follow up 大小写敏感

- merge k sorted data stream. 这个要沟通输入给的是什么, custom 了 stream 还有存入minHeap 的data. 问了复杂度

- 出了一道把0都移到前面的题, 我问他输入是array 还是list 大叔说随便, 问我有啥区别 swap 的话没什么区别, 大叔开始以为我要insert 然后问没区别? 然后开始问如果是用户要提 5 或者 prime 或者 even 怎么办? 开始我说写一个isValid function 在里面修改要求, 又进一步构建了一个类 里面有 isPrime() 啦 isEven(). 然后大叔说如果不是test数字怎么办, 我说用 general type 改成了大 T. 然后大叔又说改成大 T 这些function 该怎么办。其实就是构建个interface 然后让different type 分别去继承, 我当时没get到他是想让我答这里。

- 问了一道  $m * n$  数多少个square 的题. 开始还以为要用dp 在格子上标了数据, 跟小哥解释说看看有没有规律, 然后思路来了。 只要traverse 边就可以了, 外层length 里面i,j 循环。小哥让优化, 然后推了公式乘起来, 但还需要最外层一个 loop,  $O(n)$ . 小哥让继续优化, 我就开始推通式, 然后推了一会儿小哥说我们不在这儿推了 follow up: 矩阵中有 invalid 的格子, 有这些格子的都无效。给了最naive 的解法, 每次去check当前square 包不包含invalid的。因为小哥迟到, 这时时间已经差不多了下一个面试官在门外了, 小哥说你要不写几行代码, 然后就简单的写了几行解释了一下, 白板上大概都是推的过程, 没写什么代码这轮。表示数学已经连小学生都不如, 求别虐。

- 一道题 spiral matrix. 开始讲思路, 写好了每个index 的range. 然后写代码。 写完之后带了个test case 进去走了一遍。然后大叔开始问怎么test, 我想着不就是写sample test 带进去打断点么。然后大叔说如果别人用你的代码发现有bug了怎么办? 后面经提醒才说了unit test。test 方面知识匮乏, 求指点。然后国人小哥说, 哎? 我发现你的代码里有个bug, 我写的时候带的是  $m < n$  的, 小哥说  $m > n$  你的代码就不对了, 让我举test case 然后我跑了一个没问题, 然后他说不用你这个, 再想一个。我就有点懵, 明明没问题, 然后我问小哥你来给一个, 然后他说你自己给, 我又给了一个还是没问题, 我说你给吧。。然

后小哥给了一个，跑了还是对的，然后小哥说哦是他想错了，写的没问题。中间推过临界条件，然后又讲了一遍那里。然后问了两个问题结束

- 交换二叉树的两个子树认为他们是相同的，判断两个二叉树是否相同 递归秒了 follow up: 复杂度，如果是general的tree怎么处理，复杂度 递归里加了个backtrack判断孩子们是否一致，加个memo记录热议两个subtree是否相同加速

- 热身move zero 变种 一开始问说不要maintain order么，大叔说不用。然后解法大叔不满意，说那要是maintain order呢，然后给出标准解法。

- 给一堆<child, parent1, parent2> 求一个人的祖先 bfs秒了 follow up: 求两个人是否有血缘关系 答求intersection。大叔不满意，以为要优化复杂度浪费了时间。最后讨论了一下双向bfs过关。

- 国人小哥，给一个function， boolean isErase(TreeNode node)，告诉这个节点该不该删除，然后写一个function，输入是root，返回一个list，里面是删掉所有该删的节点之后，形成的所有新树的root，比如树1，2，3 删掉root 1之后，返回2，3

- 然后写一个树的iterator，inorder，很简单，然后再写个function判断两棵树是否是xxx tree，具体是个啥词不记得了，总而言之就是当两棵树一样或者对称的时候返回true，否则返回false，比如1，2，3和1，2，3或者1，3，2都返回true，这轮很简单，印度小哥人也很好

- 美国小哥，L代表late，A代表absence，O代表正常出席，input是一个string，包含L，A，O，要求不能连续3次late，不能超过1次absence，就可以reward这个学生，判断这个学生的出席记录string能不能reward。 follow up: 输出长度为n的rewardable的出席string的数量

- Validate Pangram: 其实我都不知道什么是Pangram，他解释了一下，其实那个就是指这个string包含所有的字母，比如著名的A quick brown fox jumps over a lazy dog. 需要注意的就是标点符号和空格的问题。之后聊了聊怎么支持其他的字母，可以从ASCII本身考虑这个问题

- Validate Toeplitz Matrix: 我也不知道这是什么matrix，其实定义很简单，就是一个斜线上面所有的数字一样，坐标(i, j) == 坐标(i+1, j+1)。

- 本轮题目是设计两个API，情景是一个机器向另外一个机器发packet，packet的format是offset, size, data。这个设定其实就相当于interval，offset就是start，size就是interval的长度。第一个API就是类似于receive，这个function作用就是存下来发过来的packet，可以用global variable。第二个API就是read bytes，有点类似于LC158，input你希望最多读到的byte，返回最终读到的byte数目，读出来的数据需要是连续的。 Follow up: 就是如果发packet的机器出了问题，导致发过来的packet所对应的interval有overlapping，然后问你怎么办，其实就相当于LC merge interval那道题。

- Find all duplicate subtrees in a binary tree。这道题我之前好像看过，不过当时并没有仔细看，也没有做过。我当时的做法就是先找相同高度的subtree，用hash table存起来，然后再写一个function来比较两个binary tree是否一样。这个办法挺蠢的，但是我也没什么办法。

- 涉及一点DOM的知识，不过其实无所谓，本质是一个tree的问题。题目相当于是说比较两个DOM所有leaf node上面的string concatenate起来的结果是否一样。做法就是用DFS就可以了，recursion很好写。 Follow up: 就是如果现在考虑内存不太够的问题，也就

是说你不能把整个concatenate之后的string都算出来再比较，应该怎么办。我当时说的做法就是iterative DFS，然后就能一个一个读leaf node上面的string了，这样来比较两个string，有点类似把整个过程变成stream。

- 号码牌倒置歧义问题：已知有  $N$  个运动员，每个人有一个号码牌，找出哪些号码上下颠倒以后会有歧义的

- 给定一个平面上许多点的坐标，找出以其中四个点为顶点可能组成的所有矩形中，最大的矩形的面积。这里矩形只考虑边和坐标轴平行的情况。

- Leetcode 某原题，word ladder 给了源单词、目标单词和一个字典，每次可以从单词里替换掉一个字母，得到的单词必须也是字典里的单词，返回最少需要几步可以变化到目标单词。

- 给定一系列已经排好序的数字，和一个二次函数  $f(x) = ax^2 + bx + c$  的三个参数  $a, b, c$ ，将序列中的每一个数字用二次函数计算后得到的结果有序输出出来

- linkedlist, 找出最大的两个, 然后swap, 不是swap value, swap reference

- RGB颜色转换比如现有#2f3d13，有16进制的00,33,66,99, cc, ff.要把现有的数字转换成最close to这几个数字。比如#2f3d13 -> #333300

- find Nth smallest element in BST

- one edit distance

- wall, flower, matrix找到能看见最多flower的点，地里高频题 给一个two D garden , 每一个slot可以是flower或者Wall. 找一个合适的位置，让游客可以看到最多的flower.可以站在flower上，不能站在墙上。。如果被墙挡了，就看不到墙后面的花。然后游客只能竖直或者水平看，不能看对角线。。比如 [ [f, x, x, w, f], [f, f, x, x, x], [x, x, f, w, f], [f, f, x, w, x]] 这样，{3, 0} 和 {1,4}都能看到四朵花。开始说了个brute force的方法，最后优化到 $O(n^2)$ 分别水平还竖直的便利整个matrix, 记下这行的flower个数，碰到FLOWER就加一，然后每一行碰到墙了就把之前的全部更新, 然后 flower个数reset 到 0. 然后水平和竖直便利后的相加。最后找最大就可以了。。

- count of smaller number before itself

- 给很多intervals，每个长度都是1，起始是浮点数，求intervals包含的总距离。很自然写了merge intervals，然后要求 $O(n)$ 。解法先不说，大家自己想想，hint是长度都为1。接着问如何分布式处理。

- 一条直线上有很多加油站，现在可以再添加K个加油站，要求最小化相邻加油站之间的最大距离，输出添加的位置。

follow up: 问如果不是直线而是环怎么办，基本是一样的。用greedy做的，每次取出最大的间距，往里面加，但此时只记录添加的个数，它的距离用初始间距除以个数加1得到。

- 给一个数组，从右往左读，每次遇到一个数，输出右边第一个比它大的，没有的话输出-1。比如[2,3,1,4,5,3,7]，输出[3,4,4,5,7,7,-1]。因为假定是数据流，所以不能从左边开始读

- 有一个string表示地址，比如1667 Plymouth St, Mountain View, CA 94043. 然后因为某些application的限制，字符串不能任意长，所以现在要把它变成长度为n的字符串。有下面两个规则：Rule1: digits > letters（数字比其他字母的优先级高）Rule2: 前面出现的比后面出现的优先级高 因为有两个rule，所以还有一个是 Rule1 > Rule2，就是rule1比rule2优先级高 题目也做了简化，输入是一个string，只有字母和数字，按照给的rule变成长度为n

的字符串。我一开始脑抽，直接说，那就把数字和字母分别扔到两个queue里，先取数字，再取字母，直到够n个。然后他和我说，还有个隐含条件就是要preserve order。我就。。。其实应该是这样的，比如：abcde123dhs1，n=5，返回a1231。我想了想，给了一种做法，就是从后面开始扫描，如果是数字，就先留着，如果是字母，就看当前的长度是不是超过n，如果超过，就将这个字母扔掉，直到剩下n个字符。还有要考虑到的情况就是一共小于n个字符，那就直接返回，如果数字多于n个，那就等遍历到头之后再trim一下。然后大叔问了时间空间复杂度。如果字符串的长度是m，那么这个时间空间都是O(m)。然后让我把空间优化到O(n)。最后想出来要先count数字的个数，然后再从前面开始构建就可以了。然后就写代码了。写出了bug，大叔指出了之后就改过来了。

- 存同义词的字典，比如 A:B, C(A,B,C是同义词) D:F B:C 怎么把它们存下来，然后还要支持给两个单词，返回他们是不是同义词。这个一看就是union find，小哥问了时间空间复杂度，然后就写了。最后look up没让我写，说太简单了。然后也没什么follow up，因为我把union find里面减少树的深度的操作也写进去了。看来leetcode能用union find写的题一定要用union find写一遍！背下来！

- 出的题就是longest consecutive sub array, follow up: 是把这个变到tree里面

- lc 281 有k个iterator。构建时候的输入是Iterator<Iterator<Integer>>，然后让我写这个class的constructor，next和hasnext。这道题就用一个queue存就可以了。写完之后跟我说class的signature有问题，然而就“class Interleaving Iterator”能有什么问题。最后搞明白他是想让我implement interface。然后我就加上了implements Iterator<Integer>，然后next()的返回值改成了Integer。然后又问我是不是thread safe。我前两天刚看过，一高兴直接跟他说不是，在class里面定义的queue被多个thread同时访问就不行了，要加volatile关键字。然后他没做任何comment就问别的去了。然后我面完了又想了想觉得不对呀，要把next() synchronize 才行吧。我也不知道到底应该怎样，也不知道大叔怎么想的。。还有他要求next() 如果call的时候是空的要throw exception。问了时间空间复杂度。

- `` class Node { Node getParent(); Node getNextSibling(); Node getFirstChild(); } void print(Node root) { Node cur = root; while(cur != null) { System.out.println(cur); cur = nextNode(cur); } } Node nextNode(Node cur) { // 让实现这个function，然后调用print时候能打印所有的node } `` 问他用什么order打印他说随便，我一开始没想明白，直接先给了用queue的BFS做法。然后他让我不用extra space，我就又想了几分钟，最后给了一个preorder traversal的做法，写完然后跑了test case就完事了。不过事后想想好像postorder更好写一点。

- Museum map: Given a map (oftype char[][]) of a museum where '.' stands for an empty room, 'G' stands for a guardian and 'L' stands for a locked room. A guardian is able to reach neighboring(up, down, left, right) empty rooms in 1 move, but can not enter a locked room. Return how many moves the nearest guardian has to take to reach each empty room. The return value should be an int[][] whose size is the same as the input map. For an empty room mark the corresponding cell with -1 and for a guardian mark with -2. 讨论：地图是不是永远valid，是不是静态的，是不是可以放到内存里；guardian的数量相对于整个地图room的数量是不是trivial的。回答：地图永远valid，静态的，可以放到内存，guardian数量远远小于room的数量。我的做法：一开始想法是对于每一个guardian做一次bfs，得到他到每一个room的距离，然后把所有guardian的bfs结果汇总，对于每一个room取最小值作为结果输出。面试官说可以让写code。写完发现bfs中有一个变量有bug，经提示改正。面试官说

code应该是对的，但是不够efficient。我说似乎可以在一开始把所有的guardian加到bfs的queue中，然后一次dfs就可以完成。面试官说可以。没有要写code。让我问了些问题，第一轮结束。

- Given the root directory of a file system (represented by an-ary tree), return all the directories and files (return as List<Node>). 讨论：具体的Node的表示方法，输入不合法需不需要handle等。回答：可以用类似Leetcode那种方法表示，需要handle不合法输入（该情况下其实只是根节点是null）。我的做法：用Queue implement一个BFS完成。跟进1: what if the file system has symbolic links? (i.e. the tree is now a graph). 要求在原来的代码上进行修改。我的做法：保存一个hashset表示已经访问过的节点避免重复。跟进2: reconstruct the file system in another drive w/o the symbolic links (i.e. deep copy the tree) 我的做法：将上述hashset改为hashmap，key是原来的树里需要copy的节点，value是复制后的节点。依然用BFS进行deepcopy。跟进3: What if there are symbolic links? (i.e. deep copy the graph)? 讨论了一下做法，依然保留上述hashmap，如果hashmap中有的key-value-pair就不需要复制。不需要写code。

- 题目：Given a sidewalk with length 100.0, and a stream of rain drops, assuming the length of a rain drop is 1.00 and the rain drop could fall randomly anywhere w/in the sidewalk, return the number of raindrops until the sidewalk is all wet. 讨论：raindrop stream是不是无限的，raindrop stream api的形式，raindrop会不会滴到sidewalk以外，区间开闭等。回答：是无限，类似Iterator，不会滴到sidewalk以外，区间为闭区间。我的做法：一开始不太有想法，想到是区间问题可能是用树比较合适，跟面试官讨论面试官说可以你先写写看（这里吐槽一下google的面试官，大多情况下只要我有想法都会说你开始写吧，而不会先讨论清楚具体怎么写）。写着写着发现用treemap确实可以写的通，用treemap存sidewalk上已经湿了的区间，其中key是区间左端点，value是区间右端点。对于新的raindrop调用treemap的api ceiling和floor找到左右区间，分类讨论看能不能和左右合并并分别处理。写完跟面试官交流，被面试官质疑分类有不完整的地方，检查了一下做了修改，和面试官讨论通过。跟进1: 如何测试。我的做法：新雨点不和旧区间overlap: raindrop左端点的位置依次是0.00, 1.00, 2.00..., 返回值应该是100. 测试新雨点与左边区间overlap: raindrop左端点依次是0.00, 0.10, 0.20, ..., 返回值应该是1000. 测试新雨点与右边区间overlap: 99.0, 98.5, 98.0 ...; 新雨点与旧雨点完全重合: 0.00, 0.00, 1.00, 1.00.....等 跟进2: 时间空间复杂度 一开始我说时间是 $O(\lg n)$ , n是树里面的区间的个数，空间是 $O(n)$ 。面试官说再看一看条件，发现树里面的区间不可能超过100个，因此时间空间复杂度都是 $O(1)$ 。

- Given the root of a tree and a list of nodes that are about to be erased, return the forest (represented by a list of root nodes) after the erase. todo 我静静思考了几秒钟面试官立刻说你需要 think out loud. 于是开始胡言乱语说BFS，用Queue，又说erase好像不是很好操作，因为erase的节点的孩子还是要继续访问的，但是孩子又有可能是被erase的……可以在bfs出队的时候对孩子进行判断，但是又好像不太对……面试官看不下去了说你的想法应该可行，开始写code吧。写完拿了一个例子walk through可行。跟进1: 如何测试。我的回答：空树，[1,2,3]完全树根节点被erase，左子树根节点被erase，右子树根节点被erase，只有左子树左子树被erase等等。面试官说还需要测根和左子树都被erase。

- Given a list of Iterators, Design a class which implements the interface Iterator (i.e. include hasNext() and next()), which would iterate through the iterators in a round-robin way.

E.g., [[1, 2, 3], [4, 5, 6], [7, 8, 9]], next() 依次输出1, 4, 7, 2, 5, 8, 3, 6, 9. 我的做法: 存一个 deque, 每次next()从队首deque一个iterator, call这个iterator的next(), 如果空了就扔掉, 还有下一个就再添加到队尾。跟进2: 如果还需要hasPrev()和 prev() 怎么办。我说, 那就反过来好了从后面deque从前面enqueue.面试官说不行, 因为你有些iterator到头了就扔掉了。我说那不扔, 他说还是不行, 比如[[1, 2, 3], [4], [5, 6, 7]]就会有问题。我想了半天没听明白为啥接着问, 他又解释了一遍我好像明白了, 我说那再加一个variable表示iterator到了 level, 他说可以, 没有写code。

- LC332简化版, 每个城市只visit一次而且没有环。其实只要找到起点就好, 上来有点紧张, 解释了一下实质上还是拓扑排序, 然后就在想怎么能够缩减空间只存储有用的东西, 结果反而把自己搞晕了, 写的有点慢。然后大叔跟我聊了会儿美国大选, 聊了下加州的奇葩选举制度, 然后说我自创了一种选举方法(傲娇脸)——每个人选票给所有候选人排序, 然后每一轮只统计每张票的第一志愿, 如果有candidate超过50%, 那么就钦定他了, 否则将当前统计的票中的最后一名除去, 在所有选票中排在他后面的人名次上升一位。(大叔是真能讲, 还跟我说这个人除去了, 他的百分比就怎么怎么分散给了其他人...我一边听一边着急时间)然后让我设计这样一个系统。我是真的很懵比, 我以为要设计一个计票系统, 一直在思考怎么存这些百分比以及有人gg了怎么他的百分比分给别人。结果后来大叔说要我设计的是怎么存每一张选票, 每一轮完了, 踢了人以后还要重新计票的...我表示很无语, 然后说那就用linkedlist + unordered\_map来实现, 大叔比较满意然后说, 没时间了你说说看要实现哪些函数, 我说就一个string getFirstChoice()和一个void removeCandidate(string candidate)。大叔说好然后把我分析时候画的图拍下来了。其实我感觉这轮已经gg。

- 把binary array比如000110存成vector<node>, node.val是0或1, node.len是连续0或1的长度。然后给我两个这样的vector让我判断他们压缩的binary array是不是同一个。写了大概十多分钟吧。

- 把一个Binary的图用这样的vector<node>表示, 然后给我图的width, 让我把图镜像翻转。套路和Text Justification差不多就是硬着头皮一点一点处理。写完问了下我时间复杂度, 以及可能的优化。不过他全程都在敲代码, 不怎么主动跟我互动。

- median of a data stream, 我佯装分析了以后给了两个heap的解法, follow up: 是给题目加限制条件, 让我把添加data的复杂度降到O(1), 之前讨论过面经里这道题, 只要data有范围就可以用counting sort类似的方法了(复杂度和数据范围有关但是可以看成是伪常数级别的)。烙印最后说其实可以只限定median的范围, 其他可以不用限定。我露出膜拜状。

- 问了我面经题找最小矩形的, 我继续佯装思考然后给了O(n<sup>2</sup>)的解。follow up:, 满心期待会问我不与坐标轴平行的情况因为我之前思考过。结果问了个在第一问基础上找最小的内部不包含其他点的矩形。我稍微想了一下给了一个辅助的结论然后在这个基础上提出了一个O(n)的解法, 但是说完发现有点问题, 但是时间不多了我就说naive的解法是O(n<sup>3</sup>), 小哥说他其实也不知道。不过面完走出来我想了一下在我第一个解法的基础上改一下可以得到一个O(n<sup>2</sup>)的解。 <https://www.geeksforgeeks.org/find-rectangle-binary-matrix-corners-1/>

- lc 152, pangram 给定一个dictionary, 生成最短的字符串, 既是palindrome又是pangram。每个单词可以用一次然后可以用无限个单词。lc336的至尊无敌加强版, 也没有

写code一直在画搜索树然后分析time / space, 什么bfs, dfs, A\* search轮着来, 他也一直说make sense, good之类。一直就是探讨。有人知道标准思路是什么吗?

- count height of binary tree。我写了一个helper来完成recursion, 他问为什么不在主函数, 我说可以handle wrong input format raise exception之类忽悠过去。然后正题lc 109, 先用nlogn做了之后小哥说优化, 没注意过这道题的discussion, 想了半天在提示下写出了n的解法。

- lc 159, 装模作样想了下滑动窗口每次挪一个然后问time complexity说是否还有优化, 最后提示下想出来高票的nlogk解法

- 给定一个polygon, 边都是垂直和水平的, 输入是所有的点组成的1个array, 从最下然后最左的点开始。再给一个任意点问点是否在polygon里面, 点在边上也算在polygon里面。想到的方法是从点向上竖直做射线判断有多少个边与其有交点, 然后还没考虑corner case面试官就催说时间不太多我们先写出来程序。建polygon的时候有bug少考虑了最后一个点与第一个点的连线, 提示下改正, 然后面试官说了一个case发现无法handle, 就是射线与边重合的时候, 有无穷个交点然后数交点就数不过来, 想了2分钟面试官表示放弃吧太难了, 还有没有什么优化, 想了下说可以sort边。

- 在一个二维平面上的长方形(坐标取值精度是double类型), 怎样使得随意取得的点符合均匀分布。

follow up: 是如果有多个不同的, 且不重合的长方形时怎么搞? 形状任意的情况下怎么搞

- 就是找到一棵树里节点值之和最大的路径(节点值都是正整数)。路径可以包括根也可以不包括。另一道是树的最大高度。

- 找到top k的元素在一个数组内。问了很多代码怎样写好的细节, 比如别用太多map.get()什么的

- 王位继承, 长子及其后代有优先继承权, 然后是次子及其后代, 依次类推。完成birth(parentname, childname)和death(name)以及输出王位继承顺序三个function

- 给一个有向图返回所有从起点到起点的环组成的图, 输入只有一个起点

- 判断在字典中是否存在改动一个字母就可匹配输入字符串的单词, 这题时间复杂度没答好

- 字典中所有单词长度为5而且每个单词都没有重复字母, 实现一个function可以生成下一个猜测的词

- 21点游戏, 牌只有1-10点, 如果还没满17点就要再拿一张, 问超过21点的机率是多少? DP问题, 但面试时太紧张, 不知道怎么设定DP, 尤其是被一直抽牌有好几轮迷惑了。最后面试官提示, 只要用DP计算拿到我点的机率, 然后 $1 - (P(17) + P(18) + P(19) + P(20) + P(21))$ 就好

- 给一个板凳, 上面坐了一些人, 现在上来一个人, 求要坐哪个位置才能离其他人最远 我一开始只想到暴力解为 $O(n^2)$  想到后来才想到用一个矩阵纪录第i个位置左边有几个空位, 另一个矩阵纪录第i个位置右边有几个空位, 取最小值这样只要 $O(n)$  面试官问这样有k个人上来要依序找座位要 $O(kn)$ , 能不能更快?

- 费氏数列, 问说如果要查询 $1-2^{32}$ 任意一个但记忆体只能记 $2^{20}$ 个数要怎么做  $O(1)$  查询 这题我也是不太会只说了每隔k个就记2个要求f(n) 就从最近的开始算过去 面试官看来还能接受就要我算k应该是多少然后写代码

- 判断两个包含退格键的字符串是否相同

- 设计一个event queue。要求给定一个event，判断在这个event之前的一个固定时间段里有没有出现大于k个event。event可能是按时间顺序来也可能是随机来
- 24点（加减乘除）
- 一个车，可以收两个指令A和R。A就向前走一秒然后速度加倍。R就停下来然后反向。给一个AR组成的string求最后停在哪里。反问如果给位置求string。
- 一个奇怪的list，push的时候是push到头上，pop的时候按概率从头尾pop一个。问如果push了一个sorted list进去，怎么pop一个sorted出来。follow up: 二问如果pop是从头出来，而push是概率到头尾，怎么pop一个sorted。
- 给一堆长方形，随机返回长方形内一个点，概率要跟长方形的大小成正比。Follow up1: 如果要多次调用生成随机点的函数怎么办。Follow up2: 如果长方形有重叠怎么办？Follow up3: 怎么检测有没有重叠？优化检测方式和考虑下多个重叠的问题。Follow up4: 怎么检测这个函数是不是在正常工作？用significant test的时候在不同情况下怎么选p value？
- 中国妹纸给很多string, 找出它们最长的共同substring。Follow up: 如果函数多了一个parameter: threshold, 输出最长的共同substring。只需要threshold个词有这个共同substring就可以了。
- 给一个六位16进制的字符串代表rrggbb, 返回一个三位的字符串rgb, 要求他们两个的误差最小
- 二叉树，每个节点有一个字符，输出inorder顺序的string。Follow up: 有两颗树，判断他们inorder顺序输出的string是不是一样的。优化空间复杂度
- 很明显的一道union find题 给定一堆video playlist l1: [1, 2, 4, 5], l2: [2, 4], l3: [6], l4: [5]. 求里面有几个cluster。规则：1.拥有相同video的两个list属于同一cluster； 2 以上的property可以commute: 如果l1, l2属于同一个cluster, l2, l3也属于同一个cluster, 那么l1, l2, l3 都在同一cluster里。follow up: 如何优化
- KMP。两个数组 [1,2,3,4,5], [2,3,4,5,6] 求第一个的subfix和第二个prefix相同的最长情况，比如例子里就是[2,3,4,5] 长度为4。Follow up: 二维数组怎么做，如何优化。
- 设计一个不断接收TimeStamp的系统，并且在接收每个TimeStamp TSnow 的时候输出[TSnow - k, TS\_now]这个window中的timestamp 数量，已知Time Stamp都是有序的（先来的肯定小，后来的肯定大） follow up: 如果TS是无序的怎么做
- lc 778
- 类似course schedule 的题，拓扑可秒
- 给两个 string，寻找最小的 Damerau–Levenshtein distance（定义详见wiki，概括下来就是比edit distance多一个 transpose 操作），输出 shortest path。
- lc308很像的一道题，就是indexed tree
- 给个html的文件，输入用tree来表示，让你比较两个html文件显示出的东西是否一致 e.g. <html><body><p>H<i>ello</i></p></body></html> 打出来就是 "Hello\n", 这里不考虑<i>（斜体）的差异，"\n"是因为有<p>. lz用dfs做的 follow up: 如果string很长怎么办，lz想的就是不一次性吧整个string generate出来，一点一点generate，然后看两个html的前部分是否一个是另一个的substr



- lc原题的变种，就是那个两个玩家每轮从左右两端选一个拿走的题（抱歉没找到题号），原题是返回先手是否能赢，这个题让输出先手的人最多能拿到的数字和是多少。本质一样的，就是很简单的dp

- 还问了一道爆难的智力题（有可能是我太菜了），面试官一开始就说没指望我做出来。应该是因为我前一个写的比较快，就当玩考了考我（可能我天真了，其实这个才是重点？）。是给一个01组成的矩阵，每次可以click一个点，这样这个点周围的十字行的bit全都flip。问给一个matrix能不能通过这个操作变成全0。最后几乎是面试官喂给我答案。。。说是0肯定是被flip偶数次，1是奇数次，然后每个点click两次没有意义，所以只会被click 0或1次。然后就可以列出很多式子

比如 01 10 的matrix，用c1到c4表示每个点被click的次数。列出来的式子有：

$c1+c2+c3=0$  or  $2$   $c1+c2+c4=1$  or  $3$   $c1+c3+c4=1$  or  $3$   $c2+c3+c4=0$  or  $2$  至于怎么解面试官说的很笼统，就是把c1消掉什么的。。。

- 一道演员关系题：给一个list of movie, 每个movie 就是一个list of actor。实现给一个actor，返回list of actor that connected with it. (direct connect or indirect connect, connect means has been in the same movie)

- 给一个array of self defined integer。类似这样：MyInteger{int value, MyInteger connectedInteger}, connectedInteger的index只会在当前MyInteger之后。然后给一个array of index 要remove 并且 array也要被update（就是被remove掉的不能为null）。没什么好说的，就正常做了。

- lc 270 然后探讨各种办法实现（说只是探讨,不用写码，然后方法as many as possible). closest k elements of given number in binary search tree. 分析出三四种办法，分析复杂度

- 一道关于 package dependency. 决定应该先build 什么package，再build什么package。topological sort 就好

- [3,3,3,2,2,1] -> [3,3,2,2,1,1] (3个3 2个2 1个1) 设计接口getNext()

- 利口原题数字不记得了，给一堆数，看是不是能分成多个 数字连续的且长度起码为k的序列

- 二维矩阵找median，sorted in row

- 加油站问题，给你起始点A 终点B 一个数组代表每一个加油站的油价 坐标代表每一个加油站的位置 一辆车 每英里耗油V 从A出发到B最少耗油量

- 给你两个数组分别是[a b c] [c b a] a 从 index 0 移动到了 index 2... 由此可以得出第一个数组移动到第二个数组有一个 map 现在给了你第三个数组 根据之前的map 生成第四个数组

- 给你一个broken binary tree 有一条边导致这个tree不valid 让你找出来并删掉

- 给你一个固定的正则表达式pattern a\*b 再给你一个字符串让你判断他俩是不是match

- lc 505

- 给一个list 有城市a到城市b的 机票价格 比如 a b 100\$ b c 200\$ e f 200\$ ... 现在让你从城市x 到城市y 不能超过两次中转 花费的最少飞机票钱 follow up:是 打印结果.

- 设计unique ID系统

- 两个人玩游戏，一大堆数，游戏双方知道这些数是什么。一个人心里想一个数，另外一个人去猜对方想的数是什么。猜的这个人怎样才能给出第一个他猜的数使得他后面猜中的机会更大。这个题费了半天劲才明白对方到底是要我干嘛。开始想到概率论的东西。面试官同意了我的想法。follow up: 是那第二个数要如何猜才能更快的猜到对方想的是什么。然后就整个懵逼了。

- 给一堆树的node，判断这些node能不能组成一个树

- dp题。给一个grid的宽和长，求得从左下的点到右下的点所有可能的路径之和。

起点当然是左下，要求每次只能走三个方向， $\rightarrow \nearrow \searrow$

follow up: 2d dp -> 1d dp

- > 假设我们有一个数组「0,0,0,0,3,3,3,2,2,5....」设计一个class，封装成「{4, 0}, {3, 3}, {2, 2}, {1, 5}」

basically把所用重复个数和value本身封装在一起。第一问写一个function input是一个idx，找出相对应的value，比如5->3 / 8->2 第一问我就写了一个O(n)遍历 follow up: 如何封装一个extra参数，优化时间。{4, 0, x}, {3, 3, x}, {2, 2, x}, {1, 5, x} 答: binary search

- 给一个公式 $2^i * 5^j$  给一个上限N，求得所有可能的数( $i \geq 0, j \geq 0$ ) 并且increasing order输出 后来发现就是ugly number ii。我花了45分钟勉强写一个O(nlogn)，面试官心里的答案是O(n)

- 特友好的国人妹子（据说刚来一个月），先扯了几句闲话，开始做题。第一题：给一个array of binary tree node，问能否组成一个树（必须并且只能用上所有给的node，只能组出一棵树）。用hashmap存每个node的parent，然后找那唯一的一个没有parent的node为root，BFS检测环路。

- 剩15分钟第二题：定义一个横向一维棋盘，上面有两种棋子L和R若干，L只能往左走，R只能往右走，并且棋子不可以跨越其他棋子移动。给出起始状态和终点状态，问能否到达。序列对比+起始-终点状态检测秒了

- LC340。follow up: 如果string太长无法存进内存，只能用getNext()获取下一个字符，该怎么修改算法。

- 给一个整数array代表一组线段，问是否任意三条线段都能组成三角形，数学题O(n)秒了。follow up: 如果是sorted big array 并且用getNext()获取下个线段长度，问是否存在三条线段能组成三角形（这不还是数学题吗）

- 典3Sum返回T/F，不允许用hashing（包括hashset，hashmap和类似手段）以小于O(n<sup>2</sup>logn)的时间复杂度完成。这题在binary search上纠结了好长时间，最后用循环+双指针搞定。

- 给一系列近义词组(a set of pairs of strings)，问两个string是否同义（对应词为近义词或完全相同）。一开始用2d boolean array来存储近义词关系，最后改成map(word,set(word))来减小空间需求。关键点在于(a,b)近义词，(a,c)近义词，(b,c)不一定是近义词，除非(b,c)出现在set里面。follow up: 如果(a,b)近义词，(a,c)近义词，(b,c)就是近义词（不需要存在于set中），怎么修改算法。类似MST的思路搞定。