

### P003 相隔至少 k

(面经中看到多次) 给一个 string 比如 aaabb .

重新排列这个 string 的 character 让相同的 character 不相邻

follow up 相同的 character 相距至少为 k (已解决, 用 heap+queue)

a. 另一种问法: 给一个 String[] array, 和任意一个移动的 window size k .

对 array 里的元素位置进行改变, 使得 window 里的元素不重复. 要 efficient 的解法。

相距定义为  $j - i \geq k$

### P004 Top K Voted 个人

给一堆 votes(candidate, timestamp) . 问当前时刻 T 得票最高的人是谁。Follow up 问得票最高的前 K 个人

(Q,T) represents at time T, a query Q comes. Each query may appear many times.  
(q1, t1), (q2,t2), (q1,t3), (q3,t4)....

Write a class, which contains a function of enQuery(T,Q) and a function TopK which will return K query, which appear most times within t

(q1,0), (q2,0), (q1,1), (q2,2) . (q3,3) (q1,3)

### P005 平分剩余点

平面上一堆点, 找两个点使由这两个点确定的直线平分剩余所有点

### P006 贪吃蛇

贪食蛇 就是手机里面的那个贪食蛇, 不能出手机屏幕, 头部不能吃到身体的任何部分, 出了尾巴, 因为头部移过去的时候尾巴会移走, 所以不会吃到

输入:

1. List of moves: [U, U, D, L, D, R...] U: up, D: down, L: left, R : right 移动的方向, 有限长度的 list

2. List of food: [(x1, y1), (x2, y2), (x3, y3)....] 食物的位置, 有限长度的 list

3. int: width 宽度 (屏幕宽度)  
4. int: height 高度 (屏幕高度)

开始的时候，蛇在 (0,0) 的位置，长度是一

蛇每次吃到一个食物，长度就会增加一，然后得分也会加一，求最后的得分 食物是一个一个出现的，第一个食物没被吃到的话，第二个食物是不会出现的。

## P007 黑白棋 1

第二轮，白人小哥，在围棋棋盘中给定任意一个白 ( 或黑 ) 子，找出它是否被围死。这个就是个 BFS。写完后他两次说

"I found a bug", 我基本上几秒钟之内找到。

黑白棋。给一个棋盘和一个棋子的坐标，判断这个棋子是不是活的。Leetcode 也有类似的比这个难得题目。DFS/BFS 看能不能走到棋盘的边缘就好。分析复杂度。

## P008 黑白棋 2

给定黑白棋的一个 Scenario，问当前黑方可以在那些位置放棋子。

沿着上下左右两条对角线分别扫描一下。O (棋盘大小)

## P009 凸包

## P010 带权随机数生成

大概就是一个数组

1, 4, 2, 6....

每次调用一个函数，按照数组里面的数字的大小，返回相应的 Index。

比如， 上面的例子就是

1/13 的概率返回 0，

4/13 的概率返回 1。

## P011 Popular Number

3 轮 白人 · popular number 找出数组里面出现次数大于  $n/4$  的数字 · 之前面经见过 解法 binary search

给一个 int array · 找任意一个 popular number,

popular number 就是出现次数大于等于 array length 的  $1/4$  的数。

其实就是 Leetcode 169, 229 Majority Element.

第一问 unsorted array 用的 hashmap。第二问 sorted array 用的 binary search。

Boyer-Moore majority vote algorithm

a.

1 看 `arr[0]` 和 `arr[n/4]` 是否相等 · 相等则返回 `arr[0]` ;

2 如果不等, binary search `arr[n/4]` 第一次出现的位置 · 假设为 `j` · 比较 `arr[j]` 和 `arr[j+n/4]`, 相等则返回 `arr[j]`.

3 如果上一步还好是不返回 · 接着 binary search `arr[n/2]`...

4 类似的接着 binary search `arr[3n/4]`

## P012 URL 中位数

类似 Lint code data stream median, 写个 API · 有两个方法 · `addURL(String url)` 和 `getURL()` ·

`getURL()` 返回的是目前为止所有 URL 长度的中位数。lz 使用两个 heap 做的。

follow-up: what if we know the range of the input ·

比如我们知道 URL 的大小不会超过  $2k$  · 那有没有别的 implement 的方法。

## P013 股票价格改正

一道常出现的面经题 · 拿出来讨论下自己设计接口 · 使得支持两个 function:

```
onUpdate(timestamp, price)
onCorrect(timestamp, price)
```

可以理解为有一个时间流 · 每一个 timestamp 都对应一个股票的时间 ·

每次调用一次 `onUpdate` 的时候 · 就对我们设计的那个类插入一个 timestamp 和 price ·

`onCorrect` 就是修改之前的一个 timestamp 的 price。

最后 · 我们的类要能返回 latest price · max price 和 min price

## P014 Rate Limiter

## P015 分水岭问题

Flow water 问题，没太看懂问题。

具体题目如下：

```
~ ~ ~ ~ ~
~ 1 2 2 3 (5) *
~ 3 2 3 (4) (4) *
~ 2 4 (5) 3 1 *
~ (6) (7) 1 4 5 *
~ (5) 1 1 2 4 *
* * * * *
```

给定一个 grid，判断是否有点即能到达~，也能到达\*。

## P016 排序

有一堆 query<ID, start\_time, end\_time>，按照 start\_time 排序，输出是<id, start/end, time>按照 time 排序，我最开始说全排，面试官说没有那么多的内存，后来我就改进了一下但还是用了 priorityqueue 实现。然后问我如果是 multithread，每个 thread 处理的 query 是不 overlap 的，问时间复杂度是多少

a. follow up 多线程答案： $O(n \log t)$ ，t 是 thread 的个数

b. 如果有 t 个线程，那每个线程的工作量就是  $(n/t)$ ，每个线程单独排序耗时应该是  $(n/t) \log(n/t)$  吧，然后 t-merge 的效率应该是  $n \log t$

## P017 重复子树

给一颗二叉树，返回重复的 subtree。比如：

```
  1
 / \
2   3
/ \ / \
4 2 4
/
4
```

结果应该返回[ ( 2 -> 4), (4) ] 两颗树。

2. output all same sub-tree pairs.

## P018 Fraction

设计一个 fraction number 的 class，要求实现 equals 和 String toDecimal()。toDecimal 是 leetcode 原题，equals 的话注意先得到 gcd，然后除了之后再比较。还有分子分母是否为 0，符号不一样等等都需要考虑。细节挺多的。最后一个，divide float numbers，给定一个除数被除数，以及一个精度，要求在不用除法，mod，以及右移位>>的情况下做一个除法器。不难，参考 leetcode divide two integers. input: float a, float b, float c. output: result 要求：  
 $\text{abs}(\text{result} - a/b) \leq c$  corner case 这里就不写了，基本功，要练

## P019 所有正方形

跟最大正方形和含 1 正方形本质一样

给一个二维 boolean array，true 代表 greyed，要找出所有可能的正方形。比如：

```
0 1 0
0 0 0
1 0 0
```

一共有 8 个正方形（边长为 1 的 7 个，为 2 的 1 个，为 3 的 0 个）。

## P020 拿钱问题

第二题是那个有一列袋子，袋子上标有内含的硬币价值，两个人轮流取，只能取头和尾的袋子，看谁最后能取到的总价值最大，然后程序是计算作为先手玩家最大能取多少。

我以前只写过这题用 backtracking 的办法，给他说了之后他问我 time complexity 是多少，然后问我咋改进，我说用一个 hashtable 存下来所有遇到的情况，每次 backtracking 时去看那个 hashtable 里有没有这个情况，有的话就直接得到剩余最大的价值。他说行，我就开始写了，写完了 backtrack 的部分，我正准备加上 hashtable 时他说可以了，能写到这个也就行了，而且时间也到了。然后拍了照就走了。

## P021 滚球迷宫

第二轮：很 geek 的小哥不知道啥裔 这轮是 4 轮中面的最不好的个人感觉，

就围绕着一道题写了两端代码，面试官上来先问迷宫该咋走能找到最短路径，我说 BFS，

他说好，然后给了个迷宫用小球走，每次确定一个方向之后小球会一直走到碰到障碍物或者边界才会停下，如果停下的地方恰好是出口就算出去了，然后现场画了一个迷宫让我在图上画一条可行的路线，我当时略微懵逼，

居然没找到一条能走通的路，面试官微微一笑指了一条，我赶紧照葫芦画瓢又找了一条，

心想这是算考 brain teaser 么。。。然后面试官问这种情况又应该咋走？

时间复杂度是多少，如何改进，再他提示下我算是想出来了要新建一个和迷宫同样大小的 matrix，

先遍历整个图两遍，把每个空位上下左右的障碍物的位置都存到那个点的 array 里，

然后再用 bfs 就可以直接读到从任意一个点出发小球能停到的位置，这样可以降低时间复杂度。

然后我写了一次遍历这张图要用的 code 和完成预处理之后再 BFS 所用的 code，

然后时间就到了。

下午第一轮。。。一个中国小哥。。感觉不太友善。。。出了道常见的迷宫题，就是小球滚动不会停止 直到遇到了障碍物才能停下来改变方向。。bfs 可解决。。。lz 好久没刷题。。。各种漏条件。。。。没写完。。。。

### Slide Maze

有个迷宫游戏，从网上找了一个类似的，大家可以玩一下：

<http://www.bigfishgames.com/online-games/4038/slide-maze/index.html>。注意跟一般迷宫游戏不同的是，小人只有到墙或者 boundary 才停下来，我们无法控制小人走的步数。

写一个 function，输入是这个迷宫，小人的起始位置，和他的目的地。要求输出一个 move 的序列，使小人能从起点到终点。

以链接里那个游戏的 level one 为例，输出

```
('Down','Right','Down','Right','Down','Left','Down','Left','Up','Left')。
```

## P022 颜色转换

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

第四轮一个白人小哥，问的是如何把一个 24 位的 rgb 压缩到 12 位，然后保证最小的 vector distance，实话实说不是难题

## P023 炸弹人 or 看花 (超高频)

LC 361

Wall flower 问题 (bomb enemy 问题), x 是花, y 是墙, 在哪个位置能看到最多的花 (只能横竖, 不能对角)

```
X 0 0 0
X 0 Y X
0 X 0 0
```

思路: 存个 int[][], 每个点从左到右扫一遍, 再从右到左扫一遍, 从上到下扫一遍, 再从下到上扫一遍

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瓶看, 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。

找一个点能看到最多的花

1. 面经上的 flower 题, 但没细看, matrix 中有 flower, statue, 还有 empty 的地方。statue 能挡住视野, 问站在哪个空的地方, 能看到最多的花 (只可以看上下左右, 四个方向)

直接说了 dp 的解法, 为了 coding 方便, 我有了两个 dp 矩阵, 后来 follow up 是一个 dp 矩阵怎么做, 我说可以 dp 里面存 vector, 再问只存一个数怎么做, 我表示可以存 count, 面试官就说可以了。

## P024 Sliding Window Average

给一个 API ·  $O(1)$  时间计算 slidingwindow avg, global avg, update(insert) value;

LC 346

## P025 最大假期问题

Given a map, each road has a value denoting how many hours it takes to travel from adjacent cities.

Each city has its own holiday. If you arrive at the city during its holiday, you can get one gift.

However, in each week, you can only travel one time. Now you are initially placed in one city at the beginning of this year, how do you plan your traveling to get the maximal gifts.

There are many tourist sites and each has their own holiday. If you arrive there during the holiday, you can gain one gift. It costs you many hours  $W_{ij}$  traveling from site<sub>i</sub> to site<sub>j</sub>.

What's more, you can only travel once in one week. Now you are initially placed in one site, how do you plan your routine in this year to gain most gifts.

公司有很多 office，你可以每周末飞到别的 office。只有在比某个数字小的 hours 内能飞到才可以飞。每个 office 有很多假期。怎样获得一年内最多的假期。这道题应该怎么解？跪求大神指点啊

第二题，DP，有几座城市，每个月在每个城市都有不同的假期，然后每个城市有飞往不同城市的航班，求最大能获取的假期和 Path。dp(i)(j) 代表第 i 个月在第 j 个城市所能获得的最大假期。DP 方程大概是  $dp(i)(j) = \max(dp(i-1)(fromCity) + map(i)(j), dp(i)(j))$

这个题经常在面经中出现，但是到现在还不知道具体的输入时什么？有碰到过的同学可以讲解下吗？

## P026A Reconstruct Queue

restruct queue 题目意思大概是这样。一群人在一点商店的门口排好队站着，每个人的身高都不同，到中午了，大家都去吃饭，吃完饭之后回来，要求按照之前的顺序排好队，给的信息是每个人的身高和原来队伍中在其前面有多少个比他高的人。

## P026B Reconstruct Array



给一个 array 比如  $[4, 2, 1, 3, 5]$  , 根据这个 array 现在我们能有了一个新的 array => 每个数是在原 array 里, 在它左边的所有比它大的 number 的个数, 就是  $[0, 1, 2, 1, 0]$  . 题目是现在给了这个  $[0, 1, 2, 1, 0]$  要求原 array, 原来 array 的 range 是  $1 \sim n$

## P027 删除最少变回文

第二题就是给你一个 string, 你返回 remove 最少数量的 character 之后形成的 palindrome  
Follow up Remove Replace Insert

## P028 Longest Consecutive

给一个双向链表和一个存着部分节点的数组, 问这个数组里面的节点能划分成几个 group。

## P029 完成任务

题目是给一堆 people, 每人有 skills, 再给一堆 tasks, 每个有需要的 skills, 返回 bool 值是否所有任务都可完成。然后一通讨论 corner case。follow up 问 people 之间不能合作怎么做。

补充: 第一问的话没有限制, 可以多个人合作完成一个任务, 也可以一个人参与完成多个任务

第二问的话一个任务只能由一个人完成, 但一个人可以完成多个任务

## P030 最大面积子矩阵 under budget

一个 matrix, 每个  $matrix[i][j]$  有一个价格给你一个 budget 问如何求出最大面积的子矩阵

## P031 删掉三个点

平面上有  $n$  个点,  $n \geq 5$ 。每个点有  $x$  和  $y$  两个坐标,  $x$ 、 $y$  均为正整数。 $n$  个点中可以拿掉任意 1 到 3 个点。求包含所有点的矩形的最小面积。(点可以压在线上)。

## P032 可乐机问题 (题意不明)

## P033 穿墙最短路

穿墙问题 给一个  $m \times n$  的 board, board 里面存 0, 1, 和 2, 分别代表: 0 - 可申通无阻 1 - 有墙阻隔, 人只能穿墙才能经过 2 - 有建筑, 人无法经过

给定一个点的位置和他最多能穿的墙的数量, 求到他另一个指定的点的最短路径的长度。如果路径不存在, 返回-1

给任意一个矩阵, 0 表示可以走, 1 表示是墙。给个起点和终点, 找到最短路径, 返回路径和长度。你可以走上下左右, 你有一个特殊技能, 可以跳墙, 但是只能跳限定次数, 因为会累啊 (他就是这么说的) ... 每次只能跳一堵墙, 连续的墙不能跳。跳一次墙算走了一步。

## P034A Swap 变某树

第一题: 问一棵二叉树能不能通过 swap 左右子树变成另一颗树, 很简单, 直接递归解, 然后跑了跑 case, 问了复杂度。

## P034B Swap 变 BST

第二题, 问一二叉棵树通过 swap 操作能不能变为二分查找树, 这个代码比较长, 也是递归解。

swap 操作是交换左右子树的, 然后子树也可以进行 swap。

这跟 leetcode 有道题, verify BST 有点像,

可是每步 verify 的时候, 你还要考虑 swap 的可能性。

## P035 X of A Kind

Q1: Assume you have a deck of cards. Each card has a number on it with no suit. We define "X of a kind" as X cards with same number on it ( $X \geq 2$ ). Determine if the deck can be fully divided into sets of "X of a kind".

Example: 3, 5, 3, 5, 3 -> True

3, 3, 5, 3, 3 -> False

## P036 Straight

Define "Straight" as 5 cards with consecutive numbers. Determine if the deck can be fully divided into sets of "Straight". Example: 1, 2, 3, 4, 4, 5, 5, 6, 7, 8 -> True

### P037 X-Straight

Define "X-Straight" as X cards with consecutive numbers ( $X \geq 3$ ).  
Determine if the deck can be fully divided into sets of "X-Straight".  
\*/  
// assuming number on the card is 1 to 13

### P038 Partition Sum Problem (NP)

Partition a set of numbers into two such that difference between their sum is minimum, and both sets have equal number of elements.

For example: [1, 4, 9, 16] is partitioned as [1, 16] and [4, 9] with diff:  $17 - 13 = 4$ .

### P039 用户排名

题目是给你一个 board。

里面存储 user 的信息，user 有 id 和 score。

board 有 adduser(id, score) (返回 add 进去的 user 当前的 rank), findByRank(k) (这个返回 id)。

Add 如果本身已经有 id 在 board 中，需要对这个 id 的 score 进行 update。

adduser 操作 和 findByRank 操作都要 lgn

用 binary search tree 即可，每个 node append 点的个数 就行

Round3:

一堆 score 和 player，getPlayerByRank。BIT 加 binary search

### P040 自建 BST & BBST

Search Successor Deletion Insertion

至少学会一种 BBST

### P041 Maximum Guest In a Room

Input: arr1[] = {1, 2, 9, 5, 5}  
exit[] = {4, 5, 12, 9, 12}  
First guest in array arrives at 1 and leaves at 4,

second guest arrives at 2 and leaves at 5, and so on.  
Output: 5  
There are maximum 3 guests at time 5.

## P042 少了两个数

大概就是，一个数组，1112223334445556677888....当中少2个数字，找到就行了。

a. `binarySearch`(只要重复数字是连续的即可),

## P043 AA 制问题

AA 制问题：有一题是比如有一群朋友一起出去玩，然后在外面的消费可能有些人先垫付了，

比如 ABCDE 五个人出去聚会，总共消费 150，然后 A 垫付了 100，B 垫付了 50，然后大家回来之后要 AA 制结清花费，

问使得所有人平摊总消费的最少 transaction 次数，比如 C 回来之后给 A 30 块就算一次 transaction。

a. 通过 2sum, 3sum...nsum 一路算下来是面试官最推荐的解法 (不现实)

b. BFS 更好

c. 最多  $n-1$  次 transaction

## P044 安排任务

“还有一题是给你一些 task，然后这些 task 之间是有规定的前后执行顺序的，比如必须先执行 A 才允许执行 B，然后同时给你 5 个 processor，如何最有效率完成所有任务”

a. 好像是求连通图问题

b. 另一道相似问题的答案：最后姐姐说可以对每组边算个 hashcode，然后像 bucket sort 那样分给一个相应的 machine 处理

## P045 二方格取数

二维 board 里面有草莓个数 从左上到右下再回来 求草莓采摘最大个数

## P046 最小公倍数 LCM

求  $1, 2, \dots, n$  的最小公倍数。

$f(5) = 60$

## P047 杯子装水 (Twitter 题不是 GG)

给两个大小分别为  $m$  和  $n$  的杯子，以及一个目标容量  $k$ ，判断是否能够装出目标容量的水，就用 BFS 就可以了（也有数学的解法，但是面试官说不用数学的解法做），followup 是求怎么装（这也就是为什么面试官不要用数学方法做的原因）。之后又讨论了下 DFS 和 BFS 的优劣。

## P048 安卓解锁 Android Patterns (高频题)

第四个面试：问 android 的面板 pattern 和普通数字密码，哪个更 secure？然后问了 leetcode 的 351，改了一下，最后 return value 是 list of all possible patterns。

Android 手机手势解锁的所有可能性（至少连接 4 个点，至多全部 9 个），需要考虑三点一线 invalid 的情况，

如果考虑成 1-9 九宫格的话就是从 1 连到 3 这样的路径是不允许的

## P049 Max Multiple (超高频)

给一个整数数组（可正可负可重复），返回其中任意两个元素乘积的最大值。

follow up 任意  $k$  个元素最大值

## P050 电话号码问题 (超高频)

要求实现一个北美电话号码的分配系统，即 10 位的电话号码的注册、注销和查询。

```
/*
```

10 位一共有  $10^{10}$  种不同的电话号码，考虑到实际 0 不会做首位实际只有  $10^9$  个不同电话号码

```
trie can be used to solve this problem
```

```
注册一个电话号码 area -> second ->
```

```
*/
```

有如下的 API：

```
void registerNumber(String number); // 注册一个电话号码
```

```
void unregisterNumber(String number); // 注销一个电话号码
boolean isNumberRegistered(String number); // 查询一个电话号码的注册状态
String getAvailableNumber(); // 返回任意一个可用号码
```

假设有个电话公司，用户要新号码，给三个 function，一个是把用户申请的号码标记为 taken，第二个是 check 这个号码是不是已经被占用，第三个是随机生成一个号码。问你不能用不同的 data structure 设计这三个 function。然后问了每个的 runtime。也问了搜索和生成 random number 的时候有没有 fast strategy。

上午第一轮是个亚裔哥哥，第一题出的是 phone number 的 register/unregister 的 data structure。。。谷歌常考题。。

## P051 类似 24 点

coding 一道 类似 24 点。

题目是给一组 integer 问可否用加减乘除使结果等于某个 target。比如 { 1, 50, 3, 6, 7 }，

target = 60

结果是 true 50+6/3+1+7=60

b. 允许加括号的话解题思路是：permutation+add parentheses

c. 不允许加括号的话应该就是 permutation+express add operations

## P052 Word Alert

给一个 stream of characters(ascii) 和一个 hot word("a-z")，

求哪一个 character 可以和前面的连续字符组成这个 hot word，

给出一个 char，返回 true or false。

要求 O(1) 空间

follow up

是如果有多个 hot word · 如果能组成任意一个就返回 true

不要求  $O(1)$  空间

```
Example:
word = "try"
stream = "abcokdeftrying....."
output = "0000000001000....."
```

## P053 只有一个字母不同的 substring pair 数

53. 题目是给一个 string, 找出符合条件的 substring pair 个数, 条件是两个 substring 只有一个字母不同。

LZ 用 dp 做的,  $A[i]$  表示用 index 为 0 到 i 的字符能找到的 pair 个数,  $A[i] = A[i-1] +$  第二个字符串以 i 结尾以 j 开头,

第一个字符串以 k 结尾 ( $0 < j < i, i-j \leq k < i$ ) 可能的匹配个数。

算法时间复杂度是  $O(N^4)$ , 复杂度还是很高, 哥哥想 follow up 的,

结果没有时间了

lz: 允许 overlap, 一开始 LZ 写的是没 overlap 的, 后来被中国哥哥提醒了下

ask: “两个 substring 只有一个字母不同” 指的是 一定有一个字母不同 还是 最多有一个字母不同?

lz: abcde 和 abdde 这样, 有且只有一个字母不同

ask: 就是说长度也一致? 然后每个 index 的 character 都一样, 除了一个不一样, 对吧?

lz: 是的, 字数字数

## P054 字符串压缩 (超高频)

字符串压缩问题: 经典的地里出现过的 String 压缩编码解码类似题,

后悔当时看到没有好好写过一遍. 给一个 String 比如 "abcdffffffffffxyz",

写两个 methods, encode 和 decode. encode 就是比如 "ffffff" 变成 "7xf",

decode 就是要变为原字符串. 我说 "ff" 怎么办, 他说变成 "2xf" 你不觉得更长了吗?

我才明白了, 应该是 encoded 后的 String 要比原来的短

写完以后他就问我如果原 String 本来就是 "5xt" 这种结构, 我 encode 应该怎么处理?

"1x51xx1xq" 就好了.. 还讨论了好多种情况, 最后一种是 "1aaaaa" 这种情况怎么变,

我说 "1x15xa". 他说这是 6 个字符, 能不能只用 5 个? 实在想不出来, 这时候第三个小哥进来了,

韩国哥哥就过来告诉我说,其实看做 1a 和 aaaa 两部分 encode 就好....

a. 特殊处理数字和 x, 其他单独的字符就不压缩了

第二题感谢地里面经, 是 string encoding, 就是 aaabbb 变 3xa3xb 那一道题,

写了 encode 的代码, follow u 带着一步一步发觉有哪些问题,

比如原始 string 就是 3xa 怎么办, 回答是把数字和 x 都编码。

再问怎么简化一点, 答是只用 encode 数字不用管 x。

再问如果有重复数字比如 2222a 呢, 答压缩数字 4x2a。

再问如果有一百万个数字是 0123456789012345.....的循环, 最后跟一个 x 怎么办,

答如果 x 结尾不用管, 不压缩了就。整个过程一半自己想出来一半靠他提示,

他其实一直再问, "so what's the general rules?"。。。。

就是说分情况讨论之前提到的几种情况, 保证能够顺利 decode。

并且不要求每次 encode 之后 string 都变短, 但是总体上 average 是变短的。

这轮代码写得还行, 就是最后 general rule 答得不好, 感觉一般。

而且没留时间问问题。。

之后的题是 implement an encode function。decode 的 func 已经给了, input 是 string, output 也是 string。decode 是如果看见一个数字然后又看见 'x', 那么在 output 里加上这个数字数量 'x' 后面的 char。例子, input = "4xe", output = "eeee"; input = "abc10xq", output = "abcqqqqqqqqqq"。encode func 应该尽量减少 output 的长度, 然后保证 decode 出来的 string 是 encode 之前的 string。最后的 encode func 应该保证不会有任何不同的 input 最后 encode 完有一样的 output。

第二题感谢地里面经, 是 string encoding, 就是 aaabbb 变 3xa3xb 那一道题, 写了 encode 的代码, follow u 带着一步一步发觉有哪些问题, 比如原始 string 就是 3xa 怎么办, 回答是把数字和 x 都编码。

再问怎么简化一点, 答是只用 encode 数字不用管 x。再问如果有重复数字比如 2222a 呢, 答压缩数字

4x2a。再问如果有一百万个数字是 0123456789012345.....的循环, 最后跟一个 x 怎么办, 答如果 x 结尾

不用管, 不压缩了就。整个过程一半自己想出来一半靠他提示, 他其实一直再问, "so what's the

general rules?"。。。就是说分情况讨论之前提到的几种情况, 保证能够顺利 decode。并且不要求

每次 encode 之后 string 都变短, 但是总体上 average 是变短的。这轮代码写得还行, 就是最后

general rule 答得不好, 感觉一般。而且没留时间问问题。。



## P055 广告盈利

华人小哥。

题目：

Given an array of  $Ad$  (profit, start\_time, end\_time)  
and a timespan  $[0, T]$ , find the max profit of Ads that fit the timespan.

小哥进来就说了句中文：怎么样啊？

顿觉好温暖，中文寒暄了几句就用英文开始了。

先说了穷举法  $O(2^n)$ ，然后说了贪心法（不是最优解），最后用 DP 解决。

小哥态度灰常好，给了很有用的提示，就像自家人啊。

## P056 Initialize Candy Crush Saga

design an algorithm to initialize the board of Candy Crush Saga. With  $M \times N$  board,  $Q$  types of candies. Rules: no 3 for run after initialization, must contain at least one valid move at the beginning

## P057 Matching

题大概是这样的，首先，一个二叉树，每条 edge 有 weight（想象成 graph），matching 代表了一组两两没有公共节点的 edges。

现在让找一个 matching，满足两个条件，一是 cover 所有节点，二是 edge 之和最小。

这个题目我怎么感觉如果有匹配只能有一个唯一匹配呢.....

## P058 Shrink Matrix Minimum

$N \times N$  的矩阵，给一个值  $M$ ，求一个新的矩阵，新矩阵的每一个元素是原来矩阵  $M \times M$  中的最小值

## P059 Where Rain Goes

一个  $n \times m$  方格图，有一个位置被下雨了，问最终能从哪些边界处的格子流出水

a. dfs 吧，从数字大的到小或等的，如果是边界就加到 return 里面

## P060 Min Multiple with Digit One

给一个正整数  $n$ ，求能被  $n$  整除的且只由数字 1 组成的正整数的最小值，如果找不到返回 0，否则返回这个数的长度

## P061 TypeAhead

小哥在 Google 8 年了，太资深了

题目是，在发邮件的时候，比如输入 `ben`，下边会提示名字 (`FirstName`, `LastName`) 或者邮件以 `ben` 开头的人，设计一个类来完成这个提示功能。假设每次我们返回最多 10 个这样的结果。

Follow up I，如果希望返回的结果是 alphabetic 有序的，比如输入 `ben` 的时候，`benaa` 在 `benbd` 前面，怎么设计。

Follow up II，如果我们希望 FN 是 `ben` 开头的在 LN 是 `ben` 开头的前边，比如 `ben Back` 在 `ben Smith` 前面怎么办。

## P062 Zigzag Iterator with Prev

开始的题目是 LeetCode 的 Zigzag Iterator

比如我们有一个 `Iterator<Iterator<Integer>>`，这个里边是 `iterator`

i1 1, 2, 3

i2 4

i3 5, 6

然后结果返回 1,4,5,2,6,3

Follow up 是，如果这些 `iterator` 都有 `hasPrevious()`，`previous()` 方法，意思就是后退一步，你的 `class` 也应该有这两个方法，来后退一步

比如我们现在结果返回了 1,4

这时候原来的 `iterator` 变成

i1 2, 3

i2

i3 5, 6

如果调用 `previous()`，变成

i1 2, 3

i2 4

i3 5, 6

有一个情况是如果现在结果返回了 1,4,5,2

这时候原来的 iterator 变成

```
i1 2, 3  
i2  
i3 5, 6
```

那么调用 previous 的时候怎么知道调用 i1.previous() 还是 i2.previous()

最后姐姐跟我大概讨论了下 concurrent 怎么办

用个 lock, 或者用 singleton pattern, 对这个 synconize 这个 instance

最后要走的时候, 我问了下姐姐对 previous 的最优解是什么, 当时可能面完太放松了。。。小姐姐说的没记住, 可能的意思是对于每个 iterator, 我们 keep track 什么时候调用了这个的 next, 然后后撤一步。好像不太对。。。sorry 我忘了原话是什么。。。

## P063 Moving Average No Deque

题目大概是, 每次用户会调用一个方法 double next(double v) 然后函数返回的是这个数之前的 windowSize 个数的 average

比如 windowSize 是 3, call 了 next(10) next(11) next(3) call(1), 第一个返回 10, 第二个返回 10.5, 第三个返回 8, 第四个返回 5

因为我用了 Deque 来保存之前的数据, 我以为他会问精度的问题, 我记得面经里有人发过, 结果没问。。。所以 Follow up 是 如果不用现成的 Deque 这个 class, 你怎么办。好像用个链表更好写。。。我作死说可以用一个 rotated array 来模拟这个功能, 其实也挺简单

## P064 判断两棵二叉树先序遍历是否一样

讨论一下, 最后用 BST iterator

## P065 Streaming Quantile

给一个巨大的 data stream，假设都是 integer，内存和硬盘都存不下，问如何在输入完之后，找到特定的 quantile，比如 50% 的，那就是求中位数。允许有一定的误差比如 10%。写一个 API，需要哪些变量和方法，这题的关键是如何不均匀滴分 bucket，使得满足误差要求

## P066 给字符串数组编码

文件读写，输入是一个 word list，讨论了半天，最后明白过来就是 string encoding, decoding. 一分钟写完 encoding.

1. String encode / decode

如何把一个字符串数组编码成字符串，然后反向破解。

正准备写 decoding，面试官说“慢，我们换一个方式”，decoding 改为给一个原来 word list 的 index，要求返回 index 对应的单词。

白人大叔很耐心的提示了半天，最后想到在文件的开始写入每个单词结束的位置，但是位置信息的位数不固定，比较难 decoding.

最后又提示了半天想出来用二进制表示位置信息这样位数就固定了

## P067 车牌号

We look at Polish cars licence plates and try to find a word from the dictionary that includes all the letters from the licence plate. The shorter the word, the better. The licence plates start with two or three letters, then they are followed by 5 characters, from which at most 2 are letters, the rest are digits.

Your goal is to write code that will find the shortest words for 100 licence plates. You are given a dictionary.

E.g. for the licence plate "RT 123SO" the shortest word would be "SORT", for "RC 10014": "CAR".

## P068 覆盖所有 Interval 最少几个点

give a list of intervals, find min number of points which will intersect all intervals

## P069 Most Overlap Interval

Giving lots of intervals  $[a_i, b_i]$ , find the interval (point) which intersects with the most number of intervals

## P070 Interval Search Tree

第三轮 亚裔小哥 做 backend 的 题是 merge interval 的变形，结合他的工作情景，是有一个 date range 的类，要实现 insert a range, delete a range, sort, check if a date is in some range。写完然后问时间空间复杂度，然后他说他实际工作中是用 ArrayList 存，没想到我用 TreeSet 存好像也不错~ 然后就是问问题了

## P071 镜子和光线

给一个二维的棋盘，棋盘上每一格上可能有镜子，方向可能是/或\，从棋盘外射进来一束光线，问射出的方向和射出的格子。

Follow up: 问随机生成棋盘的方法

Follow up: 如果棋盘大怎么办

## P072 Nearest Neighbor 2D Tree

有两个 set 的 points，set A 和 set B，又给了一个 double 类型的参数 delta，求 set A 里面每一个 point 最近的且距离小于 delta 的 set B 的 point，有点绕，就是给一个 point A，就返回一个最近的 point B，并且这个距离要小于 delta。lz 开始用的暴力解法，复杂度是  $O(n^2)$ ，后来想出来一个先排序再搜索，复杂度降为  $O(n \log n)$ ，三姐一直说 interesting，我也不知道是好是坏.....

## P073 血缘关系

给你两个人 A 和 B，自己设计数据结构，判断他们是否有血缘关系。我的答案：一开始看不懂这题，紧张的要死，后来终于想到了，这不是二叉树嘛。设计两个 `TreeNode(int id, TreeNode left, TreeNode right)`，left 和 right 是 root 节点的 parents，判断 A 和 B 是否包含有 id 相同的节点。我给了俩个答案，第一个是递归，第二个用 HashSet。都问了时间复杂度和空间复杂度。由于 HashSet

方法要遍历二叉树，我还介绍了下遍历二叉树的三种方法 (Recursive, Iterative, Morris Traversal)，面试官好像对 Morris 遍历比较感兴趣，让我简单讲了下。

## P074 0-9 概率

给一个 infinite array 只有 0 - 9 设计一个 `def getprobability(n)`: 得到某个数出现的概率。  
我用的 reservoir sampling 做的，然后我补充如果直接 hashing 会 overflow。然后考虑到 multi thread 情况，需要写两个函数，一个专门产生 sampling list  
一个专门计算概率。这里要有做个 checker 看看产生的 list 是不是有效，也就是 0-9 数的概率和要为 1

follow up:

现在你得到概率，你怎么按照概率产生刚刚的数。

两个方法，第一个直接用刚刚的 array random index 取数，但是问题是如果是 multithread 调用这样做有问题。第二个方法：定时产生所有数的 accumulate probability 根据这个 probability array generate number 即可。答完烙印还挺满意，拍了照片。。

## P075 密码箱

一堆密码箱，每个密码都是四位 0-9 之间，算一个暴力破解序列，包含所有可能的四位序列，让这个序列尽量短，给了一个贪心算法，代码写的比较长，而且没 bug free

然后出了道 4 位 passcode 破解最短序列题

密码箱，每个密码都是四位 0-9 之间，算一个暴力破解序列，包含所有可能的四位序列，让这个序列尽量短 没有找出最佳解法

## P076 比赛夺冠概率

就是有一堆 player，每个人 beat 其他人的概率已知。然后已知初始的对阵表，问给定一个 player，问他最后夺冠的概率是多少。

## P077 记录扑克牌序

解释就好：如何用最少的 bit 来记录一副扑克牌的顺序（假设 encoder 和 decoder 都由你来设计）

followup 是如果 space O(1) 怎么做。。Time 是多少

第一题就是 permutation，答案是  $\log_2(54!)$  向上取整。

## P078 文件寻址

2：很大的文本文件，比如 Linux 运行产生的日志文件。要在常数时间找到第 N 个日志，怎么存，怎么找。

如果用到了 hash map，假如文件最小 1b，最大 40000b，平均 1000b，总大小 20T，hash map 要多大？有点蒙逼...讨论了一堆...后来讨论了存 offset 地址，问寻址方式 balabala...具体记不清了反正就是很操作系统的东西

## P079 最大相同子句

给 N 个文本文档，找最大相同子句，以单词为单位，连续空格看作一个空格

Suffix Sort

## P080A Number Of Triangles

给 N 个点，判断能组成几个三角形。在一给定误差范围内，斜率的差小于误差范围看作一条直线不能算三角形。如果把点换成边，判断能组成几个三角形。时间复杂度是多少，怎么优化。（印度大哥还不让用 python build-in 函数，要用啥你得自己写）

## P080B Number Of Triangles In Graph

自定义一个无向的图，然后，求这个图里有多少个三角形。三角形的定义是，三角形 abc 中，(a,b)，(b,c)，(c,a) 都相连。半小时后拿到 onsite...

## P081 Find String in Comments

```
String = new String("\\/*");
```

有一个 file，里面有很多 comments (// or /\*)，然后还有个 input string，如果这个 file 的 comment 里 contains input string 的话 return true...

## P082 Hamming Distance

hamming distance between a and b,  $a, b < 2^{64}$ . 这题很快就做了出来。就是把  $a^b \gg i$  &1 64 次。然后他就说要想办法 speed up, 说给我 64G 的 ram。我想了很久最后说可以搞个  $2^8$  的字典, 然后把  $a^b$  分 8 段比就好。他就说为什么用 8, 然后就问我  $2^8$  的字典要用多少空间。我没记空间大小的那些知识, 所以不会做... 几经提示后结论是可以用  $2^{32}$  的字典要 4G 空间, 这样比两次就好。他最后又问说如果你用这个方法, 但是 ram 只有 2g, 那会发生什么情况。我就说那会有 error 吧。他就说什么 error。我说不出来。他就说 "you clearly have never used win 95 swap space". 然后差不多就结束了。我觉的最后这个哥么应该给我差评了。

## P083 计算二叉树两个节点的最短距离

## P084 雨滴问题

you have 1 meter walkroad, and randomly generate rain, the rain is 1 cm.  
simulate how many rain drops to cover all the 1 meter  $[-0.01 \sim 1]$ .

## P085 Valid Unicode

这次出了一个 byte encoding 的题目, 大概是给你一种编码的格式, 然后给你一个 byte 数组, 让你判断这个数组合不合法 这个编码的定义是, 有 1-7byte 的 character, 每种 character 第一个 byte 有一个固定的前缀, 比如 1 byte character 第一个 byte 的前缀是 0, 2 byte 是 110, 3 byte 是 1110 以此类推, 8byte 是 11111111。有一个特殊的前缀是 10, 这种 byte 是每个多 byte character 的组成部分。比如一个 2 byte character: 110XXXXX 10XXXXXX。一个 3 byte character: 1110xxxx 10xxxxxx 10xxxxxx。如果字符被打段或者超长都要 fail。算法很简单, 实现起来还是有些细节要注意。这题之前面经没看到过, 就发出来给大家留个参考。

## P086 String Distance

第一题: 给两个长度相等的 string, 如果两个 string 对应位置的字母不相同记为一个 distance, 如果现在能够交换一次其中一个 string 中任意位置的两个 char, 返回能够将 distance 缩到最小的两个 char 的 index (如果有多个最优解只返回一个)

## P087 $3n$ 个数围成一个环



第二题：有  $3n$  个数围成一个环，取走其中一个的话会顺带去掉这个数相邻的两个数（这两个不计入总和），剩下的继续围成环，问取走  $n$  个数构成总和的最大值。

## P088 Design Frogger (Glassdoor 上这题很老 2009)

Design an algorithm to play a game of Frogger and then code the solution. The object of the game is to direct a frog to avoid cars while crossing a busy road. You may represent a road lane via an array. Generalize the solution for an N-lane road.

## P089 闹钟

从朋友那里听来的一道设计题，给定一个闹钟的 class，到一定的时间会执行 `dosth()`

```
class alarm {
    const timer t;
    void dosth();
}
```

让你设计一个 manager 的 class，能够实现 add 一个 alarm，并且当时间到 alarm 设定到的时间的时候 trigger 它。问怎么设计。请问大家有啥好的想法？

我想到的是

```
class manager {
    list curlist; //sorted
    timer global_timer;
    void check_set();
    void addalarm();
}
```

addalarm 保证按 sorted 的添加到 curlist 里面，然后 check\_set 可以是个 loop 一直 checktimer，从头开始找所有的到达时间的 alarm trigger 它，但是如果多线程就会遇到需要 lock 的情况，请问大家有没有啥好的想法？

## P090 Producer-Consumer

## P091 设计 blocking queue

设计 blocking queue

## P092 Smashable Word

给一个词典和一个 target word, 问这个 word 是不是 smashable. smashable 的定义是这个词在词典里并且它不停地任意删除一个字母得到的单词也在词典里。用了 DFS 之后提示用 hashtable 优化时间复杂度，然后问如果要提前把这个词典存好之后反复调用怎么办，按照他的提示我想到的方法是按长度 sort 词典里所有单词，按照长度 group 起来，给一个 word 之后就直接找比它长度-1 的那个 group 看有没有哪个 substring 在这个 word 里

然后上题。题目是给一个 String 数组（我不晓得我怎么就跟数组杠上了），求最长的 keyword。

keyword 的定义是，移除一个字符，剩下的字符串依然在数组中，一直移直到最后只剩下一个字符。

拿到题目之后，思考一会之后沟通了下我的思路，因为要求最长，所以我建议先根据长度进行排序（又是排序）。

还有一个提前 break 的情况是如果最短的字符串长度大于 1，那没有 keyword 出现的可能性。

这时候跟面试小哥沟通的结果是这个可以返回""

Round2: 给一个字典，找到最长的如下 sequence: [handy, hand, and, an, a]，每次删一个字母，一定要删到只剩一个，只需要一个最长的。我的做法就是从所有最长 word 的开始，dfs 每次删一个，到底为止，一旦出现了一个，整个函数就可以退出了，没有的话就试下一个长度的。此题就是 DFS 提前退出，follow up 不用写，memorization，说一下可以记录一下哪些单词是到不了的，这样可以提前退出 pruning，楼主脑子抽说是要 memo，但是说的是记下可以到的 word，说反了，此题一旦有可以到达的就该退出了。

## P093 OCR

Q2. given millions of electronic books and each book contains hundreds of pages. Given a function OCR to convert the electronic image of book pages to text file. That is to say, scan the books first to get photo copy and then convert the image to text file. So in the process, there is supposed to be 5% error. How do we tell if two books are identical considering the error? Error can come from several cases like: "word random" -> "word ran dom"; "word random" -> "worded random"; "word random" -> "word " Solution: It's an open ended question, you can propose different approaches.

## P094 Sparse Excel File

Q1. Given a sparse excel file, design a data structure to effectively store all the informations. Sparse means most of the excel are empty. Write 4 functions: void set(int row, int col, int val), int get(int row, int col), vector getRow(int row), vector getCol(int col) getRow and getCol returns the

corresponding row or column that is indexed at the input number which neglect empty cells. Solution: Using this one: `unordered_map> myMap1, myMap2; myMap1: (row -> (col -> val)) ; myMap2: (col -> (row -> val))` Follow up: what if there is a function `eraseRow` or `eraseCol` like in excel you can erase a whole row or column. Hash table solution becomes unefficient because row index and column index changes and this cause erase time complexity to be  $O(n)$  Solution : using `LinkedList`, and I have no time for coding

## P095 设计乌龟

白人·设计乌龟·一开始有两个方法·往前走一步·右转 90 度。然后说设计个方法可以执行一串命令：

e.g., "FRRF"代表走、转、转、走。然后命令要支持括号和数字：e.g., "F97[RF[F]]"代表走一步，然后执行括号里的命令 9\*7 次。

## P096 病人吃药

第一题是写一个 class simulate 一个病人吃药。一个 bottle 里有 half pills 或者 whole pills，随机拿一个，如果拿出来的是 half pill 就吃掉，如果是 whole pill 就弄成两个 half pill 吃一个放回去一个。初始给的 pills 不一定全是 whole 的。一开始用了 array list，跟面试官的思路有点不同，交流了一下改用两个 integer 又写了一遍。第二题是假设一开始给的 pills 是 start state，然后给一个 end state。求一个病人随机吃药，从 start state 吃到 end state 的概率。比如一开始 2 whole, 1 half，求吃成 1 whole, 1 half 的概率。LZ 一开始说 build 一个 search tree 这样直观点，他表示不需要把 tree 专门建起来，LZ 就直接一边搜一边算没有用 extra memory。

## P097 Remove Back Edge

给一个 directed graph 的 start node，这个 graph 里可能有 cycle，

如果 remove 一些 edge 可以使这个 graph 不含有 cycle，

并且从 start node 依然能访问到所有这个 graph 里的 nodes，

这些 edge 就是 back edge。要求打印出所有的 back edge。

写完拍了照又 follow up 了一下，问存不存在这种 graph，你可以 remove 不同的 backedges 使得这个 graph valid。LZ 一开始觉得是不存在的，但是想不出来证明的方法。

后来画了个图发现是存在的，但 LZ 自己没一下子看出来。。被提醒了下。

7. detect cycle in a graph. lots of follow up questions. They do care about space complexity.

## P098 不规则图形

然后第二题是给一些类和 function:

```
class Point {
    int x, y;
}

class Rect {
    Point p1,
    Point p2
}

class Shape {
    int inside(Rect r);
    draw(Rect r);
};
```

shape 是不规则的图形，给了一个 function 判断这个 shape 和 rect 的重叠关系，

如果 rect 完全在 shape 里，return 1.

如果 rect 和 shape 有 overlap，return 0.

如果 rect 完全在 shape 外面，return -1

以上是已有的 function，给一个长方形的 screen，再给一些 shapes，

要求写一个 function 在 screen 上画出这些 shape。

写完后又问了一题，如果有 A 和 B 两个 shape，A 和 B 有 overlap，

用给的 inside function 写一个新的 inside function，

判断一个 rect 和这个 overlap area 的关系，同样返回 -1，0，1

然后问这个关系是什么，LZ 一开始没看出来，

其实就是  $\min(A.\text{inside}(\text{rect}), B.\text{inside}(\text{rect}))$ ;

然后问如果把 min 换成 max 的话这个算的 area 是什么样的。应该是 A 和 B 的 union.

## P099 Number Of Dividers

第二题是给你一个 number, 还有一个数组包含这个 number 所有 prime factors, 问这个 number 的 dividers 的个数

## P100 Music List

music list, if shuffle is possible or not, test cases

要求你写一个函数来判断用户使用的是 random 模式, 还是 shuffle 模式。

你有一个 music 的播放列表, 里面的歌曲 unique, 但是播放列表的长度未知。

这个音乐播放器 APP 有两个模式: random 模式和 shuffle 模式。

random 模式就是每次随机播放列表里的一首歌;

shuffle 模式就是 shuffle 列表里的歌, 然后顺序播放, 放完以后重新 shuffle,

再顺序播放; 现在给你一个播放历史记录,

## P101 法里数列

题目是给你一个 positive 的值  $K$ , 然后按照 fraction 的值的小到大输出所有  $n/d$ , 其中  $1 \leq d \leq k$ ,  $0 \leq n \leq d$ , 还有一个要求输出的 fraction 不能有 duplicate, 比如  $1/2$  和  $2/4$  是 duplicate, 这种情况只要输出  $1/2$ 。

```
1: 0/1 1/1
2: 0/1 1/2 1/1
3: 0/1 1/3 1/2 2/3 1/1
4: 0/1 1/4 1/3 1/2 2/3 3/4 1/1
```

## P102 含有 1 的正方形

给一个  $N \times N$  grid, 上面有一些格子里面标了 1. 问总共多少方块了有 1, 所以方块包括只含 1 的那个, 还有任何含 1 的那个 所以  $2 \times 2$ ,  $3 \times 3$

## P103 灯泡开关问题

给你一堆灯泡。可以 flip 一个范围开着变关, 关变开, 然后问这么干了  $k$  次以后, 随便问你一个灯泡是开着开始关着, 怎么做。

## P104 找出多出的两个字符

Given two strings A and B. We know that B is two chars more than A. Find the two chars in  $O(n)$  time  $O(1)$  space.

## P105 最短曼哈顿距离

第三个是求曼哈顿长度的，给一个 2D 数组，

里面有 1, 2, -1 若干，问所有 1 到 2 的 manhattan distance 最短的是多少。

每个 1 到所有 2 的距离取最小的，然后再在这些值中取最小的

## P106 Response ID Generator

然后他看看还有时间又问了一题，说有几台 server 收到 client 的请求之后发出 response，要实现一个 response ID generator，满足生成的 ID 是 global unique 和 sequential，我当时第一反应就是 spanner，就说用个 GPS 一样的全局 timer 生成 timestamp，配上 server ID，组成一个 64bit 的 ID，然后他就问应该是 server ID 在前还是 timestamp 在前？这样做有什么好处？会有什么问题？然后又让我想另外一种设计思路，他给提示是用一个 server 专门负责生成 ID，我就想到 zookeeper，然后就给他讲讲怎么做，有什么好处，有什么问题。

## P107 Index of First Larger After Self

G 家面经题，返回数组的每个元素是原数组中右边的第一个大于该位置元素的 index。

## P108 Larger Than One Percent

给你一个数组，从这个数组中挑出最少的数字，它们的和要比整个数组的总和的 1% 要大。要求 Linear time 解决

## P109 生成迷宫

<http://www.jamisbuck.org/presentations/rubyconf2011/index.html>

## P110 Union Iterator

题意不明

接着 coding question 是写一个 Union Iterator，parameter 是两个 Iterator，

output 应该是这两个 iterator 中 elements 的 union。

ie:

it1:{1,2,3,4}

it2:{3,4,5}

output should be:{1,2,3,4,5}

Follow up: 如果这两个 iterator 非常大，RAM 存放不下怎么办

B-Tree? Bucket?

## P111 等比数列求和

最后一轮，求  $\sum(n^i)$  就是  $1+n+n^2+n^3+\dots+n^N$

## P112 Top K int in a large stream

This can be done in  $O(n)$

2K windows, Quick Select, throw K away.

有  $N/K$  个 batches，所以是  $O(K * N / K) = O(N)$

## P113 最少指数表达

input: int n

function: 将 n 用 2 的指数表示，使得指数表达式的个数最少

output : int num ( 指数的最少个数 )

e.g: input = 28

$28 = 2^4 + 2^3 + 2^2 \Rightarrow \text{num} = 3$

$28 = 2^5 - 2^2 \Rightarrow \text{num} = 2$

所以 output = 2

## P114 猜单词

第四轮：游戏，猜字。给你一个字典，player 1 脑子里想一个 5 位单词，这个单词一定来源于字典，然后告诉你字典中每两个单词间相同单词的数量，你需要做出最好的猜测，能最有利于接下来的猜测。。举个简单的例子，三个单词 a, b, c。如果告诉你 a 和 b，a 和 c 都差 3，然后 b 和 c 差 2，你最不应该猜 a，因为接下来这属于无效猜测。。。。这道题楼主答得不好，所以也描述不清楚，大家凑活着看吧。。先让写个类似 StrStr 的 function，然后写如何做最优猜测。。

第四面：国人大叔，一直友好和蔼的交谈，整体感觉最放松的一轮。先是一个 warm up 问了一个 guess number 的游戏，用最简单的 binary search。然后 follow up：如果给的是一列单词，A 心目中有一个目标单词，让 B 来猜。每次 B 猜一个单词，A 只会告诉他猜中了几个字母。举个栗子，如果 A 心中的是 APPLE，B 猜 ANGEL，A 告诉 B 猜中了 3 个字母（A, L, E，位置无关）。问 B 如何猜最聪明。再 follow up：如果 A 会作弊故意刁难 B，B 知道 A 会刁难自己，应该如何猜。



## P115 Design Google Calendar

design google calendar . 要求分析如何存 data, 如何 invoke user events, 如何 handle 100000events per second, 然后要写了一部分 thread safe 的 code 实现如何 invoke event.

## P116 Key Value with Time

写一个类, 类似 hashmap, 但是有 (key,value,time) 这三个参数。有 set 方法和 get 方法。

调用方法:

```
set(k1,v1,t1)
set(k2,v2,t2)
set(k1,v2,t2)
```

Get 方法有两个返回的可能 (对应上面的 set 看)

```
get(k1,t3), if( t1<t3<t2) return v1. if(t3>= t2) return v2. key 不同的时候和 t3<
t1 的时候都返回 null
```

## P117 Log System

(1) 第一题是设计, 打印接收 message 的 log system。message 包含 id, time 以及是否结束。打印出来 id, 开始时间, 结束时间。用 priorityqueue 做的。

问一下楼主 message 的输入是什么? 打印出来的东西有什么要求? 谢谢!

没有输入。这个 function 包含一个 listener method 接收 message, 打印 method 被调用打印。

譬如 每分钟打印一次。message 可能有 up to 1min 的 delay。所以打印时要考虑当前时间减去 delay 时间

(2)

1. 一个系统, 记录任务进入时间和完成时间

1) 一个任务完成就马上打印, 并从系统中移除

2) 扩展: 按照任务开始时间排序。一个任务完成以后, 如果系统中开始时间在前面的任务没有完成, 你必须等前面所有完成了再打印并删除。

然后问这个系统设计有什么问题

(3)

implement 一个叫 logger 的 interface, 后来发现地里有原题。

题目是酱的，logger 里有俩个 method，分别叫做 startRequest 和 endRequest，input 都是 request id，和时间，return 是 void。

每一个 request 都有独一无二的 id，并且开始和结束都需要分别调用那两个 method。当一个 request 结束之后，需要打印出来它的 id，开始时间以及结束时间，但是，要按开始时间先后打印。

题目很简单，但是感觉自己没有好好把握这次机会，主要是第二面，基于我的 code 的一些改进，答得不好，整个人反应及其迟钝。

白人小哥，又一个上来就直接问问题，不过小哥一直笑呵呵的，但是我总感觉是戏谑的笑容。。。

问题是，有一个 simple log system，可以给每个 team 服务，每条 log 的形式是<team name，message>。每天结束的时候，每个组会调用 API 返回，只能返回给定数量 message 比如 100 条，而且这是所有 message 中的 100 条，不是某个组的 100 条。而且该 API 每天的返回信息是固定的。让你设计一下怎么返回，让每个组都满意，因为如果某个组返回的信息太少，那个组肯定不满意。所有 message 都存在 disk 上，你能拿到所有 message。

我给了一些比较简单的意见，比如优先级，基于时间，或者基于比例或者用随机数，都不行。因为他说这是个简单的系统，没法学习，如果每个组定优先级，那谁都说他们组的信息更重要；如果说时间，那每个组都在每天的最后写 log。。。再说随机数或者比例，这样的话会变相鼓励某个组多写 log，以提高被选中的数量。简直就有点懵，然后还给了个 case，比如前面三个组都有上百条，但是第四个组只有 1 条，buffer 只有 100，按比例的话，第四个组就没有 log 了。。。我给了另一个方法，是先拿出 50% 的 buffer 让所有组均分，剩下的部分再想其他办法。他说均分的方法可以，然后那会 30 分钟过去了，他说先写点 code 吧，我说好，然后就把这部分 code 写完了，中途有些 counter 小 bug 被发现了汗颜。。。最后我给他的方案是，先拿 50% 保证每个组都分得有，而且需求比较小的组可能还会有剩，剩余的回收过来，然后不停地迭代平均分配，直到所有的都分完。感觉有一个小亮点就是，本来需要写两个 function，我说可以改传递参数，直接重用，讲了下思路，就结束了。又是一个没有提问的一轮。。。。

## P118 Design Battleship Game

设计 battleship 这个游戏，似乎是个经典游戏，可我完全没玩过。。

最后跪了，HR 告诉我是第一轮 battleship 这题 feed back 不好，其他轮都很好，第一轮面试官说我没有选择最佳的数据结构。

## P119 Rearrange Order

然后是一道关于按照固定的顺序重新排序，多少次排序才能回到初始顺序的题目，我们总是把第  $x$  个位置的元素移到第  $y$  个位置，比如【0, 1, 2, 3】就是完全不换顺序，一次排序就回去了；【3, 2, 1, 0】是位置 3 和位置 0 互换，位置 2 和位置 1 互换，两次排序就能回到原顺序；而【1, 2, 3, 0】就是每次把一个位置上的元素往后移动一个位置，那显然转一圈就回来了。

举个例子，原来的 input 为 [20, 34, 7, 9]，排序要求为 [1, 3, 2, 0]

一次排序后：

[20, 34, 7, 9] -> [34, 9, 7, 20]

二次排序后：

[34, 9, 7, 20] -> [9, 20, 7, 34]

三次排序后：

[9, 20, 7, 34] -> [20, 34, 7, 9] （此时和初始顺序一致，所以是用了三个 iteration 恢复原样）

排序规则 [1, 3, 2, 0] 的意思是把在位置 1 的元素排到第 0 个位置，位置 3 的元素排到第 2 个位置，位置 2 的元素还是在位置 2，位置 0 的元素放到位置 3 去。排序规则不变，每次只是 Input 被按照规则改变了顺序。

白人大哥，声音有点沙哑，一会严肃一会微笑的，又是上来就 coding!!! google 家好直接。。。

这题很常规，也是面得最顺的一轮。题目是，设计一个 class 来 re-arrange array。一开始会给两个 array，这两个 array 存在一个 pattern，然后给你一个 input array，让你根据 pattern 给 output，例如：A:1,2,3,4 --> B: 4,3,2,1 这样的 pattern，如果在给你一组数是 5678，那你就返回 8765。相当于对应 index。

一开始有点懵，就说，对每个 A 中的元素，在 B 中 scan 来找他对应的 index，这样建立一个 index mapping，时间复杂度是  $O(N^2)$ ，想了两分钟，脑袋简直停滞了。。他说我们先写个 code，然后就把这个 code 写了。写完了以后 follow up，他说你的 code 如果有 duplicate 就不 work 了呀，我说那我加一个 flag array，每次选了一个位置之后把对应位置置为 1。再 follow up 他举了个例子，其实就是 A 和 B 中的元素不完全对应，意思就是 check validation，然后我给写了。这时候如果一个 client 他需要用多个 pattern 你的 code 会是怎样的，就是声明一个新 obj 然后初始化，再调用就行了。这时候开始要优化了，毕竟  $O(N^2)$  太慢了，突然才发现自己傻逼，其实就是一个 double mapping，先用一个 map 记录 A 中每个数字的索引，然后第二个 map 建立 A 的索引指向 B 的索引就行了，对于

duplicate element，我们只需要把第二个 map 的 value 设成一个 list，每次取其中一个就行了，他假设相同元素无所谓先后顺序。顺利写完 code，基本上没什么 bug

## P120 Implement Caesar Cypher

很容易，面试官像个教授，讲了很多密码学历史（我并没有相关背景）不像在面试，像听 lecture...

## P121 Matrix Class (LC308 again.....) (超高频)

第三轮：经典题，二维 matrix，给坐标写 getSum 和 update function。follow up: (1) getSum 怎么  $O(1)$ ，update  $O(n^2)$  (2) 怎么两个都是  $O(n)$

```
write a matrix class
set(int i, int j, int value);
sum(int i, int j); //可求(0,0)到(i,j)整个 square matrix elements 的和
Solution
```

1. 2D BIT 解决之

2. 简单点儿也可以，只需要进行 col dp，每次更新复杂度  $O(col)$ ，求 sum  $O(row)$ ，应付一个店面我觉得也不错

(2) 设计一个数据结构表示 matrix 可以无限 scale 的那种，然后有 sum, set 等 methods 吧，我是想到 hashmap 做的

1. matrix sum

给一个矩阵，一个 update 方法，更新给定坐标值；一个 sum 方法，输入两个坐标，求以这两个坐标为左上右下的子矩阵的 sum（早上说 binary index tree 不会真就考了）

## P122 2sum BST

导致第二题没写完，只写了个 outline。好在面试官表示知道我的想法了.....

（这题谷歌居然也要问）

1 在 BST 中给定  $[min, max]$  求在此值域里的所有 node 之和。简单递归

## P123 最少操作 01 排序

第一轮：面试官迟到了。。给一个 string，只含有 a 和 b，a 可以变成 b，b 可以变成 a，也可以不操作，返回操作次数最少就可以得到的 sort 的 string，用了 word ladder 的方法。。也不知道对不对了。心理素质太差了。。。

Solution

过一遍看左右数量，或者过两遍

复杂度最优可以  $O(n)$  时间， $O(1)$  空间

(word ladder.....难怪楼主挂了.....)

## P124 Sad Face

第三轮：给一个 API，叫 `getchar1024()`，总能返回 1024 个 char，里面有个：(，这个 sad face 前面都是 valid，后面都不要，实现 `getchar(int n)`

## P125 同颜色边长

格子里，给一个入口点。找和他颜色相同的所以格子组成区域的边长。和 boarder 连接的那条边不算。

## P126 Decode String

A

第四轮是中饭之后，白人大哥，Ph.D 做过九年 faculty 的，decode string，输入

`a[4]bcd[3]efg`，要返回 `aaaabcbdddefg`。

string character 都是小写 26 个字母。很快写完。他检查是对的，看着很满意的样子。下面问 encoder 那边怎样实现，

只说 idea 不要写 code。我说了一下他也说好，显得很满意。还剩几分钟聊天。我让他介绍自己，才知道原来是 UIUC 的校友。

B

然后是那道以前出现过的 `3[abc2[x]]` 解码变成 `abcxxabcxxabcxx`。说明了原字符串不会有数字出现

韩国人，给定这类型的字符串：

`"3a2[mtv]ac"`，decompress to: `aaamtvmtvac`，括号可以嵌套。

这个我觉得不是很难，大概花了 15 分钟理清了思路并写好了代码，大概就是找匹配括号递归解，面试官也找不到 bug 表示认同。

但吊诡的地方来了，面试官说把这种字符串 compress 回去...这显然有多种情况，于是我问是不是要求压缩后最短，面试官说肯定越短越好。

比如对于 aaaa，肯定 4a 比 2[aa] 好。

我思考了一会，只想到了枚举所有 substring 及其连续出现次数，然后选择 match 出现次数最多的 substring 作为压缩。

面试官觉得复杂度高了，问还能不能优化，感觉他自己语气也不是很肯定，提示了一下我有没有类似 two pointer 的解法。

我个人觉得这道题真心不简单，没什么想法，一直卡到了这轮结束...

第四轮是中饭之后，白人大哥，Ph.D 做过九年 faculty 的，decode string，输入

a[4]bcd[3]efg，要返回 aaaabccddefg。

string character 都是小写 26 个字母。很快写完。他检查是对的，看着很满意的样子。下面问 encoder 那边怎样实现，

只说 idea 不要写 code。我说了一下他也说好，显得很满意。还剩几分钟聊天。我让他介绍自己，才知道原来是 UIUC 的校友。

## P127 森林照相 (几何题)

什么都不给。告诉你站在一个森林里。你有个任意形状的地图 (他也没给任何定义)，树木的输入也没有任何定义，全部自己设计。

地图里面有所有树木的位置。你有个照相机，找到你能照进照片的树的最多数量。

我就假设照相机有个最大角度。这个可以自己 design。这个哥们感觉蛮严肃加任性，你问他啥他都反问你你说你觉得呢？。。。汗滴滴。。

## P128 Cosine Similarity

第一轮: machine learning design: 有 1 million 个 word embedding。找出一个 permutation，使得相邻的 embedding cosine similarity 总和最大。

## P129 Find ngram skip-gram

找出句子的 ngram, skip-gram. 要求空间最小。

## P130 QuadTree 黑白图片问题

### Quadtree Intersection

二叉树压缩黑白图片，一个图片递归分成 2x2 四部分，如果一个区域颜色一样就设为叶子节点，算黑像素比例

follow up 是给两个图片，把白色视为不透明，黑色视为透明，重叠在一起，返回一个图片，都用二叉树表示这个递归不难，感觉做的不错，最后出了一个小 bug，在他提示下改了

quadtree。大概就是说一个黑白 image，用的是 quadtree 表示的。黑的就是 0，白的就是 1。问求这个 quadtree 表示的 image 里黑像素的比例。

然后 follow up 就是求 combine 两个 quad tree 后产生一个新的 image 的 quadtree，rule 是就是只有两个 images 都是黑 merge 后才是黑，其他都是白。

美国大叔，人很 nice，就是面经里的比较两个 quad tree。然后 return 一个新的，这题虽然简单，但是很容易出 bug。最后面试官引到了一下帮助发现 bug。

## P131 Min Unique Prefix

第一个面试：warmup 问题，给一个 string，找出第一个不重复的 character。没啥难度，直接第二题。给一个 list of strings，比如 ['chips', 'chocolate', 'orange', 'cake'], return ['chi', 'cho', 'o', 'ca']。意思就是找出可以 uniquely 表达每个 string 最短的 prefix。给回的 list 里的 prefix 的顺序不重要。我当时的想法是先 sort，然后值得注意的地方是第一次当两个 string 的第一个 character 不一样的时候。

## P132 Network

第二个面试：有一个 shadow。上来问 network 的问题，accessing a single page, loading forever, what could be the issue and solution?

## P133 Social Network

第二题，假设有一个 social network，找出其中一个人的第 k 层的所有的朋友。例子，A 和 B 是朋友，B 和 C 是朋友，那么 A 的第一层朋友有 B，第二层朋友有 C。直接 bfs 解，node 的 structure 自己写。

## P134 画图题（几何题）

临时换了一个亚洲大哥。哥们出了一个画图题，给一个

```
class {
    draw(boolean),
    move(x,y) // absolute coordinate
}.
```

draw 为 true 的时候是画线，draw 为 false 的时候是抬起笔，move 就是把笔移到某个位置。

然后让画一个  $m \times n$  的矩形，每个格子长度为 L。要求中心对称。

（有俩 bug，一是算 float 的时候两个整数相除，2 是 for loop 时  $\leq$  写成  $<$ ）

follow up：如何一次性画完这个矩形，运行 overlap 某些 edge。

follow up：如何一次性画矩形，不允许 overlap。

follow up：minimize overlap 的 edge，如果证明为啥你的方法是 minimize 的。

（只能达 brute force 所有的可能性，然后算出最小的一个。然后问 brute force 有何优化（答：memorization，貌似不对。））

是画一个矩阵。应该是只有  $m, n$  都为 1 的情况下能一笔画（就是个矩形）。他的第三个 follow up 就是在不能一笔画完的情况下，提供一种方法 minimize overlap 的 edge。

感谢楼主分享。还是没看懂第二题啊。 $m \times n$  的矩形，这里  $m$  和  $n$  是 cell 的个数么？每个 cell 的边长是 L？



## P135 number expressions

第一次电面：假设除了用 1 和 0 表示数字的 binary 方式，还可以用 2 来表示，给一个数字，求有多少种表达方式。第一题的 4 应该是 100,20,12，答案是 3， $8 = 1000, 200, 120, 112$  所以答案是 4。

## P136 顺序输出二次函数

这大叔没看出来是哪的，开始看名字像日本，后来发现是老毛子。虽然题很简单，但是不知道他什么毛病，也不满意的样子。 $f(x) = ax^2 + bx + c$ ，给一个 ordered input，结果也顺序输出。这人上来迟到 7 分钟，导致写的时候时间紧。最后匆匆忙忙写完，也没问啥别的，最后 take picture 走人。

## P137 target 是否在 source 的子树里面

4 轮 不只是哪人，全程无表情。给一个二叉树，一个 source node 和一个 target node，问是否有一个路径从 source node 到 target node（沿着 parent 到 children），LZ 很天真的说用 dfs 做下，写完然后问如何在 O1 time 和 space 找到，LZ 没想出，然后提示是说先 preprocess 下，每个 node 加一个 label 的 range，举个栗子 [1,10] 代表这个 node 下面的节点范围是 1-10 这样通过判断 target node 的 label 是否在范围内就可以确定，感觉像 segment tree，最后没时间写代码了。跪了。。。

Solution

并不是 Segment Tree。。。。

就是给每个 node 一个编号，连续的，按 preorder 的 index 就可以，然后每个子树都是一个 range，这样 query 的时候就很容易判断 target 是否在 source 的子树里

<https://www.clear.rice.edu/comp314/lec/week3/Depth-First.htm>

关键是面试官说 O(1) space 属于胡扯淡....

## P138 Read One

上周 GOOGLE 电面，一个题，已知标准的系统函数 read，有一个功能每次就读一个字节，然后要经常读，在一个 loop 里，问如何优化。结果就是一次调用 read 多读一些，放在缓存里，用没了再读。

## P139 itoa

char\* itoa (int a) , then char\* itoa (int a, int base)

老美白人。注意 corner cases, 字母的范围 (0-9, A-Z) 决定了 base 的范围, 还有 char buffer 的大小。

## P140 match pattern

```
design pattern data struct, and there is a bit flow , check if match any
patter, rules :    0-0, 1-1, x-don't care
e.g,              bit flow  10101010 10101011 ...
                   pattern   1010101x 10101000 ...
```

貌似棒子。为了简单假设 8 个 bits 一组, 自己设计结构, 那么 前两个 8bits match 第一个 pattern, match 哪个 pattern 输出那个 pattern 的序号。然后最开始我不是很明白他要设计成什么样, 他有点不高兴。

我仿照 IP address and mask 用了 array 和 mask array 存 pattern, 然后遍历查找, 要求写代码。后来让改进说效率低, 占空间, 我就用了 trie 存 pattern, 然后查找, 面试官说 trie 他第一次用, 可能 work 他要研究一下, 写了代码, 我感觉他有别的解法。这轮可能跪了。

Solution

感觉是 NFA 的节奏

## P141. detected loop in networking and in directed graph

烙印。先问怎么在网络里避免环的产生, 我说有协议做这个像 OSPF。然后感觉他不满意, 就问怎么在有向图里 check 是否有环。输入输出都自己设计。然后我用 DFS 写了个, 有个 bug, 有一种 case 过不了。后来想改, 他说算了, 然后问我怎么能不用额外空间完成这个, 说我这个每次都要记录此 visited 的 point, 想了半天不知道怎么能不用空间, DFS BFS 貌似那个都需要空间。估计跪这了。

## P142 找后代

有向图, 给一堆 nodes 有, 每个里面有 id, 有 parent\_id

实现获取当某点的子节点 function 1

实现获取当某点的后代点 function 2

follow up:

先用  $O(n)$  time 做了, 然后问如何做到  $O(1)$  time

取后代的时候注意什么问题?

1、自己是自己的父节点

2、nodes 里面有环 ( 如何防范 ? )

### P143 Inorder Connect Nodes

给一个 tree，每一个 treeNode 多 2 个指针 pre next 但是是空的。要求按 inorder 的顺序把指针连起来。

Lz 用 recursion。写的不好。

followup 是怎么实现 insert 方法。还问了时间复杂度和空间复杂度。

请问 insert 方法是指什么意思？谢谢， 就是问你怎么维护链表的顺序。。。

### P144 Tournament Tree

完全二叉树 parent 是孩子中的最小值，请找出整棵树第二小的值

example:

```
*      2
*    /  \
*   2    3
*  / \  | \
* 4   2 5  3
```

\* In this given tree the answer is 3.

解法 一开始装傻提出  $O(n)$ ，两分钟后给出了  $O(\log n)$  的解法

follow up 的 多叉树

### P145 大数加法

big integer 的加法 这里我一开始没有问清楚 没有考虑负数的问题 导致跳坑一下子跳不出来了 但还是把正数加法做了出来 出现负数的也弄了差不多 就是整合出了一点问题 我最后乱了阵子

刚刚面完的狗家电面，面试官是个国人大哥。一上来也不介绍什么，立刻上题目，非常容易的一个题目，给两个整数数组，每个数组分别表示一个整数，然后相加求和，返回一个数组表示和。

### P146 Single Complete Cycle

Given a circular array of relative indices, find if it is a single complete cycle.

Example:

a = [1, -1, 2]

```
a[0] -> a[1] -> a[0]
a[2] -> a[1]
return true
O(1) space
```

## P147 All Prime Numbers

第一题：打印所有的质数 (print all prime numbers) · 我已开始没理解清楚范围，设了个 `i < Integer.MAX_VALUE`，为了第一题秒时间，我没多想就用了俩 for loop 写出来了，他说可以，这样能保证小范围的质数短时间打出来

第二题：follow up · what if keep printing to the number that greater than the biginteger by using java, 没想明白,我说 java 里面有 biginteger 的 library · 大于 long 的长度 · 他说这样用第一题的方法的话越到最后时间越长 · 两个数间隔好很久 (假设 n 到正无穷..)

因为是个  $O(N*N)$  的解法 · 有没有更好的解法 · 大叔很有耐心 · 一直在等我想 · 也给了些 hint · 我最后没想出来 · 后来时间到了 · 感觉药丸。

后来他推荐说经常上 project euler 上面刷刷题 · 我说好的 · 估计是暗示我刷的不够吧 · 本来专业就不一样 · EE 出身半路出家转 CS · 最后一学期发现 EE 工作很难找于是赚了 CS · 还是挺吃力的 · anyway · 先写出来给地里的兄弟姐妹们看看面经吧 · 攒个人品

Solution

筛法求质数  $n \log \log n$

线性筛 约为  $\sim n$

<https://xuanwo.org/2015/03/18/number-theory-sieve-method-to-find-primes/>

## P148 Max Elements in One

有一个 double 类型的数组 · 找满足  $[a, a + 1)$  的最长序列含有的元素的个数 · eg.

[ 1.0 , 1.3 , 1.5 , 2.3, 3.5], 最长的是 [1.0 1.3 1.5], 应该返回 3。这种小学 5 年级的数学 (数组操作) 题对我来说是很困难的 · 我直接说暴力解法呗 · 然后面试官提醒了下用 greedy 的方法。然而代码还是写的很艰难。。。大家可以自己写写  $O(n)$  的方法练习一下。

## P149 Binary Watch

比如：

```
1000 >
8
000011 >
3
```

时间就是 08:03

让输出所有亮三个灯的时间

## P150 Grid Class

write a grid class

1. set the height and width
2. set value at arbitrary point
3. get max value of the grid

一开始问我用什么结构存这个 grid，我说 vector，他问为什么，我说可以 random access。

然后做题，先直接 brute-force，2 问  $O(1)$ ，3 问  $O(n^2)$ 。他说取最大值太慢了，我就用了个

priority\_queue (他还问了我 priority\_queue<int, vector<int>, less<int>>参数的意思，

说他不太清楚。。)，2 问变  $O(n^2)$ ，3 问变  $O(1)$ 。他说不行，set value 太慢了，我说创建一个

vector<int> maxVallue(m, 0)，记录每行的最大值，这样 2 问复杂度是  $O(n)$ ，3 问是  $O(m)$ 。

他又否了我的想法，说如果是一个  $1 \times 1000$  的 grid 复杂度还是很高。我说 binary search tree 可以都  $O(\log n)$ ，他说有 duplicate 怎么办，后来我说用 multiset，这样都是  $O(\log n)$  并且可以重复。

(PS: 因为太久没用 multiset+时间太赶，用法写错了。。。multiset 是不能改变值的，只能找到了然后删了再加新的，当时没时间了就跟 set 一样直接 `ms[grid[i][j]] = newValue` 了。。。应该是 `ms.erase(ms.find(grid[i][j])); ms.insert(newValue)`。) 悲剧

## P151 interweaved iterator

然后问如何优化，一共写了两个版本

一个 stack 和 queue 的合体，有时候从前面 pop，有时候从后面 pop，问如何复原原来的序列 (原来是排好序的) 这个好像之前面经里就有，可惜当时没细想，还是面试当场想的。。

这个题的下一问，如果 pop 出来的数字有时候放最前面，有时候放最后面，如何复原

后面俩题都是只讲了思路，没时间写代码了

## P152 Abbreviation

单词的缩写 apple --> a3e 这一轮应该是 lc 原题吧

第二问 每个单词可以有多少种缩写，返回所有缩写

(这里楼猪 sb 的写了一个 bug，小哥让 run 一个 sample，run 时发现有个 bug 简直想要抽自己。。)

第三问 给一堆单词 如何使得缩写不重复

第三问没来的及写代码了，只讲了思路。。。只能怪那个 bug 了。。

### P153 Lamp Overlap

也是经典题了，一个屋子  $M \times N$ ，一堆灯，照耀半径是  $r$

问是否有灯 overlap

写了 naive 版本，优化版本

问还能怎么优化，又说了几种能利用屋子大小来优化的，这个题似乎还可以用线段树，但我不会写

又问了一些细小问题，但没有换题了，就聊天了

### P154 Left Sum Equals Right Sum

国人小哥，非常友善

一个数组里找某个 index，使  $\text{sum}[:i] == \text{sum}[i+1:]$ ，也是经典题。我一开始用了  $O(n)$  space，follow up 就是优化成了  $O(1)$ 。

这里代码写的有点慢，但都在没给提示的前提下 bug free 了。

follow up 此时要求数组和带有权重，每个 nums 需要乘以一个 weight，这个 weight 等于和某个 index 的距离。

eg:

`nums = [1, 3, 5, 7, 8]`

假如当前处理到 `nums[2]`，则  $\text{leftsum} = 1 * 2 + 3 * 1 = 5$ ， $\text{rightsum} = 7 * 1 + 8 * 2 = 23$ 。

这道题其实也不难，我找到思路后跟面试官说了，他表示赞同还举了举大拇指（人真是太好了），但时间不够我写代码了，只写了几行。

### P155 长方形面积和

白人大叔，题目是 google 面试里遇到的最难的一题哈哈，问，给一堆长方形，求它们的面积和。长方形自定义。

### P156 File with dependency

Buildfile with tag and dependency, return one of the invalid tags.  
Toposort 搞之，但回家之后才发现自己搞错了复杂度...这里感觉会特别悲剧。

### P157 Startwith

给一堆有序的单词和一个 prefix，叫你从单词里找出 range 是以这个 prefix 开头的，我第一感觉是 binary search。回头想了一下这题要是多次查询的话应该用 Trie。

### P158 QuadTree Nearest Neighbor

第二轮是个白人大叔，感觉像刚嗑药似的非常嗨，不过人不得不说非常的 helpful。。LZ 代码基本上是他一步步指引着写出来的，只是感觉他在旁边比我还激动。。。题目很麻烦是他们现在在做的 maps 的一个功能，大概是一个四叉树，然后让你调用他们已有的 API，寻找距离最近的前几个饭店，主要是讨论，code 时间很短。

### P159 Rain And Bar

第三轮是 reverse shadow，国人大哥面的，题目就是给你一个 matrix，里面的数字代表 bar 的高度，现在说降雨量如果高于 bar 的高度水可以漫过去，降雨量 0 开始每天+1 这样，问最早第几天水可以有一条路径从 src 漫到 dst。这轮也是讨论 optimization 很久，最后用 bfs 写一个 subproblem。

### P160 LinkedList Add One

第四轮是一个印度大姐面的，人挺 nice 的必要的时候也给 hint，题目就是单链表版 addOne，不允许修改原链表，然后要求时间  $O(n)$ ，空间  $O(1)$ 。

### P161 (LC340) Longest Substring with At Most K Distinct Characters

这题如果 cannot fit into memory 应该怎么办呢？多谢！！

哈哈，我 onsite 就背面到了这题，我是这么做的：用 hashmap [char, int] 碰到已有的单词，更新新的位置，没有则往里面加，如果 map size 超出 k，删去所有 map value 中最小的那个 key；  
输入时 infinite char stream

### P162 K Dimension Neighbors

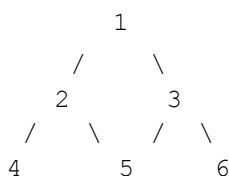
2 维坐标找与给定点相邻的点，follow up k 维坐标找与给定点相邻的点。

比如二维下给定  $(0, 0)$  这个点，那么相邻的点有  $(-1, -1)$ ， $(-1, 0)$ ， $(-1, 1)$ ， $(0, -1)$ ， $(1, -1)$ ， $(1, 0)$ ， $(1, 1)$ ， $(0, 1)$  这 8 个点

fellow up 一开始没什么思路，小哥给了个提示，让我算 3 维和 4 维的个数，我报给他答案后又沉默了差不多有 3、4 分钟（想想真是笨，都提示得那么明显了），然后想到了用 divide and conquer，然后简单解释了一下，开始写代码，大概 6、7 分钟写完了代码，小哥让我跑下 test case，然后找到了 bug，修改完之后小哥又问还有什么要改的吗？然后我看了代码大概半分钟左右，小哥说剩 5 分钟了，我们来问问题吧，面试完后看了下代码，有一个很明显的逻辑错误，但是写的时候太慌张了，居然没有看出来，但是大方向应该是对的，不知道能不能 onsite。

## P163 Number of Unique Paths in DAG

第二题，



如上图，有好多节点，给一个目标节点，求出从 root 到目标节点有多少条最短 path。比如，目标节点为 2，就有 1 条 path，目标节点 5，有 2 条 path

这题是个坑，我跳进去了。写完之后，三叔说，我不喜欢你的 code，但你写的也是对的。

$(m + 1) * (m + 2) / 2 \geq n > m * (m + 1) / 2$ ,  $f(n) = C(m, k)$

## P164 环形链表 insert

在一个排序的不重复的环形的链表里插入一个数。

## P165 iterator with memory

他又出了道数据结构题。。。说有这样一个 list，每次 call iterator 之后，后面做的任何 update list 的操作不会影响之前的 iterator，比如说 list a；

```
a.add("build");a.add("cool"); Iterator<> ia = a.iterator(); a.del("build"); 对于 ia 他的输出的东西 还是 build, cool;
```

问怎么实现这一行为。。。

## P166 Remove Subtree from Tree Array



给 Node[] nodeList, 每个 node 有 int p\_index 表示它 parent node 在数组中的位置, parent node 必在 child node 左边, root 在第 0 个, 其 p\_index 为 -1。问, remove 第 k 个 node 和它的子树, 并把剩下所有 node 按顺序尽量向左摆, 剩下 set null。TimeO(n) SpaceO(n), followup 是如果 spaceO(1) 怎么做。Time 是多少

## P167 Snapshot time

下午第二轮是出了道。。。说不清楚的题, 就是给你一个 snapshot, 里头有排序好的 item 的投票, 但不知道投票的数量, 你有一个数据库, 里面有 item 的 timestamp, vote 类型 ( up vote 或者 down vote ), 问你怎么知道 snapshot 的时间。。。暴力 dfs 把所有符合排序的时间都得到。。。

## P168 进制转换

计算题, 把一个数转化为 16 进制和 7 进制

## P169 Pascal Number

Pascal number (int row, int col)

分别用 recursive 和 iterative 的方法做。LC 上有

## P170 Longest Palindrome sequence 随便搞

找出最长的 palindrome sequence 的长度, 如果通过改造一个 string 的话 ( Operation 支持删字母, 或者 reorder 字母 )

## P171 Null Filter Iterator

Create a filterediterator that skips over nulls. Create a class that implements iterator, takes another iterator in its constructor. Your class iterates through the other iterator except for cases where the initial iterator has a null value. In this case your iterator skips over the null and treats it as if it did not exist.

example:

```
sourceIterator = 'a', null, 'b'
sourceIterator.hasNext() = true
sourceIterator.next() = 'a'
sourceIterator.hasNext() = true
sourceIterator.next() = null
sourceIterator.hasNext() = true
sourceIterator.next() = 'b'
sourceIterator.hasNext() = false
yourIterator = new yourIterator(sourceIterator)
yourIterator.hasNext() = true
yourIterator.next() = 'a'
yourIterator.hasNext() = true
```

```
yourIterator.next()= 'b'
yourIterator.hasNext()= false
```

## P172 Word Path

第 5 轮一个老白资深程序员吧，估计我就是挂在他手里。这轮本来时间就不够他还要去找水笔拿水估计只做了 30 分钟结果没写完思路是有了不过不是最优解。题目是给你一个 String 里面有 a-z 0-9 可以转换成一个 m\*n 的 char 矩阵，然后给你一个 String 让你在 m\*n 矩阵里面寻找 s 并且输出寻找指针移动的过程（过程不唯一随便输出一个）。

Ex :

```
s e c.
t b d
a r d
```

如果让你找 star 你就输出“ddl” move down down right

我用 dfs 穷举法做每次 4 个方向做穷举貌似他不满意还找出我一个 bug。。。

总的来说题目都不难，都做出来了但是总免不了粗心出现了几个 bug 可能这就是挂掉的原因吧。大家好好努力面试细心，希望还是很大的。

比如说给你个矩阵

```
A B C
D E F
G H I
```

给你的 string 是“ECG”

然后你就找 ECG 在矩阵内的位置 (1,1), (0,2), (2,0) 根据这 3 个位置你就可以知道最后移动的步骤了

开始 (0,0) -> (1,1) "DR" (1,1) -> (0,2) "LU" 等等

时间复杂度  $O(m*n)$

## P173 House Robber Path

第三轮国人姐姐，input 一个 array，都是 positive int，求输出和最大的 subarray，要求

output 不能包含 input 里 adjacent numbers,  $O(n)$  time。e.g. input: {2,1,4,6,10}, output: {2,4,10}

HouseRobber 的变种，加一个 path 数组去存路径。

## P174 色子和

给了  $n$  个骰子/色子，每个色子有  $f$  个面，点数从  $1, 2, 3 \dots$  到  $f$ 。所有色子视为 identical/unique，问题是：投出  $n$  个色子最后点数和加起来等于  $s$  的情况分别有多少种。

## P175 Min Variance to Ten

说给一个 array，然后调一些加到新的 array，头尾必须要，然后给你个条件，新的 array 里的 2 个元素之间差别尽可能接近 10。

## P176 Design HTML

9 月 1 号电面面经

感觉是个白人姐姐，很和蔼，很亲切。一卡就给提示。

第一题

1. 设计 HTML DOM 数据结构

信息是

```
<html><body><p>Hello</p></body></html>
```

或者

```
<html><body><p><b>H</b>ello</p></body></html>
```

【提示是类似 tree】

(1) type url 后会发生哪些事，

建一个数据结构表示 html，就是设计 tree 吧，多个 child 然后还有一些其他的 methods 像 next 之类的

## P177 Match Plan

刚刚结束电话面试，面试官是一个印度口音超级重的大叔，

我不得不反复确认他在说什么，以及让他把问题打在 document 里面。

第一个问题是个简单的算法题，有  $n$  ( $n=2^k$ ) 个 player，排名 1 到  $n$ ，每次对阵会让最好的对阵最差的，每轮对阵淘汰一半的人，要求打印出对阵情况。例如有 8 个 player，对阵情况就是  $(( (1, 8), (4, 5)), ((2, 7), (3, 6)))$ 。  
要求就是打出这么一个字符串。

花了 10 分钟搞明白他到底想问什么，然后花 5 分钟写出了个程序。

写完他问我知不知道 `tail recursive`，我表示可能知道这个东西，但不知道这个名称。

## P178 除 36 余 1

然后一道算法题，至今未想明白，求地里大神知道，给一个 `string` 很长很长，找一个 `substring` 能不能除 36 余 1，`"agda4685766986579868687669686786786;?adfff7323"` 有 `substring` 除 36 余 1，`substring` 只能有数字，但可以很长很长，可以远大于 `long` 的范围。

## P178B 除 k 余 0

不允许 0 开头字符串，除非是单个 0

实现一个 `function`，输入是一个 `string`，`string` 里的每一个字符都是 `'0-9'`，所以它每个 `substring` 都可以表示一个整数。输出其最长的可以被三整除的 `substring` 的长度。

例如 `"1012300"`，最长的能被三整除的 `substring` 是 `"12300"`，返回它。

一开始暴力解法，后来优化到 `linear`。

## P179 Binary Tree to Doubly LinkedList

第一题 `binary tree to double linkedlist`，要求 `in-place`，挺简单的偏偏好久没刷类似的题了，讲思路，写代码，`debug`，跑 `case`，总共花了半个小时左右，中间有些小 `bug` 涂涂改改写好了。

## P180 Maze with API

第二题，给一个迷宫，当时心里很忐忑，心想会不会中奖，结果中奖了。。。他说只给你两个 API：(1) `bool forward()`；(2) `turn_right()`；不给其他任何信息，第一个 API 会尝试在当前方向往前移动一步，如果遇到 `wall` 会返回 `false`；第二个 API 就是转向；然后如果在 `forward` 的时候遇到出口就会退出，不用担心。这轮真是体现出对实际问题的抽象能力还是太弱了。。。一开始说很 `naive` 的方法就是遇到墙就 `turn right`，其实不是，因为常用的沿着一边墙的方法对这道题行不通，因为你不知道那边是墙。。。但是既然是迷宫思路就应该是 `DFS/BFS` 求解吧，一般都是 `DFS` 找出口很暴力。但是这道题的问题是，没有给迷宫的其他信息，怎么保存状态呢？我说要不用一个 `vector<pair<int,int>>` 来保存

path，记录在某个方向前进了多少步，在用一个 map 记录走了那些方向，这样遇到死胡同就用最后一条记录来 roll-back，小哥说因吹丝挺。。。。然后写到一半，他才突然又提示说，你可以从 graph 的角度思考，晕。。。。然后我就说设计一个 Point class，来边走边建图，Point 里面用一个 vector 存四个方向的 point，然后用一个 flag 数组来记录哪些方向走过了。他说 make sense。不过最后一轮这实在太累了，脑袋有点转不动，代码没写完，给他讲了思路，他就拍照了。。。。那十几二十行代码，感觉拍下来好惨。。。。他还特意留了时间，然后跟他问了几个问题。LZ 只能按捺着心情和疲劳强行问了几个，就结束了。。。。

## P181 Moving Zeros

1.1 给一个 integer array，写一个 function，输出一个 array，把原 array 里所有的 0 都放在开头，其他的顺序无所谓。

eg 输入 [2, 4, 0, 5, 0] 输出 [0, 0, 2, 4, 5];

follow up: 如果要求改成把所有 7 放到开头，或者把所有奇数放在开头，或者把所有偶数放在开头，怎么设计你的 class 使这个 function reusable。

## P182 Count of Codes

3 写一个 function，输入一个 char[][ ]，一个 int len。char[][ ] 里每个字符都是 distinct 的。输出 int count。

count 的含义是 char array 里能生成的所有长度为 len 的 code 的个数。

code 的含义是可以通过 char[][ ] 生成的字符串。code 里的下一个字符可以是当前字符的上/下/左/右字符，也可以是它自己。但是不能是 '\*' 或者 '#'。

举个例子：

```
input:
char[][ ]:
a b c
d e f
* # x
len: 2
```

可能的 code 有：

```
aa, ab, ad, bb, ba, bc, be, cc, cb, cf, dd, da, de, ee, eb, ed, ef, ff, fc,
fe, fx, xx
```

一共 23 个可能的 code。

输出 count = 23。

## P183 Top K in a Heapified Array

给了一个 array，该 array 满足 heap quality，比如就是相当于把一个 heap 按照 level order traversal 排成数组。要求写一个函数找出里面的 top K，不修改原数组，数组很大而 K 基本很小。

## P184 Group of Five

Given an array of integers find if they can be partition into groups of five consecutive numbers.

Example:

```
[1,2,3,4,5] true
[1,2,3,4,5, 9, 10, 11, 12, 13] true
[1,2,3,4,5,6,5,4,3,2] true
[1,2,3,4,5,2,3,4,5,7] false
[1,3,5,6,7,8,9,2,4,-1] false
[1,3,5,7,8] false
```

## P185 生成随机数

randomize and return an array with value from 0 to n-1

## P186 Design String

白人小哥，一头飘逸的（男生中的）长发，

时不时会遮住左眼，看着相当高冷，但是偶尔又笑得很和蔼，

但是马上又变回去。写 code 的时候，偶尔往回看的时候，发现他似乎在看着天花板若有所思。。。

第一题，给定一个 string，比如 abc，设计 add/replace/delete，

三个 function 都会传入一个 index，并且同一个 index 不会被用两次。

index 还有个特殊的地方，就是它只会是原始数据的 index，比如 abc，

那就只有 012。如果调用 "add(1, a), add(2,c)"

会变成 aabcc，也就是操作会发生在原始的 index 上。

给的解决方案就是用一个 map 存<index, iterator>

然后对 iterator 进行操作，写了一些伪代码，他说 ok 这是 warm up。

## P187 Range Maximum Query

6. given a list, the query will be a range, return the maximum number within this range, what if there are 1000 queries per second.

## P188 Max Coin III

1. there are  $m$  piles of coins, each pile is consisted of coins with different values. you are allowed to take  $n$  coins home. However, you can only take coins from the top of each pile. What's the maximum value you can obtain.

## P189 Remove duplicate characters in a string (LC 316)

## P190 find longest consecutive path in tree

not necessary binary tree

## P191 Cut Ten

第一轮：

给一堆 positive integer 如【10, 15, 20, 25】之类，然后可以对 integer 用上 cut 的方法(如果大于 10，则一次 cut 产生一个 10)，看在有限的 cut 的次数下，做多能 cut 多少个 10 出来。

可能描述的有点绕，举个例子：10 不需要 cut，20，cut 一次就产生 2 个 10，30，cut1 一次，产生一个 10 和一个 20，cut2 次产生 3 个 10；15 最多能被 cut1 次，会产生一个 10。

我说 greedy 就行。然后 coding。然后一题之后就问问题了

## P192 Symmetric Points

平面上一堆点，判断是否关于某个垂直于  $x$  轴的线对称

## P193 Partition Matrix

一个含有 interger 的 matrix，找出一个点使得左上角的 submatrix 和 到右下角的 matrix 的和相等

## P194 Number of Coin Sum

给一堆硬币，可能有重复，问可以组成多少种不同的面额，没要求 coding，我说了 brute force 和 dp 两种方法，也向面试官确认有没 better solution，她表示没有，自己也没多想，接着就问问题了

## P195 Resize Image

一共四十五分钟，先扯了十分钟我的简历和我以前做的 project，然后就过了十分钟，然后第一题是给一个 image 的 size ( $x, y$ ) 和 webpage 的 size ( $\max_x, \max_y$ )，然后在 image propotion 不变的情况下能输出的最大

size · 这题就是要 handle error condition · 因为 pixel 不能是 float · 然后还得考虑 0 · 以及 (2,1,1,1) 返回的是 (0,0)

## P196 Low-Cost Search

第二道题目是 guess number game with cost · 写个方程返回 best first guess number · 我看了这题就知道是 dp 了 · 然后等我给他解释了我的 optimal substructure · 结果就剩五分钟了 · 还有一点没写完就到时间了 · · ·

<https://projecteuler.net/problem=328>

## P197 Word Square

BALL  
AREA  
LEAD  
LADY

以上是个 words square · 就是沿对角线对称 · 第一题问怎么 validate ; 第二题给一些 words 问怎么生成所有可能的 word square · 回溯写完之后优化 ·

## P198 Maxium SubBST

给一棵二叉树 · 求一棵最大满足 bst 子树的 size ·

## P199 Disjoint Interval Tree

interval data structure

给一堆 interval · 用一个数据结构储存 · 遇到 overlap 就合并 · 写一个 insert 写一个 print ·

## P200 Min Permutation within Distance

google 电面题 · 给一个 string · 和 max dist ·

返回一个最小的 string permutation · 每个 char 移动距离不超过 max dist

比如 goog · maxDist=1 返回 gogo

## P201 Inorder Predecessor

## P202 Morris Inorder Traversal



## P203 Bigram Ratio

lower-case letters:

A: bobo

B: bob

bigrams for A: bo, ob, bo

bigrams for B: bo, ob

b-->o countA=2,countB=1;

o-->b countA=1,countB=1;

union of bigrams: bo, ob, ob = 3

intersection of bigrams: bo, ob = 2

return ratio intersection count / union count :  $2/3 = 0.6667$ , float

### Solution

countA + countB = union + intersection

thus we just need to calculate intersection

first we use hashmap/trie to count for each k-gram in A

then we search each k-gram of B in the datastructure, decrease the count if found and increase the intersection.

trie is better in the second loop. because it has sub linear search miss performance.

## P204 Binary Search

一个 sorted array of intergers, 用 binary search 找指定的 value. 如有 重复数得找到第一个。

## P205 Rearrange Array

第二题, 一个 array, rearrange 成为另一个 array, 现在给了这两个 array, 求是怎么变化成第二个 array 的. 挺简单的就用了 Hashmap 秒了.... 萌宠哄滑-涓c浜+竺録嘴逆

然后问我, 那现在给你原 array, 也知道了是怎么变化的了, 所以我们现在可以用原 array 求出变化后的 array 对吗? 但是我要 run 这个 method 好多次比如 k 次, 怎么最快能求出 array 被 rearrange 了 k 次以后的结果? 最后我就推倒出求 LCM.

面完他亲切的用中文跟我说, 我是他见过面的最好的, 时间复杂度最低 trade off 也说的好的. 谢谢小哥给了我信心~么么哒~

P206 Swap Nodes Reference in LinkedList

P207 Sort Array Square

第一题：给一个 sorted Integer array, 返回 square 后的 array.主要是要考虑负数的情况，很多 trick 的地方比如全都是 negative 或者全都是 positive

P208 Sort with Slot

然后给一个 unsorted slots with numbers，比如：【3,1,0,2,x,4】，x 表示位置为空。每次只能移动数字到空的位置，不能直接两数字互换。让排序这个 slots，并且唯一的空在最后，比如【0,1,2,3,4,x】。数字只包含 0 到 n-2

P209 24 点 II

b. 给四张牌 返回四张牌的 expression 可以算出 24 点（带括号）

P210 Logging Trees (GCJ2015)

P211 Smallest Kth Element in Partial Sorted Matrix

有一个 2D Array，行和列是 非递减。找到这个 array 的 smallest Kth element

P212 Keys and Words

1.2 电话上的每个数字都对应三个字母（除了 1 和 0），设计一个 function，输入一个电话号码，怎样输出所有的这个电话号码可以组成的能 make sense 的 word。

dfs + trie，时间不够了，大体说了下思路。

P213 Number of Squares

给一个  $m \times n$  的矩阵，求一共多少个正方形。

P214A Merge Stones I

“石子合并 (<http://blog.csdn.net/acdreamers/article/details/18039073>)”，

一开始问了随机合并任意两堆用什么算法，我说这题我会，用贪心算法，然后就问了时间复杂度和空间复杂度，最后得出结论是用 Heap 最省时间。（Huffman Algorithm 霍夫曼编码）

## P214B Merge Strones II

然后问了合并相邻两堆（看那链接里的文章），写了方程式，他还问我你知道平行四边形法则吗？  
我说知道，可以把石子归并时间复杂度简化到  $n^2$ ，然后就没问了。

## P215 Concurrent Users

给一堆 `logon time` `logoff time` 和 `user id`

求任意时间在线人数

OST

## P216 Number Of Squares formed by segments

Round 1. Consider two kinds of line segments. Each line segment is either vertical or horizontal. Their coordinates are all integers. Given a list of line segments, how many distinct squares are there?

## P217 Implement External Sort

## P218 Simple Maze Problem

#1 Maze problem. Given an array filled with 0,1 and 2. 0 is the start point, 2 is the exit. 1 is the wall and can not pass through. Determine if there is a path to go out. (So easy!) Follow up: discuss the difference between BFS and DFS, w.r.t. Computational complexity and Memory cost.  
Follow up print one of the path if exist

## P219 longest subarray less than target

Given an positive array, find the longest subarray which the sum of the subarray less than or equal to the given MaxSum.  
Follow up: contains negative numbers

## P220 Filll in Partial Sorted Matrix

一个  $5 \times 5$  矩阵，填入  $1 \sim 25$ ，每行升序，每列升序，求一共有多少种方法

701149020

## P221 排版最少 Cost

给一个宽度  $w$ ，和一系列单词，每个单词的长度均  $\leq w$ ，

现在将单词排成若干行，每行的单词之间有一个空格，每行的单词加空格长度不能大于  $w$

定义最后一个单词到行末的距离为每行的 `Cost`，

总 `Cost` = 每行的 `Cost` 的平方和

求最小的总 Cost

## P222 轮流取数

AB 两个人轮流取一个 array 里的数，都只能从头或者尾里选一个取，A 先。B 采用的是 GREEDY 方法，也就是 B 一定是取头、尾里面最大的那一个。问 A 能取到的最大值。比如说 2 4 7 3，A 可以先取 3，B 肯定取 7，A 再取 4，B 取 2，这样 A 取到的和是 7。另外就是 A 先取 2，B 只能在 3 和 4 里面选，B 因为是 GREEDY 所以肯定选 4，那么 A 就能取 7，这样 A 取到的和是 9。9>7 所以返回 9。

## P223 At Least K Neighbors

给一个无向图  $G(V, E)$  还有一个正整数  $k$ ，Graph 表示  $V$  表示点， $E$  表示边，要求找到  $G$  的最大的 subset  $G'$  使得  $G'$  中每个点至少每个点至少有  $k$  个 neighbor。

## P224 Insert Comment Block

Round1: 给 google doc 的 comment block 写一个 insertion 函数，默认每行只有一个 comment，block 高度已知，block 的顺序要跟他们的 line 对应。

我就很直接的建了一个 linkedlist，每个 node 是一个 block，然后插入就是找到这一行应该去的位置（想写个 binary search，但是因为后面的操作是 linear 的，面试官说不用写），然后插入，然后挨个往后看有没有 overlap，有的话就要更新这个 block 的 start, end，加上 offset，直到没有为止，特殊情况是要加在第一个上面。

## P225 Most Possible Sentence of Length K

Round4:

一堆 word，每个 word 和每个 word 都有一个比例  $P(w_2/w_1)$ ，代表在这个语言的所有句子里面， $w_2$  出现在  $w_1$  前面的概率，现给一个  $K$ ，求在这个语言的所有长度为  $K$  的可能性最大的句子。

举例：The cat is cute: ->

$P(\text{START}/\text{cat}) * P(\text{cat}/\text{the}) * P(\text{is}/\text{cat}) * P(\text{cute}/\text{is}) * P(\text{END}/\text{cute})$  START 和 END 就当是特殊字符，跟所有字母的概率已知

DP  $dp[i][w]$  表示在位置  $i$  上以  $w$  单词结束的最大概率  $dp[0][w]$  就是所有 word 跟 start 的比例

转移方程是： $dp[i][w] = \max([dp[i-1][w'] * P(w/w') \text{ for all words in words}])$  复杂

度 - - 》  $k * N^2$

## P226 Save Modification

两个 google doc 的 template, json 形式, 一个是 user, 一个是 template, user 做了一些修改, 求怎么最有效的存 user 的版本, 其实就是求两个 nest dict 的 diff, 很 efficient 的表示出来, 先设计个 schema。

在任何 level 都只存在着: 加了一个 key value pair, 删了一个 key value pair, 以及修改了 value。其实也就是简单的 dfs, top down 的。楼主绕了很久才恍然大悟, 最后一轮太累了, 大概的代码如下, 循环还需要加个相对 path, 好写进 db, 一个比较巧的地方就是第二个大 for 里面只顾及到了 add 和 modify, 加上第一个 for 才能顾及 remove, 而且它们不会有重复。

## LeetCode

054, 042, 200, 295, 308, 329, 271, 140, 353, 345, 354,  
140, 200, 305, 361, 351, 057, 298, 128, 281, 298, 128,  
288, 167, 46, 47, 31, 246, 247, 248, 340, 066, 029, 159,  
218, 171, 168, 266, 267, 253, 363, 010, 340, 317, 160,  
056, 121, 122, 123, 188, 309, 281, 149