

Basic:

1. what is hash table? how to implement in C? how to deal with conflicts? (linkedList, linear probing, double hashing...etc)什么是hashtable, hash function, key为string怎么办.Describe how hash table works ?

2. 为什么要用tree, 为什么要用Queue, 楼主给举了几个例子
接下来树遍历时间复杂度空间复杂度

3. What is OOP? can you describe Inheritance, Polymorphism, and Encapsulation.

4. What is the difference between object and class ? (I have no idea on this one..)

5. what is the difference between abstract class and interface?

6. 如何自己实现一个ArrayList

7. Array,ArrayList,LinkedList区别

8. inheritance,heap

9. 基本数据结构概念：比较哈希表和二叉树，操作复杂度。电话本用什么数据结构。前缀树查找的时间复杂度（这里差点说错）。

10 .LinkedList和ArrayList的区别，我先简单叙述了一下，一个类似于Array，但可以自动扩容，查找指定index的时间是 $O(1)$,插入时间 $O(n)$ 。一个是链表，插入只需改动一下指针，所以只需 $O(1)$ ，但是查找指定index需要遍历，需要 $O(n)$ 。后来他就问我如果要插入到指定的位置各需要多长时间，我说都是 $O(1)$ 不过过程不一样，arraylist查找 $O(1)$ 插入 $O(n)$,linkedList查找 $O(n)$ 插入 $O(1)$ 。他又问如果插入到最后要多长时间，我说如果linkedList有尾指针的话都是 $O(1)$ ，他说good。

11. java和C++的区别

12. trie tree

Design

1. design 题, n by n tic tac toe

如果construct扫雷的那个map。有个trick的地方就是如何判断一个空地周围有几个雷, 一种方法是8个if statement, 我写的是array的, 然后一个循环就好了, 他说很好, 不是每个人都能想到。感觉他很满意的离开了。

2. Design an elevator system.

3. design phone book

4. design vending machine

5. 设计一个hotel的设计

6. 设计一个parking lot系统。

7. 设计一个餐馆预约系统。

设计一个餐厅。然后他说如果他和他朋友第二天6pm要来吃饭, 但是餐厅已经订满了, 怎么schedule。还问了怎么拼桌子来满足客户的要求。

8. 最后是OO Design题: 一个游戏场景里有很多cat, dog, fish, bird等等动物, 怎么设计类。这题比较自由, 任何合理的设计都是可以的。我提到了abstract class, polymorphism, factory pattern, 等等。这种开放题就要尽量show出你的OO Design知识。

9. OOD design a library management

10. 设计纸牌

ctci

1. Given a binary tree, write a function that returns true if and only if it is a binary search tree.

2. Find loop in the given linked list

3. fill color题，大概是输入是一个图和图上一个pixel的坐标，还有一个新的color，要求把和这个pixel相邻的而且颜色与这个pixel相同的pixel换成给定颜色，注意“相邻”(adjacent) 这个属性是transitive的。我用dfs递归做的，写完后要求不用递归再做一遍，但是时间不够用了，就大概解释了下解法。

4. tree traversal,不用recursion, level by level,地球人应该都知道吧

5. 找magic index, cc150原题。

6. 如何用两个Stack实现一个Queue

7. find lowest common ancestor in Binary tree. What is the complexity of running time and space?

8. Find lowest common ancestor in BST tree. What is the complexity of running time and space?

9. 二题就是给一个MxN的矩阵，有一个robot从左上角要走到右下角，问有多少种走法，标准careercup原题当然不能跪。。

others

1. Given two lists of integers, write a function that returns a list that contains only the intersection (elements that occur in both lists) of the two lists. The returned list should only contain unique integers, no duplicates.

For example, [4, 2, 73, 11, -5] and [-5, 73, -1, 9, 9, 4, 7] would return the list [-5, 4, 73] in no particular order.

2. Find the K closest points to the origin in 2D plane, given an array containing N points. You can assume K is much smaller than N and N is very large. You need only use standard math operators (addition, subtraction, multiplication, and division).

3. 给出大量三维空间内的点，找出离原点最近的n个；

4. Give student result structure:

```
struct Result{  
int studentID;  
string data;  
int testScore;  
}
```

给一个result的vector，返回一个map<ID, 最高5次平均分>

5. 求数组中出现频率最高的数。

6. Merge 2 sorted linkedList

7. 写出一个把数组内容随机化的算法。

8. 判断一个数是否是2的整数次幂。

edge case:n是否小于1, 小于1则false, 等于1则true。并且若末位不为2,4,8,6, false

1:while 循环除以2

2: $n \& (n-1) == 0$

3:对数算法

9. ? ? ? ? string找duplicate 两个string判断是不是substring

10. ? ? ? 设计一个系统, 访问这个系统可以添加item, 可以返回10分钟内被访问次数最多的排序

11. ? ? ? 给一个数组, 给一个窗大小, 让返回每个窗里面的最大值。比如{7,6,5,4}窗大小2, 第一个窗是7, 6最大是7, 第二个窗是6,5最大是6, 这样一直到最后面

12. 解释factory pattern

13. 知道一组数的range, 怎样在 $O(n)$ 时间内给这组数排序

14. 给了道3 sum, 要求给出用hashmap和不用hashmap的解法。

15. 应该是bar raiser, 在a家呆了9年的platform组的一个manager, 出了道设计题, 让设计一个reader interface, 然后要设计一个readerConcat class, 实现reader interface, 并且有两个reader的成员变量, 要求给用户一个两个reader对应的file是一个整体的假象。最后给了几个high level的设计题。基本上这个面试官相当的nice, 不管你答的怎么样都很乐呵。不过LZ面完就觉得应该是挂了。。因为设计的时候很多地方都要提示才能想到。没办法, 之前从没想过IO类型的设计题。。

16. how to reverse a string?how to Reverse word;

17. how to find the missing number in an array?

18. how to find a number who occurs two times in an array?

19. 2multiple/3multiple

20. fib num

21. given one array, find one number emerge odd times.,example:
[1,1,2,2,1] return: 1

22. given a tree, find every node's nearest neighbour. example:

```
      A
     / \
    B   C
   / \ / \
  D  E F  G    :A->null B->C C->null and so on
```

24. 问的非常基础的quick sort还有数字全排列的问题

25. From leetcode: Given n non-negative integers representing the histogram's bar height where the width of each bar is 1, find the area of largest rectangle in the histogram.

<http://www.1point3acres.com/bbs/thread-72014-1-1.html>

26. 题目就是有millions of files, 如何替换file中的某一个字符串

27. How to find the url in web log file? (use grep and regular expression)

然后噩梦就来了.....帅哥说咱亚马逊大公司, 客户多, 你这玩意儿回头让我们用的话肯定特别浪费空间, 你想个办法省省空间。还说我不需要写代码, 就让我从 higher level上面考虑。我擦我个半路出家的能把代码写出来就谢谢天了, higher level啥玩意儿我都不知道。我先说可以用个and吧, 他就说那实质上就是找交集, 我现在就是在问你怎么找交集~然后我说那用个brutal force也可以啊,

就是会慢死。人家就说我不是具体问你怎么用代码实现啦，就是从高层考虑。我很直的说我还真不了解higher level，容我考虑一会。考虑了不知道多长时间我投降了，说您给我个提示好不。他给了俩提示：1) 大问题可以分割成小问题来解答；2) 我们的disk很大一个。我就说分成小问题的话我刚刚那个代码就可以实现，但是怎么再把它们结合为原来那个问题呢？然后又想了好久好久人家估计是等不及了，直接和我说可以把俩list分割成小的list互相比，但是这样很浪费时间。我说是啊，基本上时间复杂度是立方了。然后他说是，所以还可以在disk里头装个哈希表。

这时候只剩5分钟了让我问问题。我先马上虚心问他那个哈希表和我一开始提出的HashSet原理不都差不多么，既然能哈希表的话为啥不能把这个HashSet装进那个disk里面。他解释了一下。然后我又问他回头我假设真的入了Amazon做实习那我都要干些啥。他回答完问题就拜拜了。

28. 二叉树遍历。

29. 找出二叉树中the kth largest number

30. 输出字符流中频率最高的字符的频率（略拗口。。）以及扩展。

31. 大小为N的数组，所存值为1到N-1，其中有一个重复的值，如何找出这个值。

32. atoi

33. 循环链表中删除节点的data值为指定的整数：

34. 不用*和/实现两个整数相除。

35. n个点找前100个距离原点(0,0)最近的点。我说计算距离原点距离非常简单，就是 $\sqrt{x^2+y^2}$ ，主要问题在于排序。我说可以全部算完了以后sort一下取前100个点，这个写起来比较简单，因为java里边可以直接sort。不过数据量n如果远大于100的话，我觉得可以用大顶堆Heap来做，建立一个size=100的

Heap, 然后前100个point直接放进来建堆, 然后100个以后的算出距离以后和堆顶比较, 如果比堆顶大, 就直接舍弃不要, 如果比堆顶小, 把堆顶删了然后把新元素放进去。他刚开始问我关于Heap的问题, insert和remove的时间, 还有查找最大数的时间。insert需要 $O(\log 100)$ 来维护, remove因为就remove堆顶所以是 $O(1)$, 查找最大数还是堆顶所以也是 $O(1)$ 。他问我整个算法的复杂度是多少我说是 $O(n \log 100)$ 约等于 $O(n)$, 因为插入堆的维护时间是 $O(\log 100)$, 一共有 n 个元素。所以是 $O(n \log 100)$ 。他问完这一大堆都40多分钟了, 所以让我简单写一下程序, 假设堆已经建好并且insert, remove功能都有了。

我就简单写了一下对于单个点, 如果 $\text{distance} < \text{heap.top}$ 就把 heap.top 删除把新的 p 插入到堆中。否则当 $\text{distance} = \text{heap.top}$ 的时候就舍弃这个点什么也不做。

然后traversal n 个点的集合, 每个执行一下上述过程。最后输出一下Heap, 就是把heap里边的每个point的 x 和 y 坐标输出一下。

36. 有一个int数组, 每个int的范围都是1到10000, 找出最小的不存在的数。比如 $\text{array} = \{3, 1, 4\}$ return 2 (大概意思是这个。。)

37. 输入两个string判断是不是anagram(只判断字母, 但是string里可以有其他字符, 比如"cat", "act", "t c a", "at[c]", 都会返回true)

38. 有一个unsorted int array有100个元素, 每个元素都是1到100之间的, 如果有重复的元素, 把它找出来。(不能修改数组, 空间 $O(1)$)。。。。

39. 输入两个string格式的version number, 如"1.23.456", 比较哪个大。我先把string转换成linkedlist, 然后再逐个比较。

40. 输入一个int array, 如[1,23,456], 输出由这些int构成的最大的数, 如456231。

我的想法是排序后输出, 排序的比较需要重新定义, $a > b$ iff $ab > ba$ 。

41. mirror tree。

42.设计一个sparse matrix类，方便进行转置运算。

43. zig zag tree printing

44. 怎么判断string1是不是string2的subsequence

45. reverse string但是保证space的位置不变

46. 一个array of integers,找median 和 average

47. data stream里面不断读数据，怎么找最大的前100数，立马祭出minHeap

48. 一个是给一个binary tree,return the number of nodes with only one child in this tree,

49. 给定三个集合，找出三个集合的重合。

50. 假设有一棵树，被拆分成了许多条边。问如果给定这些边，如何还原这棵树。

51. 一个编程题，100000个数，1到100000，缺一个

52. int A[] 和 int B[], size分别为m、n，说出两种方法，找 intersection，说完之后让再说一种（包括时间、空间复杂度）。

2-sum不用code，很基础，先用快排+双指针扫描做了一遍，他说可以又问了一下能更快不，稍微考虑后，提出了hash table的做法，面试

53. 判断两个字符串是不是anagram，统计字母频率，瞬秒

54. 给一个string array，按anagram把string分组，我先给出了快排的做法

55. 给一个数组 大小为n, element范围为0~n, 并且是升序排列, 没有重复, 找出missing number

56. given a large number of words: aa, ab, abc, bc.....
write a function to find possible prefix as many as possible. e.g. abc
return ab, cuz ab is in front of abc.
follow-up: time complexity?

57. given a sorted linked list which contains integers, write a function to remove all duplicates

58. tell me what is the closest power of 2 that equal to 10^9 (billion)?

59. 找出一个integer的二进制表示里有多少个1 (用right shift / bit wise and

60 . merge两个sorted linked list。都要求写在纸上然后读出来。

61. 文件系统下有50000个文本文件, 如何找出里面含有电话号码的那些文件。
我就说了用find | grep | wc, 然后也没说具体怎么写命令, 他说没关系, 是个bonus题而已。

62. 找出m和n之间的所有prime number, 如何设计test case。

63.large scale merge sort (说思路)

64. 统计一个list里面不重复元素的个数

65. Package dependency. 假设你定义了一个package class. 每个package里有一个list, 这个list里是这个package所dependent on的。比如你有你个package A, A的list里有package B, package C, 那么A就是dependent on B and C.

dependent on 可以理解为当你更新B或C的时候, 你必须更新A。给你两个 package A 和 package B, 判断A是否dependent on B. 考虑到 A is dependent on C, C is dependent on B, than A is dependent on B即可。

66. Implemente the collision functionality in hash table

67. First, define a structure to represent a node, then implement the tree printing function.

```
A
 / | \
B  C  D
  / \
  E  F
```

print

```
A
-B
-C
--E
--F
-D
```

68. 输入数据流, 求出word数, 段数, 字符数。

69. python 和 C++ 的区别

70. binary tree, each node has a weight(sum of its subtree), find the biggest one

当时我就根据这个tree, build一个weight tree然后找max

71. distributed hash table,
是从hashtable引导这边的

72. find the 95% end of a stream(integer)这道就是给一个stream, 数字不停变化, 找95% 大的数

73. <http://stackoverflow.com/questions/9195856/java-how-to-search-a-string-in-a-big-file>

74. Given an array of integers find the mode and its frequency. mode 是 the number with the most frequency

e.g. {1, 2, 3, 6, 10, 3, 5, 6, 3, 3} mode=3 frequency=4

75. Given two sorted arrays of integers, write a method that will merge these arrays together and return the resulting sorted array.
(Merg Sort, 不过面试官要求自己coding一个完整的merg sort)

76. Reverse a linked list in place

77. 在一个很长的string里找最长的回文substring, 输出这个substring的长度