

利口205

判断一个字符串里有没有重复字符

check if two string is an anagram

bit operation, reverse a 32 int

第一题是类似lc上的原题，水浪能到左上和右下的那题。第二题是3个有序列表，从里面分别取出3个数字，a,b,c，要求找到最小的范围，就是 $\max(abc) - \min(abc)$ 最小，lc上有类似的两个列表的题。第三题是从一个二分查找树中找到topK的node，地里有帖子讨论过，G家的题，我面G家时候碰到过这个题包装好的。基本就是不动树的话 $O(n)$ ，可以修改树的话，创建树的时候每个节点存一下左子树的节点数量，可以 $\log N$ 。告诉他我见过，然后描述了2个思路。最后10分钟又做了个题，也是太复杂描述不出来了。。。完美解法应该是一个从来没有听说过的数据结构，他们肯定也没有指望我能用出那个数据结构，而是想看我的思路。最后一分钟给了一个思路上比较接近的了，然后就结束面试，他们告诉了我那个数据结构，中文学名叫四叉树。

题目是找两个数组的intersection

LC236

write a function find all the strong linked items and weak linked items. define: if buyer a bought item A and item B, buyer b bought item B and item C, then item A and item B are strong linked items (same buyer), item A and item C are weak linked items (buyer a and buyer b both bought item B).

算法题目是实现LRU cache

李扣山武酒，加延伸，如果节省空间（用bucket计数而不是每个都存），如果在牺牲一点精确度的情况下，更省一点（只留固定比例的日志，比如百分之一的）

Coding:

Given a list of strings: ['running, 0800, 1000', 'swimming,1000,1100', 'eating,1100,1200','running,1115,1330...']. The meaning of the strings are 'activity, start time, end time'  
prob1. Write a function to output all the unique activities in the list  
prob2. Write a function to calculate the duration of each activity in minutes  
prob3. Write a function to merge all the time intervals in the list that have overlaps, leave the non-overlapping intervals alone, and finally output all non-overlapping time intervals after merging those that overlap.  $O(N^2)$  solution not acceptable.

implement fooService throttling control。Requirement是每秒之内只能 accept 100 个requests，超过100 个 request，就throttle

## 两个list找重合，翻转单链表

电面直接是coding，给了一道board game 的编程，侧重oo design

一步步由浅入深引导你去写

initialize() board, 放置棋子，开始游戏，每进行一步，打印棋盘

三个嵌套的sync的function 用上一个返回值做下一个的参数。改写成async的function，无返回值，用callback完成

面完想到应该是多线程。。。

## 利口 402

party = [person1, person2, ...] couple = [(person1, person2), (person3, person4), ...] 找出 party 里所有的 couple where 夫妻两个人都在party里。

给一个DAG of jobs，箭头A->B表示A要在B之前执行，求一个合法的执行sequence

给一个string source, 比如 "aabcfc", 给一个list of word, 求list里面可以用source组成的word

有点忘记了，似乎是number of islands + 一个类似BFS的题

设计一个游戏排名系统 + 给一个BST，找出在BST中有多少个比N大得数

用链表实现队列。打印树的层级遍历之字遍历

算法，interval那几道题，单词查找那道hard(一个印度manager出的，打出来之后，明显态度翻转了)

给一个先增后减的array（没有重复数字）和一个数，查找这个数是否在array里，返回T/F。

Followup：时间复杂度是多少。如果array里面的数有重复，时间复杂度最差多少。

先是一个肥波拿起数列，我当时都愣了，以为理解错了，怎么出这题。面试官还以为我不懂，耐心滴跟我解释这个数列是怎么回事，还提示我说用两个变量搞出来。第二题是BST问题，实现一些基本的插入、查找、判断相等、删除的方法

找数组里第一个不重复的数字

链表倒转

倒转域名，也是lc原题好像

Python实现一个函数，生成一个list包含前n个菲薄垃圾数

设计一个函数，输入n表示n秒，输出n秒的自然表示，例如3721就是1 hour 2 minutes 1 second

生成一个完全二叉树，树的节点是children指向parent。返回一个array包含所有叶子节点

求两个集合的intersection, 要求guaranteed running time并且不能使用额外空间（output不计入额外空间）。这里有点小trick，因为input是set，并没有说是hashtable还是treeset，所以不能调用input set的contains函数，否则有可能出现O(n)以及O(n log n)两种情况。

Q1 : reverse a string Q2 : LC Move Zeros

加密的3萨姆 (李口易武)

有回文字符串

expire map

实现一个ExpiringMap的两个操作, put操作和get操作, 超时了取出来为null, 没超时就能取出正确的value。

```
public void put(K key, V value, long durationMs) {}
```

```
public V get(K key) {}
```

ex:

```
// put(k, v, 1000)
```

```
// get(k) -> v (less than 1000 ms has passed since put)
```

```
// get(k) -> null (more than 1000 ms has passed since put)
```

给一个2d maze, 每个格子有相应的数字, 所有0到2的路径里, 返回cost最高的路径。

1. Valid Palindrome

2. Shortest Palindrome

coding 给一个array让做一个quick sort; 如何实现heap。

黎挺 伞霸气 找出字符串里第一个 uniq char.

先说用 map 记 frequency, 第二遍 return 第一个 frequency 1 的

followup 一遍 iteration, 我说用两个 set, 一个碰到就加进去, 并且把字放在 candidate set 里, 再碰到就从 candidate set 取出来。最后想了一下, 没有办法 make sure 是第一个。忘了最后提示下写没写出来了

a find prime factor of a given number.

设计一个游戏排名系统 + 给一个BST, 找出在BST中有多少个比N大得数.

黎挺 齐伞鹅 变种, 给 Event(trainId, type(in/out),timestamp, user, station...), 求每个车 peak time. 就拿个 map group 起来然后 sort timestamp 然后进加出减就好了 (我刚开始还习惯性的分了两个 array, 最后看没必要啊直接 sort 全部 run time 判断就好了)。然后 follow up 如果每个站有多个 in/out, 先下后上, 只算过后的 peak 怎么办。答 customize comparator 把 out 的排在 in 前面就好了。那么下一个 loop 里的 Iteration 加的时候总数就已经减过 peak 了。白叔表示很满意, 剩 5 分钟让白女出个题。

黎挺 鹅柳坝。先开始装逼说可以 sort 然后查, 但是其实可以把所有的 index 加起来再减 num 最后 return. 白女很满意表示超出预期。白叔说可以改进运算虽然 runtime 一样。我一拍脑门说直接用  $n(n+1)/2 - \text{sum}$  就好了。

黎挺 斯午 DP 秒了

两个排序好的array合并

tree level-order sum  
關哩扣原題 武拾陸

找到一个sorted list里面target的最小和最大index  
 $a = [1, 2, 4, 4, 4, 8]$ , target = 4, return [2, 4]

how to multiply two large numbers ? Leetcode Multiply Strings

LC98, 但是Node里面的数据类型未知, 不能当做整型。

LC32, 不同的是需要输出最长的valid parenthesis。用dp做的, 但是吃饭的时候面试官说其实他心中的完美答案是stack的做法。

LC341, 要求必须是real iterator, 不允许unpack数组。最后 $O(\log n)$ 空间和 $O(1)$ 时间。

给一个list of pairs of words, 每一个pair 代表一对近义词, 然后给两个 sentence. 判断两个sentence 是否相似, 相似的定义是对应index的word要么相等, 要么是近义词。follow up是如果近义词属性, 可以传递, 怎么解决。(用什么data structure)。

BST变成顺序双向链表, 要写recursion, 方法是每次返回最左边和最右边的节

给一个bst和一个target, 输出小于target的最大值

binary search, 真的就是找一个sorted array里面有没有target

写个iterator class, 遍历二维数组, 其实就是leetcode的flatten 2D vector

group shifted string

jump stairs原題

find top k in DAG , graph leaves with associate with a weight. if graph too large, how to efficient not to read all of leaves

given clicks: url, session, query e.g. ([www.starbucks.com](http://www.starbucks.com), starbucks) ([www.wikipedia.com](http://www.wikipedia.com), starbucks) ([www.peetscoffee.com](http://www.peetscoffee.com), starbucks) ([www.starbucks.com](http://www.starbucks.com), starbucks nearby me) ([www.peetscoffee.com](http://www.peetscoffee.com), starbucks nearby me) ([www.peetscoffee.com](http://www.peetscoffee.com), peets coffee) find co-query sorted by count (starbucks, starbucks nearby me) 2 (peets coffee, starbucks) 1  
mapreduce job

given 1billion strings around 1TB, 64GB memory, how to find 1000 string exist in the system (pick first 8 prefix, sorted and build b+ tree, also sorted 1000 string and search it)

mapreduce : find all query that with the same R1, how to deal with performance issue when one key is too busy. (add random id and run second reduce job)

coding题还挺直接的，就是一个topological sort的job scheduling题，然后同时找到这个tree里面有哪些duplicated structure。比如说给你一个tree，你要返回重复的subtree，再返回一个它的serialized version.

然后就是design题。问了一些server的问题，一些error handling case，怎么去partition database，然后让你把interface写了一下。

Just two sum. First method: Hashmap Second method: Sorted and two pointer

已知一个有向图的所有边，HashMap<String, Set<String>>, 求图。 自定义interface

问了两道题 第一道是如果让你test 把一个string转换成int的function好不好用你会怎么测试  
第二道是 app的lifecycle 这道题答得跟屎一样 应该是跪了， apple doc上面就有

设计停车场

一道题， 外星人字典。

设计推荐系统 + 简单实现（沟通后是要求建个无向图）

1.max stack 只要实现O(1) pop push peekmax ; followup : 优化space complexity  
2.reserve words, leetcode原题，但是写出O(n) 算法三哥并不满意，说还有更好的写法，虽然也是O(n)。最后也不告诉我怎么写，估计是真的tricky。有知道的po一下代码哈

题目考了一个给一个string，找到第一个unique word。居然最后让我用ordered map搞，就可以一个pass

parse log 文件，很多行log, 找某些字符后面的数字，判断数字是否在范围内，不在范围内报错。

max3 elements in an array。

followup:

1. 输入是数据流怎么办
- 2.thread safe
- 3.数据量很大，大集群如何加锁？答的zk，问了很多细节，不过不难
- 4.zk如果被设置了最大qps怎么办？答得信号量，然后面试官说不行，他是要在极限qps下依然提高集群性能。最后面试官告诉我标准答案了，好像是caching？？？（这个答案忘掉了有点可惜）

给一堆点组成的一个大圆，再给一个圆圈外面的点，找出圆周边上距离那个点最近的点。输入输出自己写，点怎么给也是自己写。写好后自己写Test case讨论复杂度和corner case。

permutation的变种

计算器变种224, 227

matrix组成岛和海洋的那题，0 是海洋，两个1在一起算一个岛。求岛个数。

### 第一轮

给一张图，每个图节点有一个值，节点可以理解为城市，节点值理解为人口数量。给定一个节点，规划路径minimize in bound values for each graph node。解答思路大致和拓扑排序一致。

left join related SQL question

### 第二轮

实时排序，返回前100首最受欢迎的歌曲

### 第三轮

使用mapreduce实现对log文件的分析

### 第四轮

data modeling questions。如何建表，怎么query表，什么时候用kafka，什么时候用spark。细节问题较多。

### 第五轮

SQL related question

1. Design a game, 我游戏玩的不多，就说设计俄罗斯方块吧，把几个data class设计了一下，后来说一下UI和ViewController怎么work

2. Coding, 设计LinkedList class, sort a linked list, 找一个array里面最大和第二大元素打印出来，把一串字符串压缩（根据前面一个String的形式压缩，把公共的prefix变成数字，比如abcdef和abcyio, 就把第二个变成3yio），valid sudoku, fibonacci数列 - 感觉问题非常随机，面试官随便想个问题来问，难度都不高，但是要求用OC写。Apple一直强推swift，自己公司里却不用也是让人无语了。我最近都在用swift，一下子写oc真是不太习惯，估计语法错误肯定是有的，但觉得应该不是因为这个failed

3. Multithreading, 自我感觉应该是挂在这一轮了。一开始就是问了一个非常open的问题，什么是thread, process, 把你知道的都说一下。run loop是干嘛的。然后coding是设计一个thread safe dictionary

4. bq, 中午和两个hm一起吃饭，聊现在做的项目，遇到过什么challenge，和同事意见不一致的时候怎么解决的，项目没办法按时做完怎么和manager沟通，如果给你一周时间去学习一个新技术，你打算学什么，如果组里技术交流让你做presentation，你打算讲什么。。。。。

coding : input, Binary search tree, target number, 要求找出离target value值最近的node. 问了好几遍Return node 还是return value

mutex vs semaphore? livelock? thread starvation? priority inversion

round 1 系统设计分布式系统，有n个主机，p个服务器（ $p \ll n$ ），然后需要从n个主机中随便选一个固定的机器当作领导（随便怎么选），唯一要求就是其他所有应用去服务器请求领导的时候，永远都返回固定的那个机器。然后就是考虑CAP原则，如何保证永远都是同一台机器当领导。。。

round 2 第一题记不得了。。。第二题太难描述了。。。就算了吧。这是面了10个公司里面，coding题中唯一一个我思路比较混乱，没有找到正确方式，直到面试官提示的很厉害了，才勉强get方法的一个题。

round 3 吃饭 加 聊天，他家太多confidential的东西了。。。

missing number I

## 第二题 calculate Window size average

double calculateWindow size average(int a) 是 function, 每次叫这个 function, 然后有一个 Window size (是一个 global integer), 每次 pass 进去一个 integer 然后让你求出最近 pass 进这个 function 的 Window size 个数的 sum 的平均值。最后他问如果多个 engineer 都在跑这个 function 怎么办, 因为要用到 static, 我没答上来。

Code challenge 就是一个空的 8\*8 棋盘, 给定起始位置和终止位置, 问 king/bishop/knight 最少需要多少步能从起点走到终点。如果不可达, 要告知用户不可达。

唯一可控的一些 coding 题目: Atio. LC8. 然后就是一个 Binary search 找 closest element (全部是实际应用场景, 你不猜根本想不到), 还有一个就是解释 map/hashtable 的具体实现, 为什么是 O1 啊, 什么情况是 o(n)

### 1. 白人小哥很活泼

#### 1.1a find prime factor of a given number

脑子抽了没反应过来, 说找到 all the prime number < num/2 然后看能不能整除。

小哥很配合的说那给你一个 input 是 array of prime numbers 吧

不给 array of prime, 改给 isPrime(num)。说跟前面差不多, 从 2 开始试到 num/2, 然后试 num itself

不给 isPrime 了, 然后我终于想起来最简单那个从 2 开始除原来的数的方法。迅速写完

2. 因为来得晚 11 点才开始面, 所以第二轮就吃饭去了。其实是一个半小时的吃饭+coding HM 面。题目是给了一张纸上有 code, 让看看哪里有问题修改。主要是 singleton, thread safe, null check, exception handling, sql injection prevention 之类的。

3. 居然碰到一个前公司同一个组在我入职之前三个月离职的。让我介绍自己干嘛的, 我照常说一半他突然提到我前组员名字, 吓成狗。然后开启吐槽前公司模式。

最后聊了个 design instagram。感觉比较重视 database design。

### 4. 年轻白小哥 new grad 天竺女影子

4.1 给了张纸上有三个 Object schema, 大概意思是每个 show 有好多 season, 每个 season 好多 episode, 每个 episode 有一组 actor 整个 show 都不会变, 但可能每集顺序不一样。episode 里是有 season number, episode number, show title, actors. 然后 actor instance。要求给一个 episode 求 getNextEpisode。给一个 API getShowByTitle 返回一个 List 可能是不同的 show 同样的名字。就简单的按照条理挨个找。我自己多讨论了所有 corner case 比如这一季结束了给下一季第一集, 或者这一个 show 整个结束了应该抛 exception



follow up 问那如果都结束了你拿到这个exception怎么办。答用别的 service getRecommendation 放别的 show trailer。

然后聊了聊 project

## 5. 白叔白女

5.1 design traffic management system. 弄了几个表感觉没说到人家心里去，然后出了个题。就是黎挺 齐伞鹅 变种，给 Event(trainId, type(in/out),timestamp, user, station...), 求每个车 peak time. 就拿个 map group 起来然后 sort timestamp 然后进进出出减就好了（我刚开始还习惯性的分了两个 array，最后看没必要啊直接 sort 全部 run time 判断就好了）。然后 follow up 如果每个站有多个 in/out, 先下后上，只算过后的 peak 怎么办。答 customize comparator 把 out 的排在 in 前面就好了。那么下一个 loop 里的 Iteration 加的时候总数就已经减过 peak 了。白叔表示很满意，剩 5 分钟让白女出个题。

## 6. 天竺 SDET

问了一些 testing 相关的问题。让我想象一下他们的 architecture。怎么 automated 各种 case。

Technical Question：就一道medium难度题30mins 找一个not null的binary tree的所有left child之中在最下面那层最左边的那个 返回这个node root也被当成是左孩子之一  
leetcode上面有类似的 但是没有原题 GeeksforGeeks上面应该能查到这个题的详细解法 题意有点难理解下面是几个例子

[1] 就返回1 [1, 2, 3, null, 5, 6, 7]返回6, 因为5虽然在6一层的左边但是5是个right child [[1, 2, 3, 4, 5, 6, 7, null, 9] 返回4, 因为最下面一层虽然只有一个9, 但是9是right child 有left child最下面的一层有4和6两个left child 4在6左边于是返回4

解法：跟平常recursive的题很像，只不过这个题除了需要在recursive call里面给一个int level的 variable之外，还需要加一个boolean（楼主用Java）的isLeft variable check每次call的node是不是left child 先loop root.left再right

之后跟一般题目一样，设置一个global variable存储最下面需要的那个node，还有一个global variable记录存储下来那个node的level 为了之后判断新的node是不是在已经存了的下面层 并且最后输出就好了

第一轮：一个data file, 里面有运动项目，比赛得分，开始时间，结束时间，要求1. 找到 valid data 2. return 总时长大于一分钟的项目

这一轮是全程open source，小哥说：我们只关心你能不能写出来，至于是Google着写出来还是自己写出来的，我们不关心。

我是在两周前的店面中得知他们要考Python，所以从头学起了。男友跟我提过要去看一下文本处理功能，所以我虽然背不下来具体function，但框架是知道的，所以很容易Google到，不过上来就如此真刀实枪，我还是有点蒙。最后的情况是，超时了十分钟写出来了。小哥要我别担心，说大部分人都写不完.....

第二轮：HM来了，说是带我去吃饭，结果把饭拿回来了，说你先吃你先吃，我跟你说话，结果说了五分钟就让我design一个停在仓库里两头有车库门的算法。两头分别是门，中间有俩sensor，车停在俩sensor中间，只知道sensor被没被碰到。design一个算法要求车不能撞在门上以及 Maximize车库门关着的时间，画出finite state machine....



follow up 1: 如果你不知道车一开始的位置怎么办

follow up 2: 如果其中一个sensor坏了怎么办

这一轮我当时要跪倒了。经理带我去吃饭的时候说，你知道那个machine吗？我说不知道（后来想起来多年前一门课学过...），他很开心的说：太好了，我就希望大家不知道，这样我可以给你讲，然后表示：你选你爱吃的，我拿个速食，你吃我说，我就当真了...

最后的情况是，每一步都在他的引导下画出来了，但真的每一步都需要引导。。这是全新的内容，无从准备，他给我讲了五分钟我就得自己上手，可谓真刀实枪之最了

第三轮：

1. 聊天，其中有一个问问你有没有distributed system经验，我没听过这个词，说没有。他小声说了一句：我们其实想找这方面的人，but it's okay. 我立马急了，问他可否解释一下什么是DS，我认真听他解释了之后立刻想到自己还真有项目经验可以说，于是就简单说了两分钟。他表示可以的，这个算。但其实我才知道他比较着急coding，因为上一轮HM的画图面超时了。

2. coding：input, Binary search tree, target number, 要求找出离target value值最近的node. 问了好几遍Return node 还是return value，最后时间太紧只有五分钟，他先让我举个例子走了一遍，然后可以写。我从if not root 开始，他表示时间不够了，你别担心这个了....

我的代码非常弱，这题我查了一下是lintcode 原题，不是leetcode.. 我思路没问题，但肯定代码跑不了。不过苹果并不关心代码可否跑，只关思路。面试官甚至表示你忘了function可以写伪代码。我写完以后他匆忙拿着我的纸走了

第四轮：非常坑的一轮BQ，第一次遇到BQ还不断的想要逼问你她想听到的标准答案的。

排着问behavior，能有十多道，能想起来的有：俩组员吵架你咋办，你经理质疑你的时候，你stressed out的时候，你没时间完成项目的时候，有人没做完东西blocked了你的时候... 每一个都要举具体的工作上的例子，每一个都要不停的问题，每一个都要问到你说出她想知道的答案，前几个我都最终达到她说cool了，可是最后一个真的很怨，现在觉得挂在这里有点可惜....

她问：release来了你没空测咋办

这个我准备过，准备了四五条，但她每个都不满意。我说我看看到底怎么个做不完 可以加班，他说加班也做不完，我说告诉我Pm和老板，寻找队友。他说不行队友也没时间。我说我从coverage最大的开始，然后看时间。他说不行，你咋知道哪个coverage最大。跟他说申请延长deadline也不行...最后告诉我她想听到的是我去和marketing重新priority feature which cannot go into production for this release... 卧槽我们从没做过这种事啊！！哪个feature在release里是好几个月前就订好的啊！！QA不可能决定这次release里放哪个不放哪个啊！无言以对...最后他告诉我，你可以假设你就是product owner... 哎，还是太实在。。

第五轮：问项目，问一条line上的consumer和producer service，一个不好用应该怎么trouble shoot，这轮比较轻松，全程就是cool, make sense

第六轮：brain storm: 苹果手机的touch screen 有哪些你不想让它有反应的use case。我想了十多个他还在不停的问what else, 当时中午没吃饭又被虐成那样，实在有点扛不住。不过最后他挺开心的，一直说你看你想了整整一页呢（我说他记录）

第七轮：给我画个system design的图，有ABCD结构的关联互动，how to test it, what tool do you need, if A Failed, 最有可能是哪里的问题，if B failed, 最有可能是哪里的问题...if new build of A fail, what's the problem 等等等

第一轮：coding: 用java implement hashmap 需要处理 collision。

提了两个java的基础问题

问了database里的join算法，一开始答得是sort-merge，讲了一遍过程，然后说内存管够，想到用hash。其实就是想问的就是hash join的实现。

第二轮：问了下database里的ACID特性

database里的lock

coding: 给一个只包含0和1的array : 01100110 要求：相邻的数字不能相同。

所以有两种结果：01010101和10101010，输出交换次数最少的交换过程 -> [[2,3],[6,7]] (01010101是交换次数最少的，输出的数字代表的是index，2和3交换，6和7交换)

求一个无权重图里的最长最短路径。也没想出什么好方法，就遍历每个点，都做一遍BFS，他说你思路没错。后来网上查了下好像的确也没什么更优方法。

copy linkedlist with random pointer，楼主用hashmap写的，写完跑通过之后，面试官说别用额外复杂度

LC第三题

类似LC贰扒溜吧 求每个房间到门最短距离的和

三个和

判断一个数是不是BST或者balance tree

LC三十三

利口 461

Anagram