

刚刚groupon onsite结束，虽然面得不太好，但是还是发个帖子造福后人吧，由于签了NDA，不能说具体细节，敬请原谅。

今天一共面了两个不同的组，一个是front end，一个是full stack，是在palo alto的site，前前后后一共5个多小时。

groupon在palo alto有两栋楼，感觉环境挺不错的，大概有500多个人把，工程师为主。有很多snacks还有free lunch。和里面的员工聊天听说公司work life balance还不错，而且公司在换了ceo以后，已经止住了下滑的势头，又开始慢慢往上走。

第一轮，seo组的manager，白人大叔，先互相介绍一下，然后他先问了一个关于文件内容替换的题目，里面涉及到regex，楼主写得磕磕绊绊的，被他嘲笑了一下。然后又写了一个读文件的题。基本还算不错

第二轮，三哥上来出了一道用javascript写的syntax很诡异的题目，他问我做过没有，我说没做过，结果他得意的露出了龇牙，最后这个题目也没做出来。然后又问了很多前端的东西，都是一些很corner的东西，非主流，还扣得特别深。典型的被三哥黑了

第三轮，goods组的manager，像是个abc。午饭+behavior question+简单的前端问题。每次回答完他也不说对不对，心里没底儿

第四轮，system design，两个白人，一个主要面试，一个shadow，人都挺好的，让从前端一直设计到db，给大家推荐一个挺好的学习网站：

<https://www.hiredintech.com/classrooms/system-design/lesson/52>，看了以后受益匪浅。整体回答还不错

第五轮，国人，上来直接和我飚中文，有点儿不适应。直接三道leetcode，都是medium不难。希望第二轮三哥不要黑我太狠，求offer，求大米！！

第一轮是一开始问简历，把简历上的每个project都顺着问了一遍（其中看到我的简历上提到了一个database相关的project，就让我把那个database的schema大致画一下（= =），然而完全没有想到会被考那个project所以schema几乎都全部忘掉了，大概回忆起来有几个entity但是回忆的过程都特别凌乱。。。当时就感觉出来面试官比较不耐烦。。。（熟简历上的每一个project的重要性。。。））后来的话是一题technical的问题，给了一个pair<int, int>的array以及一个单独的pair，需要把这个单独的pair和那个array里的有intersection的pair都merge起来，然后插入到那个array里去（array是sorted的），array里的每一个pair是disjoint的。当时这题写的挺崩的，if的条件都没有考虑清楚，然后最后没写完因为时间不够了就让问问题了。。。面试官是machine learning组的，说到他当时在做的一个project是如何machine learning判断出这个访问的人可能会盗取用户的信用卡信息。。。感觉做的东西还是挺有趣的～）

第二轮的话就是直接三个technical question，只有第一个让写了code，另外两个都是让说一下思路。第一题是给一个array，对于每一个item，如果他的following items的最大值比当前的这个item小那么就on把这个item加到output array里去，比如如果一个array是[3, 2, 5, 1]的话那么就输出[3, 5, 1]。。。感觉是挺简单的一个dp问题。第二题的话是给一个binary tree，需要把它竖着转换成一个vector<vector<int>>。最开始面试官随便画了一些竖着分割这个binary tree的线，然而我看了几分钟都没有想到那些线是根据什么画的。。。后来我就说能不能我稍微改一下线的画法，每次到left subtree就减一，到right subtree就加一，然后把所有值相同的node放到同一个vector里。比如如果一个binary tree是[1, 2, 3, 4, 5, 6, 7]（leetcode表示法）的话，那么最后的输出就是[[4], [2], [1, 5, 6], [3], [7]]。。。面试官说可以，然后就说用一个<int, vector<int>>的map来recursively解这个题。。。第三题的话是一个string inplace process的问题，我大概记得的细节是这个string

的长度是足够存output的，然后不准新建string，inplace 处理这个string（但是具体题目记不得了），我大概想了一下可能是比如给了一个“a4bc3”的string需要inplace展开成“aaaabccc”。面试的时候卡在如何确定现在处理到哪一个index的char了。。。后来面试官提醒说可以preprocess这个array，把每个开始的字母都移动到对应的位置，然后再做一遍traverse生成output。感觉跪的原因可能主要是在第一轮面试上吧，毕竟自己做过的project连schema都画不出来有点说不过去。。。另外第一轮那个题写的bug百出所以跪了也很正常。。。我突然发现举的例子错了。。。应该输出是[5,1]，假设array末尾是负无穷

第一轮，一个分布式系统有n个node，node的id编号从0开始到n-1，每个node有一个value，每个node还有两个方法，sendNumber(int id, int val)，receiveNum()，让你写一个程序，然后所有的node同步run这个程序，最后所有的node都知道这个系统里所有node的和，第一个方法是向某个node发送一个数，这个方法会BLOCK在那直到有对应的node接收到这个数，第二个方法是这个当前node尝试接收到一个数，这个方法会BLOCK在那直到当前node接收到一个数，基本上就是把这n个node看成一个树结构，从叶子往父节点发自己node的值，汇总到根节点求和，然后根节点反向发这个和给整个系统

第二题，貌似是地里的G家面经，就是OOD DESIGN，设计一个家具的class，里面有一个方法boolean addForce(int force)，就是对家具加一定大小的力，看他会不会坏。因为要扩展到木质家具，铁质家具，椅子，凳子之类的，所以要用继承之类的方法OOD DESIGN

第二轮，貌似也是地里的G家面经，就是一个LOG FILE，每一行是一个time stamp, url, user id, 比如

- 1, www.google.com, 1
- 2, www.facebook.com, 1
- 3, www.expedia.com, 2
- 4, www.expeida.com/hotel, 2
- 5, www.gmail.com, 1
- 6, www.priceline.com, 3
- 7, www.priceline.com/cruise, 2

要求将同一个人连续访问的三个网页的组合作为一个整体，求访问次数的TOP10，比如上面的www.google.com, www.facebook.com, www.gmail.com，这个组合被访问过一次，www.expedia.com, www.expedia.com/hotel, www.priceline.com/hotel被访问过一次

基本就是hashmap统计，加heap统计top10，不过我过了log file两遍还是三遍，经过他提示，其实过一遍就可以，他说一般大家都会用我这种解法，但能想到只过一遍的会很少，不过好在我经过他提示答出来了，感觉他也不会力挺我。之后问我如果LOG FILE很大，我们要处理几年的LOG FILE，问题是什么，答案应该是HASHMAP会很大，然后他问我解决，我说用分布式HASHMAP，他不认可，然后马上告诉了我答案，也没让我再想，但答案是什么我忘了，当时很紧张

第三轮，lunch，和hiring manager吃饭，那个hm刚从纽约relocate过来，女朋友很喜欢hiking，一直问我这边trail的事情，还有买房子的事情

第四轮，设计groupon, things to do这个页面的后端服务。

<https://www.groupon.com/browse/chicago?category=things-to-do>，这个面试官是这个组的，一上来就让我问他问题，我就客气的问了一下他们组是干啥的，然后他跟我讲了一大堆，基本没记住，然后突然说我们这轮的题就是设计他们组的服务。一下子就懵逼了。而且没搞懂他要我

system design还是oo design, 不过最后的结果感觉就是根据他讲的东西, 猜想他们组后端服务要一些什么功能, 然后OOEDesign+DB Design, 但一开始闲聊用了好多时间, 后来大概讲了一半时间就到了, 最后他说很不幸, 我们组还有很多有趣的功能我们都没谈到, 感觉他会狠黑我了第五轮, word ladder 2, 但是只要求输出一组结果, 当时比较懵逼, 套用了word ladder 2的解法输出所有的结果, 然后他说我做出来了, 但是加了complexity

下面是onsite面经 :

1. There are stream of requests coming in. How to design a system to get latest request, or requests in last 5 mins, 10 mins or 1hour?
2. Find minimum coins to represent a target amount, both iterative and recursive
3. How to find intersection point of two lines by given two pair of points of each line?
4. Regular expression match with * and +
5. Find all possible paths in a complete binary tree sum up to a target? The path could start from any node and end at any node but it cannot be bottom up
6. Ask question about how to find a word and its description. What about single machine ram cannot hold all data? What data structure could be used to hold data in disc?

onsite :

1. manager 问了behavior questions。话说是个印度manager, 人很nice。之后去参观公司, 没有一味鼓吹公司的号, 而是给了我很多对于entry-level engineer的建议。很感谢。然后问了一个tech问题: 给定一些数字pair (0, 1), (1, 2) ... 第二个数是第一个的child, construct一个tree, 返回root
2. 可爱的工程师, 问了些简历上的问题。然后问了3sum, 和我argue了好久算法, 被我成功说服。又问了pathSum
3. 机智活力的年轻工程师 问了我替换input stream里面的字母的问题 又扯到scalability 后来又问了设计oo design : 设计file system
4. 严肃的俄罗斯大叔工程师 问了我一个猜数字的游戏 8 bits 数字 有一个函数可以返回猜中的位数 让用最少的次数猜中 还有一些follow-up

onsite :

1. 一个袋子里, 有两种颜色的球, red, white. 每次先取一个球, 如果是red就立刻吃掉, 如果是white, 就放回去再去一次, 这次不管什么颜色都吃掉, 问已知开始袋子里两种球的数目, 最后一个球是白色的概率是多少
2. 一个比较senior的manager, 开始各种寒暄了奖金半小时, 接着问了一道概率题, 骰子是fair的, 问平均要掷多少次, 才能把每一面都碰到一次。楼主说了大概的思路, 不过还没有完全解除来, 答得不好, 在这里也请大牛们知道, 欢迎大家讨论。

然后面试很nice的说，咱们换个问题吧，接着就问了一些machine learning相关的概念，还问了几道类似brain teaser的概率题。。。

这一轮没要coding

3. top k elements of an unsorted array

写完又问了一道概率题，很简单，就是给了priori，如何求posterior，用 bayesian公式一套就可以解出来了

4. hiring manager

问了一些behavior的题，然后code了一道：给定一个array，找出是不是存在任意size的subset的和为0

又问了一些优化相关的问题

5. coding:

给两个string, s 和 t, 问能不能通过把s里的字符permute, 再加上一些任意其他的字符来组成t

follow up: 如果可以添加的字符也是给定的,

楼主很快写完了两题，后面面试官说没有问题了，于是聊了一会儿他做的项目，然后发现也是各种machine learning相关的知识点，不过讨论的很开心

最后recruiter跟我聊了半小时，愉快的结束了大半天的onsite

每个人都是先介绍自己，然后要我自我介绍，然后开始问问题，最后会有时间给我问问题。其实这样有个不好是我对他们做的工作如果有问题过了一个小时到最后都忘了。。。然后会问whygroupon。

1.码农美国哥哥，cv phd，不过现在做的东西好像跟cv也没什么关系。他负责一个team把ml的算法变为business可以使用的application，所以应该是属于dev。唠嗑以后开始codepad，题目是简化版的Befunge，写完代码之后开始问怎么design这个程序而不是一个，我说了一会之后他说你所说的就是oop啊！然后就说需要这个那个class，这个那个function什么的。感觉全程他一直在hint。。。

2. data scientist manager 亚裔女，一开始花很长时间介绍了一下她的team，就是负责制造让1去实现的算法的。接着问我以前遇到过的一个problem（标准的STAR），我就说了。然后问了个问题说你要预测deal的performance但是没数据怎么办，我就说fallback啊，那fallback去哪呢，就去nearest neighbor啊，之类的。接着问你会random forest乎？答曰不会，又问会svm乎，答曰大概会吧。具体讲了一遍之后她表示满意（可是我还没讲怎么train。。。）。最后问data science 的三个phase，准备，建模，投产，你喜欢哪个，我说喜欢1，2。

3. data scientist，中东女，是2的手下。先问的behavioral问题和2一样，我没多想答案也就一样了，没说两句她就说你是不是刚才跟2说了（可能刚才在边上听？），我说是，她说那你不用说了，我们继续。然后开始问如何tackle一个data science的problem，**【此部分内容需要积分达到133以上方可查看，您当前未登录或积分不足。您也可以点[这里](#)购买VIP通行证，然后去论坛阅读原帖】**问的是一个很简单的硬币问题（一个硬币扔100次后结果是40/60，问它是fair coin的可能性是多少），但是我没什么统计背景所以讲的不是很好，转了半天终于找到点了。最后她说你在right track但我们move on吧，问了她问题就结束了。

4. pm男。互相自我介绍之后先问我为什么要读ms，然后一些好像是pre-scripted的behavioral。接着就开始来一个很诡异的问题，说要弄个系统帮助决定银行teller的工作排班表，每个teller有不

同的service they can provide, 等等, 需要哪些data points, 做好了怎么present给银行等等。扯皮40分钟之后问我有什么问题, 问完终。