

组招

小哥很nice

1. 给一个排好序的sorted array, 输出排好序的square, 如[-4, -2, 1, 3, 5] 输出 [1, 4, 9, 16, 25] 要求 $O(n)$

2. **Smallest Window**

给一些windows, 输出smallest window要求cover到每一个window

Input: [4, 10, 15, 23, 26], [0, 9, 12, 19], [5, 18, 22, 30]

Output: [19, 23]

输入输出数据结构都自己定义

SQL。。。

两轮电面

1, 一个老头 20年EA工作经验, 是个音乐家和程序员, 主要问过去的经历和项目, 题目不难 fizzbuzz. 以及怎么更好的写这个代码

2, 一个华人 base加拿大, 基本也是问项目和简单的编程

9月底onsite

五轮onsite + 一轮吃饭

第一轮 电话, base德州 一个engineer level 3的白人大哥, 问了一些linux的基本操作啥的, 还有 sharding之类的 好像没coding

第二轮 电话, base德州? 一个美籍印度人, 问了行字符大文件去重, 先按prefix之类的方式分类, 再在小文件上去重。开始没想到, 后来才想到。

第三轮 吃饭 两个面试官 一个同组的印度妹子

第四轮 带我吃饭的华人大哥, 吃完饭EA附近散步看了看环境还聊了聊之前项目的技术细节, 问了简单的编程, 删除字符串中的指定字符

第五轮 电话, 印度女Boss, base 西雅图 wordladder 2, 全程代码以及说思路。

第六轮 吃饭的另一个人, 大文件排序。基本就是split, sort, k-way merge sort. 但感觉他是想按 sharding的方式先split 再merge, 给他解释了半天priority_queue怎么工作的

感受 白人华人问的问题都不难, 感觉有点放水的感觉。印度哥们问的问题都打出来了, 但是感觉没那么溜。而且女boss有点明着要挂我的感觉。

总之跪了, 具体原因不明, 给的理由是只了解算法的基本, 而不是算法在工作中的具体应用, 这我也是醉了。

电话面试: 就一个题目, 两个单向链表在某一个地方重合, 要求给出尽可能多的解法;

1. 先给了一个 $O(M*N)$ 的解法, 两层循环;

2. 修改数据结构, 里面加了一个visited变量, 第一个链表遍历的时候设置为true, 然后遍历第二个链表, 遇到true的那个节点就是;

3. 放到set里面判断

4. 想了一下不能通过两个指针移动的原因是长度可能不一样, 先各自遍历看下长度, 然后把长的链表前面先移动一部分, 然后每次移动一步;

2. 实现String的indexOf(sub)和从一个string当中去掉某一个字符remove(ch)

发一个intern面试中比较好玩的一道题目: 面试官 白人男

开始问了15分钟简历, cloud 的好处坏处, mapreduce的原理等问题。

一串数中找missing number:

example: (1, 3, 2, 2, 5) n = 5, 4 is missing, 2 is duplicate

所以要return的就是4

开始提了两种, 一种是用hashmap, 一种sort之后做。面试官不满意时间和空间复杂度。让我时间O(n) 空间O(1)的方法。

这道题目好像和lc中的missing number 不一样, 思路和lc中另一道题很像 (First Missing Positive)

感觉这个题还是很好玩的, swap, swap, 然后就可以得到答案了。

我当时面的时候没有做过first missing positive, 所以当时是靠面试官的提示才做出来的。

电面:

1. 一个0和1组成的矩阵, 把有1的行和列都变成1, 我是先走一遍把位置存下来, 一边写还在一边想更好的方法, 结果写完了也没有followup

2. linkedlist找intersection. 直接说这题做过了, 于是面试官很开心的就不面了

然后因为有一个pending offer来不及去onsite, 就直接又约了三轮电面

第一轮:

聊简历加上问了一堆基础知识

1. number of islands

2. 设计一个bank system

第二轮:

第二轮的面试官就是给我电面的, 所以也没有多聊别的

设计一个ebook reader

第三轮:

给一个input array, 一个数, 找到subarray, 这个subarray的模数最大。

例如: input array: 3,3,5,6,7,9. mod: 5. 返回: 3,5,6, 因为 $(3 + 5 + 6) \% 5 = 4$. 可以有多种解, 返回一种就行。

问了一下hashmap, hash function之类的

问了一下什么是singleton, 怎么实现singleton

电面面系统设计真的太痛苦了, 感觉还是在纸上/白板上写写画画更容易讲清楚思路呀..

我是用的dp.

假设用一个叫dp的array用来记录, 如果只要知道最大模数就一边更新dp一边更新全局最大值就好了.

如果要返回哪一段可以得到最大模数, 我是用了另外的变量记录subarray的起始index和结束index.

ebook reader是要求在线阅读，然后要考虑不同设备之间的同步，不同设备不同页面大小的同步这些bank system的话就是要可以实现一些基本的功能，比如存取转账之类的

两个系统设计都没有要求写具体实现的代码

同步问题应该是记录最后看到的页面的那个Byte就好了

啊 我被拒了 那个subarray的题 后来想了一下 感觉是没有办法做到O(N)的 最优的情况还是N*K, K是想要取模数的number

bank 那题 我也不知道他最后满不满意，因为一直是电话说的，想给他画个框图解释都没办法。大概就是设计一下会需要实现哪些功能，然后设计一下database，讨论一下database的同步这样ebook reader就是说，假设你有两个设备，一个手机一个电脑，你一开始在电脑上读，然后切换到手机去读的时候，怎么直接从之前读的地方继续开始读呢？也就是一个书签的功能。考虑到不同设备的页面大小，可以存当前设备这一页的第一个byte，然后切换设备了之后就显示那个Byte之后的东西就好了

B) EA [onsite failed]

面的是他们家的data scientist组，onsite后得知他们总共就5个人。我觉得我整体表现不错，都基本答上来了。可惜还是没给offer，可见bar相当高。面完5轮onsite又加面了一轮coding，想让我去他们的data engineering的组，最后考虑了几天还是拒了我。

一轮 Phone screen 就是问一些machine learning model的基础

- 1) Decision Tree, how it works? for classification and for regression
- 2) Describe cases/datasets when decision tree does not work. Is it decision tree linear?
- 3) Cross validation and Regularization, Can cross validation use together with regularization?
- 4) a dataset with two features,
- 5) rbf and polynomial kernel, when to use which
- 6) does kernel method work with other classifiers than SVM
- 7) find intersection of two arrays [1,1,2,2,3,3,3,7] [1,2,3,3,5] => [1,2,3,3] two solutions : 1) sort and two pointers 2)hashmap count, find intersection of key, take min of counts

二轮 offline coding 2个小时写一个 mini版的search engine，用的就是 inverted document index. 用户提交一个query (word1 & word2 & ...)，返回含有query中word的文档的交集。要发代码给他们过上million的文本集，我把自己的paper放在项目文件下，通过命令行读入内存，建立索引，然后interactive地让用户query。1个小时就轻松搞定。面试官也表示很满意。

onsite5轮 + 1轮加面

他们家面试确实很exicting，一个题目各种follow up。

Round 1: 老外

a) 爬楼梯，一次1层或者2层。果断dp解决. 问时空复杂度，答O(n),空间O(n) 可以改进空间O(2)

b) 如何更快比如 O(log n)，给了我一点小提示用矩阵，想了一会，想到用矩阵power iteration

$$F(n) = A F(n-1) \quad \text{where } F(1) = [1; 1] \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad F_n = [\text{fib}(n); \text{fib}(n-1)]$$

$$F(n) = A^{n-1} * F(1)$$

c) 让我更快，我说再快莫非是 $O(1)$ 。然后我告诉他fibonacci数列有analytical solution可以 $O(1)$ 求到。follow-up如何利用golden ratio快速求fib(n)

Round 2: 国人，背景很牛，他收下了我的膝盖。问了一个开放题，如果我只有地毯的一块布的小sample，如何自动产生整个地图，使得 1. pattern 不要重复，否则比较难看 2. 边缘比较平滑

这个题目的背景很有意思，就是游戏公司在设计游戏的时候如何把游戏地面的背景图案自动生成。Artists只能人工绘制少量的pattern，如何利用这些小pattern产生视觉上和谐的整个地图

因为是开放题目，所以我扯了好多。先上了个game of life，然后瞎扯了multi-variate gaussian model. 最后上了KNN做regression.

最后他给出他喜欢的一种解法，用2块布搭起来 (overlap 1/2) 然后找到一个optimal的cut(折线，使得线上像素跟周围的像素的颜色值的差异最小)。

于是根据这个setting写了一个三维矩阵从最左到最右的找cost最小的path。DP轻松搞定

Lunch：逛EA，拍了很多照片 植物人大战僵尸，titanfall, starwar 等等雕塑，看的我眼花缭乱。他们家饭不免费，lunch的manager请我吃的。

Round 3: 老外

a) encode and decoding string leetcode 271.

b) 他给出一种解法 比如 3#1#4#2I Love EA => I Love EA, 说如果meta data 3#1#4#2没了怎么在string的内容中通过一个英文字典找回那些可能的word segmentation leetcode 140

c) follow up 在所有可能的word segmentation中如何找到最可能的序列. => 利用一个corpus，建立Language model using bi-gram, Maximum Likelihood Estimation

d) follow up 我bi-gram的做法会bias words个数少的segmentation 因为 概率是小于1的，乘的越多就越小。=> 从corpus里面算p(len)，放回原来的 MLE, 变成 maximum a posteriori estimation

e) follow up 如果string可以由多个句子构成怎么办？=> 经提示，topic model (LDA)

Round 4: 烙印 coding 设计一个文件系统可以支持wildcard搜索 比如 /d/*/folder*abc/file*123 其实就是实现一个trie然后递归解决

感谢地里的内推，国人大哥面的，题略难。

1. 给一个总的time interval，比如[1, 100]，再给一些tasks的time interval，比如[1, 50], [51, 100], [1, 99]，每次只能连续执行一个task，然后问最多能在给定的总的time interval里执行多少个task，刚才那个例子就是最多只能执行两个[1, 50]跟[51, 100]。follow up: 如何优化时间复杂度，避免重复计算。

解法：一开始想了个DFS的解法，每次加一个interval以后再找start time比这个interval的end time要大的time interval，然后一直搜下去。follow up的话，没有特别好的办法感觉，给了一个dp解法，先对所有的task按start time排序，然后遍历所有task， $dp[task.end] = dp[0..(task.start - 1)] + 1$ ，记录其中最大值，国人大哥觉得make sense。但是如果时间是float的或者interval跨度很大就不能这么做了，不知道有没有更好的解法。

这道题就是求最大的不重叠interval数量啊，算法导论上讲greedy的那一章专门有说过，和meeting rooms是同一个模型。只不过这题多加了个区间限制，[1, 100]，不过也很好办，在每个iteration里都判断下当前检查的interval是否已经超出限制区间就行了。

第二轮是leetcode上的一道原题，search a 2D Matrix II。offline coding是把后缀式/前缀式转成中缀式。

第一轮 phone, 写了两道题, 第一题是intersection of two linked list, 第二题是**关于Trie的**, 已经实现好了Trie的add, search, 实现按字母顺序的打印, 我是用DFS和back tracking来做的

1. first k most frequent num in an array.
2. **isBalancedTree** 这个题和leetcode的不大一样, 要求是看整个tree最短的和整个tree最长的path 差是否小于2.我的复杂度高了点, 没被满意。
3. **encode, decode a tree**
4. 一个 file system, 很多文件和目录, 目录和文件名字都可以一样。怎样设计一个数据结构便于查找文件, 文件可以很多很多。

上周五电面, 是一个国人senior engineer, 45分钟的电面问了我半小时不到, 一道maximum depth of binary tree的coding, 我代码写出来了在eclipse上测了也过了, 完后问我general java questions, e.g.: 你希望java增加什么feature? 完后问我找intersection of two arrays, 我说用hashtable, time: $O(n)$, space: $O(n)$, 完后问我left join跟inner join的区别。我都答出来了。

刚刚面完, data team, 对方是师兄....

约好了三点钟面试, 久久不见有人打电话就给Recruiter发邮件催了一下, 1min后电话打来~

OK, 现在是第一面, 所以不会太难, 让你放松一下~ 首先是问了一下简历里的内容, 问你实习的内容, 其中最challenging的是啥, 问的好仔细... Then 说说你做过的最想向他介绍的project, 讲讲里面的实现, 其中有用到NoSQL, 说说你存储的形式。然后说说你写过的一个MR问题, 把思路给他讲讲~

接下来, 开始问Java的问题, 各种各样的问题 (师兄求放过...), 问Java Servlet和MVC的具体实现。

之后是算法编程题, 写一个**多叉树的level traversal**, 自己定义和构造节点, 然后写算法吧~用了BFS。然后他说如果用DFS呢, 不许用递归, 讲了一下思路就OK。

最后他说还有啥问题想问我木, 我就大概问了一下他们每天干的活, EA主要是怎样个构造。最后师兄还很nice地给了我建议。

data 组

图里找环

分享个EA的电面。

很简单, 但是自己发挥的不好。不够奇迹的让我过了, 然后做了个offline coding test。现在也拿到了onsite. 分享下面经, 然后求好运。

一个印度小哥打来电话, 口音真重啊。不过其实他题目已经表达清楚。

第一题: 一个boolean array, 里面是good bad的flag, 如果bad发生, 他之后就都是bad. 比如good, good, bad, bad,....., bad。求第一个bad。很简单, binary search.

第二题: 就问问, 没写代码。在**binary search tree**里面找topk。我感觉我答得不好, 好像复杂度就是 $O(n)$ 吧。按照inorder-traverse.

第三题: 算是第二题的扩展, 写binary tree inorder traverse。当时不知道为什么已经心凉了, 连recursive solution都写不好。然后那家伙就说好了好了, 有什么问题要问吗? 我就放弃的问, 听说EA压力很大啊。他回答: 没有的事, 我觉得挺轻松的, work load非常合理。草草的就结束了。

感觉超级烂的一次。我接着就给recruiter发了个信说自己发挥的不好，能不能再给一次机会。结果那个recruiter就回信说我是positive feedback。
求offer了！听说EA的campus非常漂亮。到时候过来补onsite面经哈。

Electronic Arts - Data Engineer Intern Redwood

地里也没见过EA的面经，glassdoor上都找不到EA的数据面经，不知道为啥。
这个是EA的Digital Platform，所有的游戏数据都汇总到这里然后处理和分析来着。
一共三轮电面，第一轮比较简单Java数据结构，还有hashtable什么的。

第二轮比较难，叫做offline coding test，一个工程师邮件发给你一道题目，两个小时后发回去，除了程序还要有一些测试文件还要有个readme什么的。

题目是给一个matrix，里面有正有负有0，从文件里面读入，然后一个机器人从左上走到右下角，每次往右或者往下走一步，不能走负数格子。对于这条路径，路径上得每一个点，检查相邻的节点，如果有相邻的节点且不在路径上，而且值是负数，就加到路径节点的value上去一起算进sum里面，最后要求输出sum最小的路径以及sum的值。

LZ当时第一反应是，我擦这不是CTCI 9.2那道题，于是乎一个劲想套用那个方法。。。结果浪费了好多时间，最后发现得用Dijkstra's algorithm，就是最短路径算法，修改relax部分，结果到最后一分钟才匆匆写完几个例子提交了，算法基础不扎实啊胆战心惊= =

最后一轮还比较好，一个小时电面，终于给哥碰到了同胞面试官，那亲切的中式英语口语是消除紧张的最佳良药~~~~~上来就问能不能实习三个月，还说如果项目没做完回学校还能继续，不知道是不是真的有COOP。然后问了问简历上得实习，然后就是上一个文档协作网站做题，第一道找least common ancestor of tow nodes in BST，LZ一开始按binary tree写的（CTCI原题），然后面试官让优化，提示说是binary search tree，所以想了一下就改成不用在检查的时候同时检查两个subtree，先比较一下node的value。

然后让我写几个linux shell scripting常用命令，一开始惊出一身冷汗以为让写shell script。。。然后就写了什么awk，grep之类，然后让写一个怎么用grep

然后考了第二道题目，一个int数组，只有0，1，2，但是没有sort过，让写个程序in place的把所有0都移到左边，2移到右边，1放中间。这里搞了老半天，还好同胞面试官人好一直在给予积极反馈和提示，最后让用一个测试例子自己walk through一下，然后给一些测试例子。

效率很高，昨天下午面完终面，今天下午就给了结果了，终于结束了快三个月的苦逼找intern生活。。这期间没事上地里看看面经和帖子，还是获得不少帮助，所以也来分享一下自己的经历，最终结论还是找intern要趁早，别想着先刷题再投，到时候内推都不一定有用了，LZ就是太晚了，Amazon内退了竟然也没有面试。

最后，弱弱问一句去EA的有组织或者其他人的么，在Redwood，感觉在地里来说有点非主流啊，怎么住什么的完全没概念。。。

问到了singleton什么时候可能会有多于一个instance出现，要怎么防止这种情况
就是multithreading的时候，所以要把singleton的getInstance方法设置为thread safe，java里就是用synchronized关键字

1. 组裡的白人工程師聊簡歷、project
2. 組裡另二名工程師，coding test，給幾個api，說他們跑的很慢，要如何用多線程來讓他們在背後執行，全部用c++來寫
3. project manager 聊簡歷，BQ
4. team lead 聊簡歷，BQ

面试的包括leetcode的题目，还有一些比较底层的数据结构的设计，也有数据库的设计，还有对于我实习做的project followup 和refractor的建议。这一块非常的inspiring, 因为我之前都没有想过这些问题，包括scalability和robustness的讨论，都让我发现了自己的不足... 而且多刷刷bit manipulation相关的题目？我其实对bit不敏感，因为一直感觉这样的代码不容易（被我这种**）readable, 虽然效率很高。所以根本想不出来.... 哎 感觉这回面的好不开心，虽然面试官都是国人，而且大家都很努力的想carry我，但是我真的好不争气哦:(

先让我自我介绍一下 然后问了简历上的一个project
接着出了一道sliding window类型的题目
debug完能work了就38分钟了 所以就做了一道题 不知道是不是太少了 一般45分钟应该面两道吧 可能是楼主太渣了
而且写出来的第一个方法有问题 被面试官提出的test case否了 后来想出了第二个方法能work 了也就是sliding window

面试官中国姐姐，人真的很不错，但是一上来就让我写sql，大三学的database已经完全忘光（大概就是读file然后去掉一些不需要的，第二题**binary search tree实现查找，插入，删除**。开始之前还问了不少简历上的问题，问的比较细。早就知道他家基本不怎么招人，就当打一次酱油了，不过也算是dream company。。

求面完发面经啊，我目前就看到reverse string in a word经常出现

没有面经 纯裸面，两轮算法，一轮SQL（这轮跪了 好难），一轮 web app 设计，一轮 股票交易简易系统设计 加代码实现，一轮boss 让设计用户 的页面并且 讲 为什么这么设计，对game 的影响等

电面一共45分钟。

第一题问了map reduce相关概念，具体情境下的分析。

第二题问了BST lookup (int val)的实现

第三题BST给定lower bound和upper bound，找到BST里面在这个范围内的所有的node val

最后简单介绍一下data platform这个组的工作内容，面试官是国人大哥，求过给offline test!

补充内容 (2017-1-21 09:15):

国人大哥好给力，过了一个小时hr followup 给了offline test:) 谢谢大哥~

补充内容 (2017-1-22 11:03):

mapreduce是问了一下什么 该怎么分shard, 什么时候应该用master node, 什么时候不应该用

[https/ http](https://http)咋建立链接。忘了，很尴尬。然后问java垃圾回收器啥时候用，咋用，原理。

写完代码后又问我咋**测试**，我说unit test。大哥明显不满意，反正楼主学了一些课本上啥的综合测试回退测试说了说，但是自己没做过。大哥评价：呵呵，看来你没咋搞过测试。我：。。。。经历不咋多。反正全程被大哥摁着摩擦 还问了一个sql

过几个case是unit test啦，应该还有其他的压力测试综合测试回退测试（兼容性测试）

国人大哥：算法题很简单，判断一个数组是不是连续的[1, 2, 3, 5, 4]连续， [1, 0, 3,]连续。0可以匹配任何元素。这个提出了几个做法。最后的大哥给的提示解很trick。比较最大最小非0元素的差和数组长度。。。。

地理好像没啥EA的面经 实际上做为一个不怎么发面试的公司 确实没什么面经
面试官是一个中国gg 人挺好的 题目是leetcode的421 暴力做完之后又用trie做了一遍
之前看有人面倒trie 特地准备了一下 但是这个题还是很难想到用trie来做的 面试官给提示了不少 然后磕磕绊绊的算是写完了
半小时之后收到了第二轮的邀请 感觉还是好心的中国gg 有所放水 不过这题真是和其他人面的题目比起来算是难的了

第一题很简单，给了一堆interval，表示起始地点，这些interval可以连成一条链，找出最开始的地点和最终地点。

第二题SQL，卒

唉一面国人小姐姐放了一马，二面被三姐姐弄爆炸--

一面：半小时简历+Java问题：讲了最近的proect，问了 interface能不能extends多个interface，hashcode是什么，hashmap的实现，什么时候会用到hashmap，还有几个问题记不住了；然后做题：给一个string，求里面的所有字符的所有permutation

二面的小插曲：本来是约在二月三号 11am, 早上10点recruiter发邮件通知面试官临时要去开个会，能不能下午开始，然后约在了2：30pm；结果下午2点就接到电话开始面试了WTF...

二面：大概10多分钟简历+Java问题：讲下project里最有挑战的问题，HashMap的实现；然后大概二十分钟两道题: Closest Binary Search Tree Value 和 Merge Two Sorted Array；最后问了两个问题结束

EA的on camppus面试，就只有30min，问了一道题，利扣239

Longest Palindrom Subsequence.

哥哥哥 面试官，介绍了一下他们组大概做什么的，让介绍下我的兴趣。
一开始先问各种general的基础知识，比如让我说出所有知道的数据结构。
然后问BST怎么validate，没打代码。
然后问怎么反转一个字符串，又问怎么不用loop实现。我？？？
没反应过来他意思是要递归做法，提示了一下才知道，还被抓到了bug，搞了老半天
最后考了个linux命令行，完全没get到他是考察grep之类的。

第一面 三哥 给了个树 问如何找到所有的子树，子树的结构一样， 树是binary tree，可以有重复的node，先定一个函数判断子树 然后遍历子树preorder标准操作，同时用一个hashmap来记录每个节点
follow up问如果里面多于一个相同的子树怎么办， 那就把hashmap里面变成value为一个list就行了呗，然后问如果树特别复杂怎么办 没答出来

..本文原创自1point3acres论坛

第二面 三姐 先问了一堆问题 比如interface 和abstract class的区别什么的 我这个答得不好 好久没看过了，然后上题 lc347变种 分分钟秒了 又问了一堆behavior 问我为什么从ece phd转成一个码农 我说我想做对人们生活影响更直接的工作 又问遇到conflict怎么办 我说我听你们的意见 又说design不符合规定怎么办 我说我新手 听取老同志意见 反正就是各种怂 今天被告知 挂了。提示了一下 设计一种hash function,然后记录每个subtree的hash值，bottom up 的遍历顺序 不要top bottom。