CS Interview Questions

~ "Real companies, real questions" ~

"I just said some shit from 162."　　　　　　　　"★★★★★"

"Are you sure that's not O(n log n)?"

"So much better than glassdoor."

"Implement a hashmap."

"Doesn't this violate the NDA?"

"Holy shit this is amazing"

"I deem this question impossible - nobody on the entire fucking internet can do it."

"Now do it in-place."

- Insert new questions alphabetically.
- Format company name with Heading 1
- Refresh the table of contents (click on table area, you should see a refresh button)

Table of Contents

A9

Given an array, which can contain positive and negative integers, return the max subarray.
Input: [1, -5, 10, 2, -7, 4]....output: [10, 2]
corner cases: Array of all negative numbers

AirBnB

Almost everyone gets this question for the phone interview [true Spring 2014]. My interviewer said that for special cases, you only have to handle double quotes, quadruple quotes, 6-quotes, etc (basically even number of consecutive quotes). """Alex""" for example is an illegal case. The solution below is working and confirmed by interviewer.

"""

Implement a CSVReader class that will read a comma-separated-values (CSV) file from disk, parse it, and print out the parsed elements with some other separator.

Rules:
The input delimiter is the comma, ","

If that delimiter is contained within an element, then that element must be quoted
If quotes are contained in an element, use double inner quotes (escape character)

Input:

John,Smith,john.smith@gmail.com,Los Angeles,1
Jane,Roberts,janer@msn.com,"San Francisco, CA",0
\"Alexandra \"\"Alex\"\"\",Menendez,alex.menendez@gmail.com,Miami,1
one,two,,four,"five"

Sample Output:

John|Smith|john.smith@gmail.com|Los Angeles|1
Jane|Roberts|janer@msn.com|San Francisco, CA|0
Alexandra "Alex"|Menendez|alex.menendez@gmail.com|Miami|1
one|two||four|five

Note: "Alexandra """"Alex"""""", blabla returns "Alexandra ""Alex""

```python
"""

def preprocess(input):
    # input row
    # out: list of column strings
    c = 0
    cols = []
    col_str = ""
    while c < len(input):
        if input[c] == ',':
            cols.append(col_str)
            col_str = ""
            c += 1
        if input[c] == '"':
            c, col_str = preprocess_special(input, c)
            cols.append(col_str)
            col_str = ""
            c += 1
        else:
            col_str += input[c]
            c += 1
    if col_str != "":
        cols.append(col_str)
    return cols

def preprocess_special(input, c):
    num_quotes = 0
    col_str = ""
```

```
    while c < len(input):
        if input[c] == '"':
            c += 1
            num_quotes += 1

            if c+1 < len(input) and input[c+1] == '"':
                c += 1
                num_quotes += 1
                col_str += input[c]

            if c+1 < len(input) and input[c+1] == ',':
                c += 1

        if input[c] == ',':
            if num_quotes%2 == 0:
                return c, col_str
        col_str += input[c]
        c += 1
    return c, col_str
test1 = 'John,Smith,john.smith@gmail.com,Los Angeles,1'
test2 = 'Jane,Roberts,janer@msn.com,"San Francisco, CA",0'
test3 = '"Alexandra """"Alex"""""", dfiasdfadf'
test4 = """
\"Alexandra \"\"Alex\"\"\",Menendez,alex.menendez@gmail.com,Miami,1
one,two,,four,"five"
"""
```

#print(preprocess(test1))
#print(preprocess_special('"San Francisco, CA"', 0))
print(preprocess(test1))
print(preprocess(test2))
print(preprocess(test3))

Amazon

Amazon Software Development Intern Interviews (tj)
45 minutes each
February 24, 2014

Interview 1: Say you're an airline company. You're given an array of passenger IDs and an array of seats. You want to map each passenger to a seat randomly. Write a function mapSeats(int[] pass, int[] seats) that assigns each passenger to a random seat.

You're told:
- int[] pass := an array of passenger IDs. Each ID is greater than 0.

- int[] seats := an array of seats. Each seat is initially set to zero.
- rand(start, end) := a function that returns a random integer in the range [start, end)
- pass.length <= seats.length

Ans: create an array of length seats called mapRand[], fill it with the numbers 0 to n-1, randomize it using the fischer-yates shuffle, then map

```
for i in range(seats.length):
    seats[i] = passengers[mapRand[i]]
```

Followup: What if we wanted to be able to lookup the passenger's seat afterward?
Ans: Save the randMap array for future use, or pass in a random seed and save that.

Interview 2:
Given an array of n integers (can be pos/neg), find the maximal subarray sum.

Ans: I gave the $O(n^2)$ time, $O(n^2)$ space DP solution but apparently you could do it in $O(n^2)$ (did not get) and then get that in O(n) in 5 lines. So I failed that interview.

Two on campus interviews, 45 min each
Feb 5, 2015 Interviews

Interview 1: Questions about yourself, what kind of role you're looking for, tell me about a project you worked on where you learned a lot and if given the chance, what would you redo. Pros and Cons between Relational vs NoSQL database.

Technical Question: Given a kindle, where you have books, which have unique integer ids, what data structure you use to model the following:

Kindle: Book4, Book2, Book1, Book3. If I click on Book1, book1 should now be on the top of my list and all other books are pushed over 1 position. Talk about pros and cons of each data structure. Now write code to implement a function movetoFront(int id)

Interview 2: Tell me about hashmaps and dictionaries. Runtime, when to use them, etc. Now talk about binary trees and BSTs. When would you use a HashMap over a BST? Vice versa

Given two int arrays, find the intersection of elements and return a set.
Now come up with an algorithm where run time is faster than O(N^2) and space is O(1), excluding the space of the set you will return.
Write a divide function without using the / operator. divide(int x, int y)

Phone Interview for Software Development Internship (2/16)
1. Return the last element before null in a linked list
2. Return the kth to last element in a linked list
3. If you were a QA on Q2, what are some edge cases you would think about?

Amazon Online Assessment
30 minutes on the internet
March 2, 2015

They added this online assessment thing where you complete 7 programming questions within 21 minutes on their site. Your webcam is turned on the whole time and you have to take a picture of yourself next to your student ID to make sure that it is really you completing the online assessment. All of the questions are of the form "look at these 10-15 lines of code, and find the one character/variable name mistake". Since you don't know what you're looking for, doing the first few questions is hard, but after 10 mins you'll get the hang of it and it will be super easy.

Note: Make sure you pick Java as your preferred testing language. The code is in Java 7, but they ask C-like questions so it doesn't matter.

I didn't remember all the seven questions, but here are several of the topics I had to debug
1. The warm-up question was "compute GCD(a,b) iteratively"
2. Selection sort
3. I think there was bubble sort/insertion sort.
   3.1. The bug was you should correct j = i to j = i+1. I think. Or it was improperly saving the max value seen so far.
4. Remove duplicate values in an array
5. Given an integer n, print "11" repeated n times. For example, if n=5 print:
   5.1. ["11", "1111", "11111", "11111111", "1111111111"]
6. Given an integer n, print the alphabet as follows. If n=5 then print:
   6.1. ["a", "ab", "abc", "abcd", "abcde"]
   6.2. They do the thing where you convert integers to characters so this probably was a C-string question
   6.3. (Note: the solution for this one is to correct the variable names in the second for loop)
7. For the last question, I forgot the question but they have a 'System.out.println()' statement as the last line and you need to correct that to 'System.out.println("")' otherwise it fails.

General advice: All the fixes are single-line minor fixes, like changing ">" to "<", or renaming a variable. You can compile your code every 15 seconds, so what I suggest is when you see the question immediately compile it and check the output; usually you can see what is wrong right away. I used guess and check a lot. Also consider skipping the first 2 questions since they are harder than the rest.

Mike and I both finished the test with 10 minutes left. Just relax! It's not too hard!

Amazon Phone Interview (tj)
45 Minutes
March 20, 2015

Talked with an Indian dude on the Amazon Product recommendation email team.
Q1: Talk about a fun project you've worked on.
Q2: Reverse a string.
Q3: Determine whether a binary tree is a binary search tree.
Q4: Do you have any questions for me?

Apple
1 Hour Phone Interview
March 16, 2015

Every semester, Apple holds a "Networking Event" where they have leads from 20+ teams standing around taking resumes. Each person is from an independent team so you need to bring a dozen copies of your resume and individually talk to each of the leads and have them remember you. I spent 20 mins speaking with one lead from the Developer Tools team (Jon Hess) and a month later he reached out to me. tl;dr of this part is if you want an interview, the lead needs to remember you otherwise you won't make it through the mess of recruiters. Also put you name in the iPad things they pass around otherwise they take forever to get back to you.

The phone interview is really intense, with lots of rapid fire questions. First part was talk about your projects on your resume, what projects are you working on right now, and what were your favorite classes. I brought up program memory layout from CS164 (compilers), and files and virtual memory from CS162 (operating systems).

1) In Unix, what is a file? Name a few examples.
2) Why is it a good thing that everything is a file?
3) Let's say that I have a socket. What's happening internally that allow me to treat this as a file?
4) Describe the memory layout of a process. (short answer: static, stack, heap)
5) What's the point of the static section?
6) What's the point of having a stack? Why not just put everything in the static section?
7) What's the point of the heap? Why not put everything in the stack?
8) What are the benefits of putting stuff in the stack?
9) Let's say I have a variable X. As a programmer, under what situations does X go in the static section? stack section? heap section? Compare and contrast them.
10) What is virtual memory? Why is it useful?
11) How does virtual memory add security?
12) What is the protected bit?
13) Explain how an OS context switch occurs.
14) What is malloc?
15) How might you implement malloc? What issues might you run into?
16) What differentiates a "good" implementation of malloc compared to a "bad" implementation of malloc? What factors would you consider for good and bad?

17) How might you optimize malloc?
18) Do you like data structures?
19) Let's say we lost the source code to ArrayList, and we want to rewrite it. What API would you choose, and what are the runtimes of each? (answer: init(default_size), get(index), insert(index), remove(index), size())
20) Let's say we insert at the end of an ArrayList. What happens? Runtime?
21) Let's say we insert at the beginning of an ArrayList. What happens? Runtime?
22) Let's say we remove an element from the ArrayList. What happens? Runtime?
23) Now do the same thing for a LinkedList. What API would you pick, and what is the runtime of each method? (answer: init(), getHead(), getTail(), insertFront(Object), insertTail(Object), insertAfter(Object), insertBefore(Object), removeHead(), removeTail(), remove(Object), remove(index), size())
24) What happens when we insert an element, remove an element, and check the size()? Analyze the runtimes of each.
25) Given a hashmap, how might you implement a set? What API would you pick, and what is the runtime of each method? (answer: init(), add(Object), contains(Object), remove(Object), size())
26) How might you implement a hashmap? What data structures would you use? What API would you pick, and what is the runtime of each method? (answer: init(), put(key, value), get(key), size(). Runtime is amortized O(1) for all).
27) What is a load factor of a hashmap?

Then we talked about Xcode, the structure of the Apple Developer Tools team, how Xcode is implemented, and what goes on at Apple.

March 26, 2015
30 Minute Phone Interview

I talked with Tony Ricciardi on the interface builder team. The interview had a ton of overlap with the previous one, so there's poor coordination between interviewers. Or maybe they're screening for the same type of thing.
1) What's your preferred programming language? (I answered Python and Java)
2) What's a weakness of this programming language?
3) Describe the memory layout of these languages.
4) What is reference counting?
5) What are some issues with reference counting? (Ans: reference cycles)
6) What are the benefits of using an Array? What about a LinkedList? Set? HashMap?
7) How might you maintain an ordering in a set? (Ans: treeset)
8) What is the runtime for operations on this structure? (Ans: it's a tree, so O(logn)
9) Give an example of a self-balancing tree. (Ans: red-black tree, splay trees)
10) What do you know about multithreading? (Ans: locks, semaphores, deadlocks, test-and-set, spinlocks, memory layout)
11) Explain what is a deadlock.
12) What's a good way to prevent deadlock? (Ans: careful programming, testing)
13) Better answer? (Ans: have acquire all resources before you need them)
14) Even better answer? (Ans: solution to dining philosophers problem)
15) If you had to answer this question again how would you do it? (Ans: google it)

This interview was kinda derp. He was asking random questions, don't think he had much interviewing experience. Anyway we chatted about what he's been working on, apparently he solo wrote the entire xcode

interface builder for the Apple Watch. It was super cool. But this project was so secret he couldn't bring it up with anyone else on his own team so wtf. He also does talks at WWDC which is pretty awesome.

Apple Onsite
March 31, 2015
Seven 45-minute interviews

Scheduled from 9:45am to 4:15
First met with the Developer Tools recruiter Maria Lippert. We talked about the my other offers, and the interview schedule for the coming day. The interview schedule was a series of two 45-minute pair interviews, followed by lunch, then 5 more 45-minute pair interviews. Each interview was "themed", meaning each pair of engineers was given a topic to discuss and ask questions about. The first interview started at 10am. The day was super long so I may have forgot all questions asked.

First interview (10am) was with Manager and team co-worker. I believe the theme was Operating Systems.

1) Let's say I have a file called `hello.c` that consists of the following:
   ```
   #include <stdio.h>
   int main() {
         printf("Hello world\n");
         return 0;
   }
   ```
   What happens when I type `cc hello.c;  ./a.out` and press enter?
   A: First, the terminal will examine your path/shell configuration and search what is "cc". My interviewer confirmed "cc" is an alias for your shell's preferred C compiler. Your shell will traverse the alias and execute "clang" with the argument "test.c" with the permissions of the current user. The C compiler starts with C preprocessor which expands all DEFINEs etc. Then it lexes (tokenizes) and parses (builds the syntax tree) for the input file and also does header stuff. Then it traverses the tree, annotates it, and verifies its syntactic correctness. It then runs code generation on each of the nodes and outputs static single assignment (SSA) LLVM intermediate representation (IR). This is then optimized by the LLVM backend. The optimized IR is then mapped to registers with an architecture specific graph coloring algorithm then linked with the linker. Then the program is loaded into memory and run.

2) I think my interviewer noticed I didn't know a ton about the linking/runtime step, so he asked follow-up questions on this.
   Q: "printf" is library function. Where does function reside in memory?

3) Describe the parts of a process memory layout. Where are functions stored?

4) What's a dynamic library? What's a static library? Difference?

5) Can I just use a static version of the "printf" function? Why or why not?

6) What are the benefits of using dynamic libraries?

7) Say I have a process with multiple threads. Talk a little about this process's address space layout.

8) Let's say I have Chrome, and Safari. Both share some dynamic libraries. How does this work?
   (A: dynamic libs are in shared memory)

9) What are shared libraries? How are they stored in memory? How might you call a method from one?

10) How does the shared library "know" what memory space to execute in? What permissions does it run with?

11) What's preventing me from calling arbitrary kernel functions?

12) As an OS designer, how might I prevent this from happening?

13) Describe ASLR. What is KASLR?
14) Give an example of an attack that ASLR prevents.
15) What is a buffer overflow?
16) Give an example how a buffer overflow might be exploited.
17) What is virtual memory?
18) List a few benefits of virtual memory. (I gave 3)
19) What is a swap file?
20) Q: A thread tries to read an address. Walk through what the OS does.
   A: The OS takes the virtual address, looks through the page table, and finds the physical page this maps to. These steps are assisted by the MMU and cached by the TLB. If the page is in memory, return that address. Otherwise, a page fault occurs and the thread is halted while the page is loaded in from disc. If physical memory is full, one is evicted (chosen semi-LRU), and another is loaded back in.
21) Talk about dirty and clean pages.
Impressions: Ran out of time

Second Interview. The topic was UI application design and MVC. Interviewers were one engineer, and one manager.
1) Q: Interviewer drew a mock layout of iTunes on the board. It was a single window with 3 dividing sections; top play bar, left menu bar, and a right songlist. He said it was called ${yourname}Tunes and asked how you would implement this UI layout.
   A: Using relativelayout (or whatever corresponding item), have a top play bar, with the left/right menu/songlist both be scrollviews.
2) On the right it is a scrollview where you can select songs. Question: if I select a song and click the "delete song" button, what happens?
   A: OS would send a clickEvent to the outer window, and based on the view hierarchy the right song list view would receive the event. This would be passed to the controller which updates the model.
3) What is MVC?
4) What are some of the benefits of MVC?
   A: Helps divide the application into distinct components. Developer can choose a different View with minor changes to the Controller and no changes to the Model. Also, can replace the backend Model with minor changes to the Controller. User will not notice a thing.
5) Let's say there are some menu bars that let users sort by Artist Name/Length/Play count, or choose what columns to see (constraints were very open ended). Using MVC, how might you design this "delete-song" scrollview?
   View: Modified scrollview. When user presses delete, width of the row for this song slowly decreases to 0.
   Controller: Concurrently, sends message to Model to delete this song
   Model: Contains two NSArrays: one reflecting what the user is actually seeing, and another for the direct structure on disc
6) How would you save this data structure to disc?
   Do what iTunes does: folder for each song; each song contains mp3 file and a metadata file
7) What might you store in the metadata file?
   plist of ArtistName (str), playcount (int), path to mp3 (str), etc
8) Let's say the user quits while a song is deleting. What happens?
   song may not be deleted, will reappear next time they open

9) How can we fix this?
   Have a finalizer method that waits for the song to be deleted before closing ${yourname}Tunes
10) What if the user did a force-quit?
    not too sure, but I suggested when you delete a song just adding a flag to the metadata saying "delete when ${yourname}Tunes is opened again". Then try to remove it, this should be safe over force quits. They seemed OK with it
11) How might you store the full list of songs?
    I suggested an array of structs. He prodded more. I suggested an array of hashes. More prodding. I suggested an array of Media objects, where Media could be a Video or Song object.
12) How might you quickly find an item in this array?
    Assign each song a UDID and the do binary search to find it. Also suggested a hashmap but we ran out of time here.

Impressions: Apparently they always ask this question. Really glad I took CS160 UI/UX development.

Lunch with Jon Hess (manager) and Chris Lattner (director of developer tools, founder of LLVM)
Ate at the cafeteria. We talked about engineering culture at Apple. Only found out who Chris was afterwards, he was literally just like another engineer. If I knew I would have asked more LLVM questions.

Third Interview: Reference counting and method dispatch
1) Describe the memory model for a C based language
2) What do you know about reference counting? Give an example.
   (have each heap object have a uint field of how many live items reference it. increment per reference, decrement per release. when count == 0 invoke destructor.)
3) How would you do this transparently? That is, implement reference counting without modifying the Object.
   (add a RC object in between the pointer)
4) What fields does your RC object have?
   count and pointer to reference counted object
5) What is this known as in C++?
   std::shared_pointer
6) Let's say I have a parent object (i.e. a Window) with 4 children objects (i.e. 4 Buttons). Draw out the object layout.
7) What is a reference cycle?
8) How can we avoid this?
   @strong and @weak pointers
9) How might you implement a weak pointer?
   (tbh i'm not too sure, but I think when weak_ptr count goes to 0, just delete the weak_ptr Object but do not delete the Object being pointed to)
10) What concurrency issues can you run into?
    two buttons try to decrement shared_pointer count at the same time which incorrectly updates the count value. basic concurrency read/write thing
11) How to avoid this?
    I suggested adding a lock for each shared pointer. Interviewer said there was a better way but due to time he just skipped it
12) New question: Implement dynamic method dispatch for a C like language. Include support for method creation and modification *during runtime*.

This question is seriously hard. The first part is virtual tables, but I had no idea for the second part. Ran out of time for the method dispatch so he just told me the answer: hashtables are your friend

Impressions: A lot of Objective C questions. Looking at glassdoor, implement reference counting seems to be a common question. Read *More Effective C++* items 28 and 29.

Fourth Interview: Programming languages. Two older guys, one from the language design team, they seemed really relaxed and not too into it.
1) What is your preferred programming language?
2) Give a weakness of the language you mentioned.
3) What is static single assignment?
4) What is an interface? Have you used interfaces before? (yes). Tell us about an instance where you did so...
5) Show (i.e. draw) an example of interface hierarchy.
6) What's the difference between a function and a method?
7) What's the difference between a lambda and an anonymous function?
8) What is the scoping of a lambda function?

Thoughts: I didn't know the last three answers but they were really happy to explain the intricacies and details.

Fifth Interview: Debugging. Someone from lldb team was there.
1) Say you have a program you wrote, for example, `int add(int x, int y) { return x + y; }` that's giving the wrong output. How would debug it?
2) What is an exception backtrace?
3) How is the backtrace generated?
   When an exception occurs, control is passed over to the exception handler. The memory of the crashed stack frame is examined and debug information generated. To access the caller's stack frames, look at the return address and jump there, and generate debug info. Repeatedly the process.
4) You mentioned gdb. What process do you normally use debug with gdb?
   recompile with -g. then run in gdb. Find the line where it crashed. Then add a breakpoint just before that line. When there do 'info locals' and 'p vars', and use 'n' or 's' until it crashes.
5) How would you debug customer crash reports? (open ended answer)
6) Say we are getting crash reports for some of the later builds but we are not seeing these crash reports for earlier builds. How do we debug this?
   examine commits using `git bisect`
7) Have you ever debugged multithreaded programs? (no)
8) Let's say a user is reporting the application to be jittery/sluggish. How to debug?
   xcode instruments cpu profiler.
9) User is getting a lot of application OOM errors. How to debug?
   xcode instruments memory profiler
10) What is a memory leak?
11) How to detect memory leaks?
    xcode instruments memory profiler

Sixth Interview: Multithreading. Interviewers were super relaxed. Easiest interview.

14

1) Have you wrote multithreaded programs before? (yes). Talk about it...
   A: I brought up the matrix multiplication with OpenMP
2) In MT programs, how to avoid cache conflicts?
3) What is deadlock? Describe/draw an example on the whiteboard.
4) What are some techniques you can use to prevent deadlock?
5) You have a very large job that can be parallelizable. How might you do this?
   Master thread creates a threadpool and slaves receive jobs from master.
6) What are some issues that can arise from using a threadpool?
   too many threads causes context switching overhead, uneven job sizes can lead to lag/waiting
7) What is a lock? Talk about the different kinds of locks (spinlocks, sleeplocks, cond_wait, test and set)
8) What is a semaphore?
9) (implement a lock using a semaphore)
10) What is priority inversion? give an example. How to fix?

Seventh Interview: Design a text editor for huge files
Q: You're writing a hex text editor and users complain the editor is very slow when dealing with very large files (100MB to 1GB+). How might you optimize this?
A: Think git. Divide the input file into blocks (4KB each) and when the user edits a block just save the diff.

Overall impressions: All design questions. ZERO coding.

AOL

1. What is the difference between Get and Post in HTML?

2. What is the difference between Left Join and Right Join in SQL?

3. If you were given a UI project, how would you go about doing it?

4. If I was a manager of a store that sold tablets and I wanted to hire you as a consultant to assess the health of my store, what metrics would you use?

Argo

2-day coding challenge (Semantic Analysis)
April 3-4, 2015

Your objective for this challenge is to design and implement a system that randomly generates sensible output by generalizing patterns found in an input text. Your system should be able to produce unique sentences based on a model that you design using the input text. Such a system could be useful for randomly generating poetry, song lyrics, or cryptic academic papers.

Your solution should:

- Analyze the source text to build a model for your randomly generated sentences.
- Randomly generate unique sentences of user-defined length.

You're welcome to implement your random sentence generator in any language that you like, provided that it provides a sensible API for testing and usage. Be sure to carefully choose your data structures for storing information about your model, and clearly explain your approach.

To test your solution, you're welcome to use any text you like, from Shakespeare's Hamlet to Martin Luther King's I Have a Dream Speech.

Solution: The general idea of my solution was to generate a dictionary of all possible n-grams (for any specified n) and basically chain n-grams in a random order (weighted by their probability) to create grammatically incorrect but somewhat cohesive sentences. Since the goal was to make this a usable API, I made a lot of tweak-able parameters, like n-gram overlap, the length of each gram, and the length of each sentence generated.

An important point was to make this fast, so use a dictionary to store the n-grams, and as many in-built Python functions (like zip and generator statements). The dictionary allows you to train on all of Hamlet in < 0.5 seconds.

Arista Networks
45 minute phone interview
March 6, 2014

Told to ssh into a server and edit some C++ code on their server with vim. They used the screen command to share the terminal session. Was given some skeleton code in a file called "tree.cpp". Here was the question:

Given a binary search tree, implement the functions `find_min(node*)` and `find_next(node*)` so that `main()` prints the inorder traversal of the tree.

Skeleton code:
```
#include <stdio.h>

struct node {
      int value;
      node* left;
      node* right;
      node* parent;
};

int main() {
      node* root = construct_bst();
      for (node* n = find_min(root); n; n = find_next(n)) {
            printf("%d\n", n->value);
      }
      return 0;
}
```

Code that I wrote:

```
node* find_min(node* root) {
   if (!root) {
      return NULL;
   }

   node* currNode = root;
   while (currNode->left != NULL) {
      currNode = currNode->left;
   }
   return currNode;
}

node* find_next(node* treenode){
   if (!treenode) {
      return NULL;
   }

   node* tmpnode = treenode;
   if (treenode->right) {
      tmpnode = treenode->right;
      return find_min(tmpnode);
   } else {
         //no right subtree
      tmpnode = treenode -> parent;
      while (tmpnode && tmpnode->value < treenode->value) {
         tmpnode = tmpnode -> parent;
      }
      return tmpnode;
   }
}
```

Followup Questions:
1) What's the runtime of this code?
In the worst case (a tree with only left children), `find_min()` can take O(n). For `find_next()`, in total we visit each edge at most twice, and since a tree has n-1 edges, `find_next()` looks at at most 2n nodes, so this takes O(n) runtime.

2) Look at the line `tmpnode->value < treenode->value`. Is there any way to do this line without using the comparison operator?
Yes, stop when we find out that `tmpnode` is in the left subtree of `tmpnode->parent`.

Bank of America Merrill Lynch

1. What's the difference between a Hashmap and a treemap?

2. What's the difference between a hashmap and a linked hash map?

3. What is a thread? What does synchronization mean?

4. Implement a singleton for me

5. Say you are given a sentence as a string input. Sort the sentence and it is case sensitive.

Benchling

Phone screen with engineer. Note that the phone screen is a Google Hangouts *video* chat, and hangouts video sucks. I had a hard time getting it to work even when I was doing a video interview with *Google*. They jump straight into an easy coding question after explaining what the company does and answering your questions. The only HR question they ask is what area you're interested in working on. Technically everyone is full stack but you can focus on backend or frontend. FYI they use Flask and coffeescript/otherscriptIdontremember.

Question
Background info (explained to you): If you're not sure what exactly the base is, you can use these codes to indicate possibilities. Write a function to generate all possible real sequences from an input with ambiguous bases.

```python
IUPAC_TO_BASES = {
        'A': 'A',
        'C': 'C',
        'G': 'G',
        'T': 'T',
        'R': 'AG',
        'Y': 'CT',
        'M': 'AC',
        'K': 'GT',
        'W': 'AT',
        'S': 'CG',
        'B': 'CGT',
        'D': 'AGT',
        'H': 'ACT',
        'V': 'ACG',
        'N': 'ACGT'
}

def generate_real_dna(degen_bases):
        possible_dna = ['']

        for c in degen_bases:
        bases = IUPAC_TO_BASES[c]
```

```
        sequences = []
        for b in bases:
              for s in possible_dna:
                    sequences.append(s+b)
        possible_dna = sequences

        return possible_dna

print '>', generate_real_dna('RY')  # ['AC', 'AT', 'GC', 'GT']
# tests I wrote
print '>', generate_real_dna('RYR')

print '>', generate_real_dna('N') # ['a', 'c','t','g]'

print '>', generate_real_dna('')
```

Technical Phone Screens
2 video chats where you solve a problem on coderpad.io. The first guy asked 1 question and the second guy asked 2 questions (went 20 minutes over time... oops). You're allowed to import/look up small helper functions, like checking that you can uses list.pop to pop from a given index.

1. given k sorted lists, each has size n => merge them into one sorted list
What's the running time? Can you do better? What's the running time now? (naive solution is $O(k^2*n)$) and the mergesort solution is $O(k*log(k)*n)$. CS170 YO.

2. Write the function eval_str so that eval_str('1 + 3 / 2 - 6 * 2') returns -9.5. Basically just evaluate the string and return the float value. Multiplication and division group closer than add/sub. You can use operatior.mul and its friends. I'm pretty sure there's a way to do this efficiently with stacks (would be easier with polish notation...) but I couldn't figure it out so I just implemented the "loop over everything to calculate all the / and * then go L -> R and aggregate the + and -.

```
# 1 + 3 / 2 - 6 * 2
 import operator
 def eval_str(expr):
     ops = {'*': operator.mul,
            '/': operator.div,
            '+': operator.add,
            '-': operator.sub}

     parsed = expr.split()
     values = []
     operations = []
     for i in range(len(parsed)):
     if i%2==0:
          values.append(float(parsed[i]))
     else:
          operations.append(parsed[i])
     print "values", values
```

```python
        print "operations", operations

        while '/' in operations or '*' in operations:
            div = len(operations) + 1
            mul = len(operations) + 1
            if '/' in operations:
                div = operations.index('/')
            elif '*' in operations:
                mul = operations.index('*')
            index = min(div,mul)
            func = ops[operations[index]]
            print "applying ", func
            print "to ", values[index], values[index+1]
            result = func(values[index], values[index+1])
            print "result is ", result
            values = values[:index] + [result] + values[index+2:]
            operations.pop(index)
            print "new values", values
            print "new operations", operations

        ans = values[0]
        val_index = 1
        for op in operations:
            ans = ops[op](ans, values[val_index])
            val_index += 1

        return ans


print eval_str('1 + 3 / 2 - 6 * 2')   # -9.5
print eval_str('2 * 3 * 4')
print eval_str('2 + 3 + 4')
print eval_str('2 * -3 + 1 + 1')
print eval_str('12 / 3 / 2')
print '-'*80
print eval_str('12 + 2 - 5 - 5')
```

3. Given a 2D array, print the items in a spiral.
```python
print_spiral([[ 1,  2,  3,  4],
           [ 5,  6,  7,  8],
           [ 9, 10, 11, 12],
           [13, 14, 15, 16]])
```
E.g. the above prints in the order of 1 2 3 4 8 12 16 15 14 13 9 10 11

```python
from math import ceil
def print_spiral(rect):
    bottom = len(rect) - 1 # rows
```

```python
    right = len(rect[0]) - 1 # cols
    left = 0
    top = 0

    x = 0
    y = 0
    num_elements = len(rect) * len(rect[0])
    index = 0
    dir = 'R'
    while index < num_elements:
    print rect[y][x]
    if dir == 'R':
        if x == right:
            top += 1
            y += 1
            dir = 'D'
        else:
            x += 1
    elif dir == 'L':
        if x == left:
            bottom -= 1
            y -= 1
            dir = 'U'
        else:
            x -= 1
    elif dir == 'U':
        if y == top:
            left += 1
            x += 1
            dir = 'R'
        else:
            y -= 1
    elif dir == 'D':
        if y == bottom:
            right -= 1
            x -= 1
            dir = 'L'
        else:
            y += 1
    index += 1

print '\n1st Test Case'
print_spiral([[ 1,  2,  3,  4],
        [ 5,  6,  7,  8],
        [ 9, 10, 11, 12],
        [13, 14, 15, 16]])
print '\n2nd Test Case'
```

```
print_spiral([[ 1,  2,  3,  4,  5,  6],
           [ 7,  8,  9, 10, 11, 12],
           [13, 14, 15, 16, 17, 18]])
print '\n3rd Test Case'
print_spiral([[ 1,  2,  3],
           [ 4,  5,  6],
           [ 7,  8,  9],
           [10, 11, 12],
           [13, 14, 15]])
print '\n4th Test Case'
print_spiral([[ 1,  2],
           [ 3,  4],
           [ 5,  6],
           [ 7,  8],
           [ 9, 10]])
print '\n5th Test Case'
print_spiral([[ 1,  2,  3,  4,  5,  6],
           [ 7,  8,  9, 10, 11, 12]])
print '\n6th Test Case'
print_spiral([[5]])
```

Blackboard

30 minute phone interview
March 25, 2014

1) Does Java support multiple inheritance? How do you get around this?

2) Describe interfaces vs abstract classes. Which is preferred?

3) Describe the rules for method overloading. Which methods can be overloaded? Also, how would you prevent a method from being overridden?

4) My interviewer emailed me an XML file and told me to open it up and describe how it would be implemented in an object. Here's the xml file:

```
<alert xmlns="urn:oasis:names:tc:emergency:cap:1.2" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <identifier>ce591532-2080-4ac1-bf52-faca52e994fa</identifier>
  <sender>145983:424335</sender>
  <sent>2013-02-15T01:17:43+00:00</sent>
  <status>Actual</status>
  <msgType>Alert</msgType>
  <scope>Public</scope>
```

```
  <info>
    <category>Other</category>
    <event>OutReach</event>
    <responseType>None</responseType>
    <urgency>Unknown</urgency>
    <severity>Unknown</severity>
    <certainty>Unknown</certainty>
    <expires>2013-02-15T04:17:43+00:00</expires>
    <headline>145983:424335:7818:Ilana Map 3</headline>
    <description>145983:424335:7818:Ilana Map 3</description>
    <area>
      <areaDesc>1660; </areaDesc>
      <circle>34.160447093000073,-118.46847958399991 8.046721</circle>
    </area>
    <area>
      <areaDesc>1661; </areaDesc>
      <circle>24.345345345345,-120.23423423423423 8.046721</circle>
    </area>
    <area>
      <areaDesc>1571; </areaDesc>
      <polygon>34.244708321000076,-118.41835446199991 34.224555760000044,-118.59962887599994
        34.203830493000055,-118.3527798159999 34.244708321000076,-118.41835446199991</polygon>
    </area>
  </info>
</alert>
```

Ans: Create a Class called Alert and add many fields to it. Use a String called identifier to store the unique ID, use a String called sender to store the sender, use a String called Status

Blackrock

1st Round

1. Tell me about what you know about Blackrock, why you're interested in working in an investment management firm, and yourself.

2. What was the hardest project you had to work on? What was the hardest bug?

3. Can you tell me what you have learned while working in a software engineering team?

Final Round

1. I have this string "xxyyyzzxxxxyyyzz." I want it to return the longest continuous substring, in which case is "xxxx." Describe on a high level how you would do this.

2. What's the difference between a hashmap and a hashtable? How does a hashmap work?

3. What's the difference between a stack and a heap? (Not the data structures)

4. Describe how the Java garbage collector works.

5. Explain to me how linked lists work conceptually and the benefits/cons of it compared to arrays.

6. What's the difference between abstract vs interface? What's another term for data structures?

7. Write a binary search for me

8. What interface does the linked list come from?

Blend Labs

1. How do you check that your code has balanced brackets, parentheses, etc.? Like a compiler does…
2. (Margaret) They had a slightly strange google doc questionnaire that had some easy-peasy code questions and things like "Imagine you've been cursed so that you can choose only 3 data structures to use for the rest of your life. Which 3 would you pick?". There was a miscommunication/fuckup somewhere (I choose to blame their HR) and the engineer that was supposed to call me missed the phone screen. When that got around to happening he was quite friendly and seemed to know what he was talking about. Write a function: "Input: an infinite stream of integers. You want to go through the whole stream and keep track of the top five largest. Now do it to keep track of the top k integers." Answer: Iterate through the stream and keep a size k min heap. Since each time you're comparing the current element with the kth largest, this is an O(1) operation instead of potentially O(k) if you had to check the entire data structure to get the smallest-largest each time.

Bloomberg

1a) Say you have a singly linked list. How do you find if it's a cycle? *Use a tortoise and hare pointer - the tortoise iterates by one and the hare iterates by two. If they are ever at the same position, return true. Otherwise if the hare reaches the end, return false.*
1b) Say you have two singly linked lists and you have two pointers already initialized at the beginning of each. Find if there are ever two nodes that are the same/connected, and return the first instance. Must be O(n) time. *He gave a hint to make the two linked lists the same length, and another hint to not use any complex data structures. I had no fucking idea. (I deem this question impossible - nobody on the entire fucking internet can do it).*

HKN Answer 1: Have both pointers advance until the very end of the list. When you are at the last node, check the data. If they are the same, they are connected. If not, different lists.

HKN Extra Harder Problem: One, Two, or none of the lists may have cycles. How do you check if they are ever connected.
Answer: Run hare and tortoise. And now here are the cases and answers.
- If none of them have cycles, run the algorithm from Answer 1.
- If one has a cycle and they are disconnected, one will run into null and one will go into a cycle. Thus, not the same
- If one has a cycle and they are connected, the hare and tortoise of each will eventually catch up and meet; thus same.
- If both have cycles and not connected, have the pointers in one of the list stop. If the pointer of the second list ever has the same data as the stopped node, same cycle.
- If both have cycles and are connected, then it's eventually the same list

2a) Say you have a list of N numbers. Find the largest and 2nd largest number and return them. *Create two variables, and initialize one to the first element and another to the second. Iterate through the list of N numbers and compare to each of the variables, replacing the variables as needed. O(n) time.*
2b) Say you the list of N numbers is ridiculously large, so that you can't represent it with data structures such as tree, hashmap, queue, etc. You want to find the K largest numbers where K << N. Do this in better than O(n + k log k) time. *I gave the O(n + k log k) solution, which he was unhappy with. Also, as I told you, he said don't use sorting on the K thing, but represent it with some data structure that isn't a list, graph, queue. wtffff*

http://stackoverflow.com/questions/7423401/whats-the-difference-between-the-data-structure-tree-and-graph

HKN Answer: Use quick select to find the Kth largest number. O(N)
Then, find all #'s greater than K, O(N)
O(N) + O(N) = O(N)

3) Didn't reach this point.

4) Say you have a list of numbers. I want to put these numbers into two buckets, one for even and one for odd. In addition, I want to see how many times a number occurs in the list. How would you go about doing this?

5) I have 8 metal balls. 7 are equal weight, one is lighter than the rest. I have a scale to see if two sides balance. What is the optimal strategy to find the metal ball?

6) What is JVM and if I transfer one Java program to another computer, will it work?

7) Java isn't the most used language here at Bloomberg. Would you be interested in learning C or

C++? Saying no to this question will most likely result in a rejection…

8. What is polymorphism and give an example (in theory and code). How would it work in C++ instead of Java?
9. You are given 2 arrays: one which stores the revenue the company made each day and one which stores the costs the company incurred each day. You want to find the maximum profit the company could potentially earn. However, the maximum profit cannot be from revenue and costs on the same day.

10. Given a string such as "A8HG2ZI", sort the string while keeping the structure of the numbers and letters intact (the number indexes must always have numbers and character indexes must always be characters)

Bloomberg Full-Time Interview, First Round (tj)
1 hour on-campus interview.
February 20, 2015

Interviewer was an Indian guy working on a stock analytics product for Bloomberg terminal. He really, really liked data structures and whenever I brought up a data structure he asked me to describe how it works internally and what were the runtimes. Interview language was either in C, C++, Java, or Python but since he asked C library functions I wrote them in C.

Q1: Implement the C library function `int strcspn(char *str1, char *str2)` (look it up). Basically, it returns the first n continuous characters in str1 that do not appear in str2.
A1: I asked him if I could restrict the input to ASCII-256 and he said yes, so this question is straightforward. He asked a lot of questions about my code, for example, converting chars to values and back.

```c
int strcspn(const char *str1, const char *str2) {
        assert(str1 && str2);
        char seenbank[256] = {0};
        for(char* ptr2 = str2; *ptr2 != '\0'; ptr2++) {
                seenbank[*ptr2] = 1;
        }

        int result = 0;
        for(char* ptr1 = str1; *ptr1 != '\0'; ptr1++) {
                if (!seenbank[*ptr1]) {
                        result++;
                }
                break;
        }
        return result;
}
```
Q2: Modify the function above so that it instead returns a C string of all characters that are found in

both str1 and str2.

A2: I reused the first half of the code above. One important thing is we need to malloc the resulting array since (1) a stack array like above will be destroyed when the frame exits and (2) resulting string could be variable sized.

```c
char *overlap(const char *str1, int strlen1, const char *str2, int strlen2) {
        assert(str1 && str2 && strlen1 >= 0 && strlen2 >= 0);
        char seenbank[256] = {0};
        for(char* ptr2 = str2; *ptr2 != '\0'; ptr2++) {
                seenbank[*ptr2] = 1;
        }

        char* strResult = (char*) malloc(sizeof(char) * (MIN(strlen1, strlen2)+1));
        if (!strResult) {
                panic();
        }

        int resultIndex = 0;
        for(char* ptr1 = str1; *ptr1 != '\0'; ptr1++) {
                if (seenbank[*ptr1]) {
                        strResult[resultIndex++] = *ptr1;
                }
        }
        strResult[resultIndex] = '\0';
        return strResult;
}
```

Followup question: How does malloc work? What does it return and what arguments does it take? Takes a size_t and returns a void* pointer to somewhere on the heap. Returns NULL if fail.

Q3: Instead of returning the overlap of two strings, now compute the overlap of two *lists* of strings. For example, if strlist1 = ["IBM", "GOOG", "AAPL", "MSFT"] and strlist2 =["CSCO", "AAPL", "IBM"], then overlap(char **strlist1, char **strlist2) should return ["IBM", "APPL"].

A3: Didn't actually code this, but take everything in the smaller list and put it in a HashMap then iterate through the longer list and check the hashmap contents for each value.

Follow Up questions: Describe how you would implement a HashMap. What is the complexity of creating a hashtable? Memory complexity? What is the runtime of accessing an element? What is a load factor? How would you implement a hashcode function for a string?

(you should know all these lol)

Q4: You are working with a team of mathematicians and they are working with a lot of polynomials of the form $a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_0$. Design a Class and API that will help them manipulate polynomials.

A4: The API was basically add(Polynomial), subtract(Polynomial), multiply(Polynomial), and evaluate(Polynomial).

Q5: What data structure would you use to store the coefficients? Best to use a linked-list, or a Vector? How would you implement a linked-list? How would you implement a Vector? What are the pros and cons of using each?

A5: Ran out of time but which data structure you use depends on what the coefficients look like. If

they are very sparse you probably want to use a linkedlist or a hashtable, but if they are very full you would rather use a vector instead.

Bloomberg Full-Time Interview, Second Round (tj)

1 hour on-campus interview.

February 20, 2015

I had this russian guy who was the leader of the Machine Learning group at Bloomberg. Apparently C++11 is *huge* at Bloomberg so, make sure you know the newest and greatest features added in C++.

Q1: Describe the memory layout of C/C++. What are the differences between C and C++?

Q2: What's the difference between malloc and new?

Q3: Census question data structure question. I think he made it up on the spot.

Imagine you are given a well-formed list of 300 million census records, each with a firstname, lastname, age, sex, address, phone number, and ID. You are trying to design a service that allows fast, non-blocking read, update, and deletes to this record list and also provides very quick searches for the following queries: (1) "Lastname, Firstname" + full address (2) Phone number (3) full address. This question is intended to check your design skills and understanding data structures, so you may not use a premade SQLike database.

A3: I have no idea. Apparently they solve this type of question a lot at bloomberg because they need to quickly look up and present data to their users. Never taken a database course before so I got rekd on this question.

Box

All the questions are already on Glassdoor.

1. Design an elevator with OOP code. ← this question =(
2. Describe a HashMap in depth.
3. Describe BST in depth.
4. Why would you use BST instead of a HashMap?
5. Write a function to see if all the nodes in a tree have the same value.

Captricity

Given an array of numbers besides 0, return an array where the number at index i is a product of every other number.

Given an array of Strings, return a 2-d structure that contains all the anagrams grouped together.

Talk through how you would implement an Excel spreadsheet.

Given a String like aaabbbbbbcccc, return a3b6c4 or the original String, whichever is shorter.

Interview with Founder/CEO: they expected me to come in but the email said phone interview..

You have 1-10MB images, with anywhere from a few to 1000 of those images. There's a blackbox algorithm that takes anywhere from a few seconds to a minute to process these images and return a result, say if there's a UFO in these images. Design the backend processes for this if say you have a club of 10 members that are using your program. ...*this was really weird and I just said some shit from 162*

0th phone screen:
- talk with VP of engineering about what is captricity, resume, why are you interested in captricity, can you work in the US, etc.

1st technical phone screen with head engineer (position interviewed for was listed as "MLE" which I assume stands for Machine Learning Engineer... pretty standard interview except the verbal questions on basic ML training/testing pipeline stuff):
- (Interviewer graduated from Berkeley a few years ago) What were your favorite classes and projects?
- verbal / general knowledge questions:
  - What does it mean for a website to be responsive (provide examples)?
  - What is cross-validation?
  - What is k-fold cross validation?
  - What is the high-level goal of an SVM?
  - What makes an SVM different from a perceptron / what is "support vector margin"?
- Implement push, pop, and is_empty for a queue using only 2 stacks.
- Discuss algorithm for natural string sorting. Ex. ["bs1-1-1", "bs1-2-1", "bs1-10-1"] should be sorted like that, but naive string comparison puts "10" < "2". Implement the sub-function that parses substrings like "23bs1" into [23, "bs", 1].

Citi
When would you use inheritance vs interface?

Cisco-Meraki
Given a sorted, rotated array, find the minimum number.
Implement a stack, find min element in O(1) time.

CrunchyRoll
(I never did the challenge btw)

Your task:

Write a crawler script that 1) finds a target _Goal_ page given a starting page, and 2) analyzes some properties of the link topology (see the solution format).

You may use one of the following languages: PHP, Python, Java, C++, C, Ruby, Javascript. No third-party packages or external libraries allowed, with the exception of HTTP and JSON libraries if your language does not have native support. Do not use any built-in functions that evaluate expressions for you, like eval().

Your starting page:

Your custom starting URL is generated from this expression:

   abs(add(multiply(203,add(add(41,61172),252)),add(61,add(18637,14))))

To get you started, this expression has already been evaluated (to 12496107) and constructed into a URL for you (see below):

   http://www.crunchyroll.com/tech-challenge/roaming-math/mtran10@berkeley.edu/1249 6107

Page types:

There are three types of pages:

1. **List page. _List_ pages link to other pages. There can be multiple links on the page, one per line. Each line is an expression that corresponds to a url for another page. You must evaluate the expression, construct the url, and follow the link. The result of the evaluated expression will be an unsigned 64-bit integer. You can assume all expressions will be valid and not malformed. Possible operations are:**
   - **add(expr1,expr2) - takes exactly two operands and returns their sum.**
   - **subtract(expr1,expr2) - takes exactly two operands and returns their difference (expr1 - expr2).**
   - **multiply(expr1,expr2) - takes exactly two operands and returns their product.**
   - **abs(expr1) - takes exactly one operand and returns its absolute value.**
2. **The url is constructed as follows:**

3. http://www.crunchyroll.com/tech-challenge/roaming-math/mtran10@berkeley.edu/[evaluated_expression]

4.

5. where [evaluated_expression] is an unsigned 64-bit integer.

6. Deadend page. *Deadend* pages do not link to other pages. It will be a text/plain response with the single word: **DEADEND**.

7. Goal page. The *Goal* page is the page you are looking for. It does not link to other pages. It will be a text/plain response with the single word: **GOAL**.

Solution format:

Your solution is a json-encoded string of a dictionary with the following elements:

unsigned 64-bit int	goal	The integer corresponding to the Goal page.
unsigned int	node_count	The number of unique pages that must be visited to exhaustively explore all pages.
list of unsigned 64-bit ints	shortest_path	The ordered list of integers that correspond to the pages that must be visited to reach the goal page from the starting page. The first element in the list will be the integer corresponding to the starting page, and the last element of the list will be the integer corresponding to the goal page. If there are more than two paths to reach the goal page, the list must be the shortest one. If there are multiple shortest paths, any one of them is acceptable.
unsigned int	directed_cycle_count	The number of *simple* unique directed cycles, or infinite loops. A

		simple directed cycle is where a *list page* contains a link to another page that ultimately links back to the same page.

An example solution is:

{

 "goal":43990,

 "node_count":382,

 "shortest_path":[12496107,7691,44562,304,20938492,43990],

 "directed_cycle_count":5

}

Varying whitespace does not matter, but please make sure you use double-quotes on the key strings as indicated, and no quotes on the values.

Dell

1. What is a pointer? This eventually came to type p* = malloc(sizeof(type));

2. What is a binary tree?

3. Say we have a doubly linked list of ascending numbers. I want to convert the linked list into a Binary Search Tree. How would I do that?

4. What happens inside the computer the moment I press the power button?

DNAnexus

1) How do you get only one commit of a branch in git to be merged into another branch? I didn't know but it's git cherrypick. Then he asked which git commands I use. I named enough that he was satisfied.
2) Given a sorted array, how do you find out the index where a given integer belongs. Obviously binary search in O(log n)...
3) Some unix command thing. I had no idea. Then he asked what Unix commands I use on a daily basis. Thanks to 162, I named enough of them that he was satisfied
4) Give an example of how a deadlock might occur. I 162'd that shit.

Coding challenge: ask Brian for the answers

1. given an array of pages, which are arrays of words, create the index like the one in the back

of a book
2. given a bunch of trains, find out how many combinations you can create. each train consists of a number of cars, each with a letter inside. you can flip a train. the rule is that when you combine 2 trains, the car from one train touching another train must have the same letter. (this question is the hardest question I've ever gotten)
3. implement the square root function (use binary search)
4. talk through different implementations of a dynamic array. what if you add 10 elements at a time, 100, 1000, blah blah blah

DropBox

Online challenge. If you get it, I can send solution to you.

EA

Given a binary tree return text where each nth line represents all the node values in the nth level of the tree.

Just use BFS on the tree and make each node also contain a level variable. Every time you get adjacent nodes from a node you set the level to one plus the current level. And add "\n" accordingly to the String to be returned.

Software Engineering Intern (tj)
30 Minute "On-site" Interview
March 3, 2015

I really don't know what EA is doing. Their "On-site" interview was at the Cal Career center and they only gave 30 minutes for the whole interview, which is absolutely not enough to see how good a candidate is. Be there early! Interview started 5 mins early and then I had a 3 minute chat with the recruiter. She asked generic questions (when are you graduating, do you have any other timed offers, etc), but be prepared for the questions "What video games do you play?" and "Why do you want to work for EA?"

Then I went to another room and talked with Jordan on the Server Data Analytics team. He understood how little time there was so we skipped straight to the coding question.
Q1: Implement a LRU cache. It must implement get(key), and set(key, value) both in O(1) time.
A1: The two data-structures you need are a LinkedList and a HashMap. This question is pretty long and he let me skip a few parts but make sure you google it and understand how to do this.
Q2: Do you have any questions for me?
A2: ayy lamo

45 Minute "Technical" Interview (tj)

March 20, 2015

It was called a technical interview but it was completely nontechnical, just talked with the manager (David Lowe) for a bit. He didn't have much to ask so we ended 15 minutes early.

Q1: Do you have any questions for me?

A1: ayy

45 Minute Technical Interview (tj)

March 23, 2015

First actually technical interview. My interviewer was Gary James and we started out with some original Java questions.

1) On a scale from 1 to 10, how would you rate your proficiency in Java?
2) What is the difference between an abstract class and an interface in Java?
3) Does Java support multiple inheritance?
4) Is Java pass-by-reference or pass-by-value?

Then he pulled up a Google doc and asked this question. Given a String containing several words, reverse the order of the words in the string. For example, "the cat sat" => "cat sat the". Do not use advanced parsing features like split or regex.

Answer: Some edge cases to worry about, think about "sat   ", " the cat sat ", or "".

```java
public String reverseWords(String inp) {
    if (inp.length() == 0) {
        return "";
    }
    char[] strChars = inp.getCharArray();
    int i, j = strChars.length-1, -1;
    StringBuilder sb = new StringBuilder();

    while (i >= 0) {
        if (j == -1) {
            if (strChars[i].isChar()) {
                j = i;
            }
        } else if (strChars[i] == ' ') {
            sb.append( substr(i+1,j) );
            j = -1;
        }
        i--;
    }

    if (j != -1) {
        // we are in a word
        sb.append( substr(0, j) );
    }

    return sb.toString();
}
```

Runtime: O(n) with O(n) space. Interviewer said this is correct and more efficient but one easier

way is to use a stack; just keep push each word onto the stack and pop them off in order.

Enchanted Labs

1. How would you go about designing a program where given an input for number of nodes and edges, be able to generate and return a random graph?

Ericsson Mediaroom

they like to ask a bunch of random crap over the phone. just have all your cheatsheets printed out. everything is from cracking the coding interview

1. What happens if you type google.com in a brand new computer.
2. Reverse the order of words in a String, when each word is separated by a space.
3. Talk through different design patterns you know
4. Implement a sorting algorithm of your choice
5. JavaScript closure/scope question
6. difference between Java, C, JavaScript, etc. (specifically what is a closure)
7. REST API stuff

2nd interview was with a female director of engineering. interesting… also there was no coding for this one

1. what are key differences between C and Java and why do you use one over the other? especially with garbage collection
2. some memory/efficiency discussion
3. design an algorithm at low-level to determine which function of which class is used by the program (if a function in class B overrides a function in class A, how does the interpreter ensure that B's function is used and not A's)
4. design a hierarchy for shapes in MS Paint. how would you draw rectangles, circles, ovals, squares

Seems like they have a theme of asking theory questions over coding.

Facebook - Data Scientist

1. Probability: Given a concentric circle with 3 circles, what's the probability that arrow will hit the 2nd circle. Assume the arrow will hit the circle no matter what. Now, it has a 75% chance of hitting the circle, what's the probability?

2. If you are a data scientists of Facebook messenger, what metrics would you use to show Mark

Zuckerberg the health of messenger? What about for other countries?

Factset
1. Given a binary search tree, write code that can convert the BST to a doubly linked list in place.

Factual

1. What is the ideal load factor of a HashMap? If you have 70 datapoints, how many buckets should you have? Stack Overflow says 75%, so this should be about 100 buckets

2. What do *, +, and ^ do in RegEx?

3. Describe storage hierarchy to me.

FLYR Labs

Phone screen with CTO to talk about your background and tell you about the company and the team(s) you're considering. No tech question, just the usual being able to talk about projects on your resume and the sort of things you're looking for in a job.

Technical interview - On-site interview with an engineer (really short for an onsite). First talk about past projects (what was the most challenging thing about X?), group work (What would you say your role in a group tends to be?), work style (pair programming? TDD? I don't think there's a right answer, just be able to say reasonable things and have reasons for your choices), current projects, how you keep up with technical things in tech, what are you looking for in a company/job. Question is a very straightforward search implementation in Python that you do on your laptop, and you're allowed to use outside resources if you want. If you don't finish it within the allotted time, you can/should finish it at home then email them your beautiful refactored and optimized solution. They focus more on seeing how you work in general rather than on specific things.

GoodRead

1. Say I have an array of integers, from 1 - 1 million sorted. Except there's an element missing, so for example it could be 102. So the array will be like ...100, 101, 103, 104…
The array starts off at 1 and ends at 1,000,001. How would I find the missing number? Binary search on the index of the array. Say you start off and look at index 499,999. It should be 500,000. If it is, search 2nd half, else search 1st half. Keep doing this and it yields O(lgN). Edge cases are if 1 or 1,000,0001 are skipped.

Second solution: Sum up all the elements in the array and subtract from the sum of all numbers from 1 to 1,000,001.

2. Say you have a chess board, 8x8. You randomly place a white knight and a black knight on the board. How many number of configurations are there so that the white knight can't kill the black knight? Look at the board mathematically. There are 64 squares you can look at. Now say you place the white knight in the middle of the board. Then there are 9 places in which the black knight can't be at, the +1 from where the white knight currently stands, giving us 55. Now do that for the entire board. For edge cases, check if the move will be legal by checking if it will leave the range of squares 1-64.

3. Say you were to come on board at GoodRead as an intern. If you were given complete control over your work here, what project would you build and how do you think it will help GoodRead?

Goldman Sachs

1. Say we have Fib Seq Code written out. If we add memoization, what's the run time? O(N)

2. I have an array of integers and I want to find duplicates in the array. What do you do?
Hash Table, if collision, add to result.

3. Say A and B are different integers. Swap their values without a temporary variable being used. If A = 5 and B = 3.
Do A = A + B = 8.
B = B + A = 11.
A = B - A = 3
B = B - A - A= 5
or
a = a ^ b;
b = a ^ b;
a = a ^ b;

4. Give me the angle between the hour hand and minute hand.

5. What is the difference between .equals and == in Java? Contents vs pointer

6. Say a = 1, b = 2, …. z = 26. What is aaa + bbb? It's not 111 + 222 btw.

7. I have an array of integers. A triple is defined as 3 numbers in an array equal in value. Return to me a boolean to determine if an array has a triple.

8. I have a function that takes a string. And I want to expand this string input. For example, if I input the string "code", my output should be "ccocodcode"
9. Given 2 arrays, a and b, I want you to return to me the element in A that appears the most in array b. For example. If no number exists, output -1

input: a = [1, 2, 3] b = [1, 1, 3, 2, 1] output: 1
input: a = [1, 2, 3] b = [4, 5, 6] output: -1

10. Create in object oriented programming a deck of cards. What classes would you use to model a card and then create a deck from there?
11. In a table where I have employee name, their department, and salary, I want a SQL statement to return the average salary in each department.
12. What's a hashtable and what data structures can you use to implement it?

The Great Googly Moogly

Recruiting process: Applied online, got rejected from "New Grad Software Engineer - Google" but got a response from "New Grad Software Engineer - YouTube (San Bruno)". Then the first-round interviewer told me that it was a completely general interview for all of Google and YouTube.... so now I'm confused.

Fuck google's bureaucracy. THREE MONTHS?!?? #bitter

Update 04/16: Was indeed actually contacted by a recruiter a year after I first applied about re-interviewing. Lol nope. Perhaps later.

1st Round Phone Screen (45 Minutes)
There are multiple ways to solve this, but the important thing is realizing that you need to escape the delimiter (catch edge cases like ending in the escaped state).
Python things: strings are immutable, so doing a += operation is wasteful. Should use string.join( ) instead.
```
1. Serialize and unserialize a list of strings:


 String encode(List<String> input) ;
 List<String> decode(String encoded);


encode(['abc', 'def', 'ggg']) => ????


decode(????) => ['abc', 'def', 'ggg']


"abc,def,ggg" => abc,def,ggg


["abc,", "d,e,f", "ggg"] => "abc\,,d\,e\,f,ggg"


["abc,", "d,e,f", "g\gg"] => "abc\,,d\,e\,f,g\\gg"


abc,d\,e\,f,g\\gg


encoding
[','] => \, => [',']
['\,'] => \\\,
['\\,'] =>
['\\\,']
```

"\" =>

"abc" => []

```python
def decode(encoded):
        lst = []
        word = []
        escaped = False
        for c in encoded:
                if escaped:
                        # word = word + c
                        word.append(c)
                        escaped = False
                else:
                        if c == ',':
                                # lst.append(word)
                                lst.append(''.join(word))
                                word = []
                                escaped = False
                        elif c =='\':
                                escaped = True
                        word.append(c)
        if word:
                lst.append(''.join(word))
        if escaped:
                raise ValueError
        return lst




def encode(lst):
        encoded = ''
        for s in lst:
                for c in s:
                        if c == ',' or c == '\':
                                encoded += '\'
                        encoded += c
                encoded += ','
        return encoded[:-2]
```

2nd Round Phone Screen

So I actually was freaking out at the beginning of this, because as I was about to open up the google doc my laptop froze, and I ended up missing the first call because I had my earbuds

39

plugged in and was preoccupied with trying to get my laptop to work... started about 5 or 6 minutes late but the interviewer seemed pretty chill about it.

Interviewer was an engineer working on Chrome OS security. So the Chrome security team is huge, and has all the automated stuff plus the team that just constantly does bug finding/fixing, but he's more of a system architect, focused on making it so Chrome OS is resilient in the face of evil.

Question was fairly straightforward - implement pattern matching with * and ?. I think the important thing is to recognize that you need to check the different states, and if you decide to go the recursion route, immediately point out that the run time will be exponential and that you'll need to use memoization to get O(m*n) instead where m==len(string) and n==len(pattern). If you can pound out a DP solution, good for you. Also asked to prove correctness. "You literally try all the cases so the answer has to pop up somewhere?" "Yes."

```
a ab abc ba
$ ls a*
a ab abc

$ ls a?
ab
```

```
Implement [a-z] * ? shell globbing.
```

```
|string| contains characters in [a-z] (all lowercase)
|pattern| contains characters in [a-z] and possibly ?, *

? matches exactly one character
* matches zero or more characters

aa matches aa
aba matches aba
a*a matches aa, aba
a*a* matches aa, aba, abab, ababb
a?a does not match aa (? always matches one character)
a?a matches aaa
a*a matches aba
ab does not match aba (the pattern needs to match the string completely)
ab? would match aba
```

```
a*a*a
aaaaaaaaaa

# optimize by storing {(string,pattern):Bool}
# memoized run time O(len(string)*len(pattern))

def matches(string, pattern):
    if string == pattern:
        return True
    if not pattern:
        return False
    if pattern[0] == '*':  # pattern = *, string = a
        if not string:
            return matches(string, pattern[1:])
        return matches(string[1:], pattern) or matches(string, pattern[1:])
    elif pattern[0] == '?':
        if not string:
            return False
        else:
            return matches(string[1:], pattern[1:])
    else:
        if not string:
            return False
        if string[0] == pattern[0]:
            return matches(string[1:], pattern[1:])
        else:
            return False
```

Coaching Call

There's this new thing where if you're applying for a new grad technical position, they'll give you a 1-hour coaching call from an engineer. You do a video chat with them (sometimes there is more than one candidate per Google employee) and they run through exactly what you should expect / what you should do during your interviews. At this point you should know all of this already, it's basically the stuff in Cracking the Coding Interview. But I guess it helps you feel better or something.

On-Site (YouTube Campus - San Bruno)

5 45-minute interviews. 3 before lunch, 2 after lunch. During lunch you get to ask one of the engineers questions about whatever. The lunch session does not affect your interview process at all, unless you do something really bad like punch somebody. It seemed like most of the interviewers ask the same one or two questions every time they have to interview a candidate, and they all took pictures of the code at the end of the interview. Everyone was really friendly. The only annoying thing was that the portable whiteboard in the interview room was really skinny so there wasn't room to write everything nicely and I spent like 10% of my time erasing stuff. The interviewers seemed to be sympathetic about this though.

On-site 1
Interviewer was some guy whose name I forget, who worked on internal tools like code quality tracking. He was kind of excited that we both did fencing so he changed the flavor text for his question to be fencing-themed lol. If you want you can replace "fencing" with "rock paper scissors" and it still works.

Question:
You are at a poorly-organized fencing tournament, which, instead of using actual pool sheets simply recorded who beat who during pools. You are handed a pile of slips in random order that say "A --> B" if fencer A beat fencer B, and you need to create rankings for the DEs (the tournament bracket). Every fencer did not fence every other fencer. In fact, a few results may be missing, there might be some duplicates, and some might even result in cycles (A-->B, B-->C, C-->A) in which case you'd just give up and return an error. If some fencers appear to have equivalent rankings, then you can rank them randomly, e.g. if A-->B, C-->B, and B-->D then [A, C, B, D] and [C, A, B, D] are both okay.

Answer:
This is a graph traversal problem (took me too long to realize this. . . ugh. shoulda been instantaneous.). Use the slips (can input them as tuples or something) to create a directed graph, then do a graph traversal starting at source nodes to get the ordering. Once I figured out that it should be represented as a directed graph, I was allowed to assume that the graph was already constructed, so I had to write the code to find the sources then traverse the graph. I can't remember if I had enough time to implement cycle-checking :'(

I did something like this... fyi haven't checked this for correctness.

```
def rankings(graph):
```

```python
    ranks = []
    queue = Queue()
    seen = set()


    # find sources. maybe make this a helper function.
    sources = set()
    sinks = set()
    for node in graph.nodes:
        if node not in sinks:
            sources.add(node)
        for neighbor in node.neighbors:
            sinks.add(neighbor)


    if not sources:
        raise Error("There is a cycle - could not construct rankings")


    for source in sources:
        queue.add(source)
        seen.add(source)


    while not queue.is_empty():
        node = queue.pop()
        for neighbor in node.neighbors:
            if neighbor not in seen:
                queue.add(neighbor)
                seen.add(neighbor)
        insert_node(node, rankings, len(graph.nodes))


    return ranks


def insert_node(node, ordered, total):
    '''Put the node in the earliest spot before its children.
    If none of its children have been inserted, put it at the end.'''
    best_rank = total
    for neighbor in node.neighbors():
        if neighbor in ordered:
            if ordered.index(neighbor) < best_rank:
                best_rank = ordered.index(neighbor)
    ordered.insert(node, best_rank - 1)
```

On-site 2

I think this guy's name was Jin? He works on the recommendations for the Music tab that you see on the homepage.

Question:

Implement an algorithm to find the length of the shortest possible number of moves to solve one of those games where there is a maze, and you have to get a marker from a start position to a goal position by moving up/down/left/right. When you do an action the marker will move until it hits a wall. Small example:
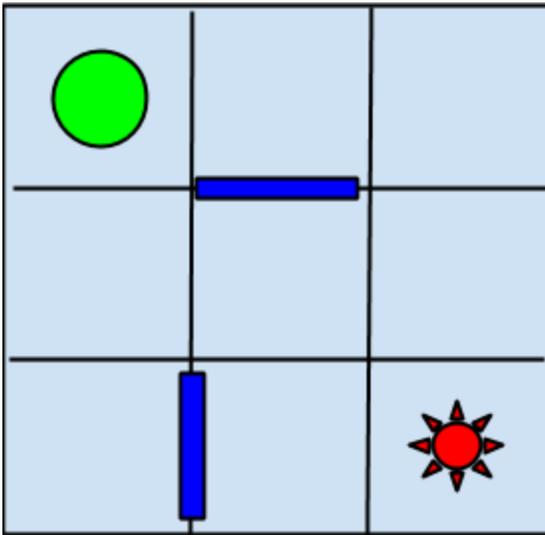
You're trying to move the circle at (0, 0) to the starburst at (2, 2).

First describe how you would implement the game representation and what kind of setup you would need to run your solver algorithm. For simplicity we're just interested in the length of the shortest path, so we don't need to keep track of the sequence of actions.

Answer:

I think he was just looking for anything reasonable. I said we could use a Board class `class Board(height, width, start, goal, walls)` where start and goal are coordinates of the start and end positions, and walls is a list of tuples with two coordinates that specify the location of a wall. For the example above, this would be something like `[((1,0),(1,1)),((0,2),(1,2))]` and you can keep track of where the marker currently is.

To solve, this is just the classic generate the graph of possible moves (make sure to not include cycles) then do a BFS traversal and return the number of steps you needed before finding a goal state / or if the board setup was invalid raise an exception or something. Once I explained what the

moves graph would look like (nodes are positions on the board, an edge from one node to another indicates that it is possible to go between those spaces. When creating this graph you would take into account the walls), I just had to implement BFS.

On-site 3
Engineer named Olive. Didn't get a chance to ask her about what she did. She was the only one who asked me about something technical on my resume, spent a few minutes going over one of my projects. I think my brain was just tired at this point / I get easily confused when I have to do things involving numbers so this one didn't go so well and I only ended up writing the brute-force solution :'(

Question:
In a number system where the digit 4 is not allowed, write a function that will take a decimal number and throw an exception if it is invalid, otherwise return the number of valid 4-less numbers that precede it.

```
count_valid(5)  --> 4
count_valid(14) --> ERROR
count_valid(15) --> 13
```

Answer:
Yeah... idk. I got that this was basically using a base-9 system, and that the answer would be recursive, but I couldn't actually figure it out. Ended up redoing the problem so there was an is_valid checker and just recursively counting up the valid preceding numbers for a linear time solution. >__> Apparently I was overcomplicating it.

Lunch
Yeah free food! The engineer I talked to was a guy named Valerie. He works on the YouTube database and its API. He moved out here after doing undergrad at MIT and worked at Oracle for awhile until he got bored then worked at a startup. Then the startup got acquired by Oracle around the same time that he was contacted by a Google recruiter, so he nope'd out of there and into YouTube. He said that he thinks for pure CS going back for a masters won't really help much, unless you're trying to change careers (MBA or specializing in biotech or something). The coolest non-engineering thing he got to do because of work was see Hillary Clinton, and he likes how there's always interesting talks being put on. For on-boarding, apparently you go through the standard Noogler camp, then another one for YouTube, and whatever team you're on will probably also have a ramp-up period. They're trying to make the YouTube onboarding more specific.

On-site 4
Guy named Charlie who was the most neckbeard-y person I've ever seen in my life (no fedora, though), but he was pretty amiable. Didn't have time to ask him what he did.

Question:

This is a vastly simplified version of a [very common](link)problem.

You are given a rectangular box in which you are to display some text. You have:
-   the box's height and width in pixels
-   the string which you want to display
-   the font of the string (entire string is in one font), the min and max sizes of this font (ints)
-   functions widthOf(char, font) and heightOf(font), which return a int value in pixels.

What is the maximum font size we can use to display the entire string in the box? If the text can't fit in the box, just return the min font size. Don't worry about splitting the string nicely.

Answer:
The general approach should be fairly obvious, for checking if a string at a given font size fits in the box. To get the largest possible font, run a binary search. Y'all can google how to implement binary search, but the edge cases to think about with the fits_box(string, h, w, font, size) function is what if you have a box with width 2px and length 300px and nothing fits, or if the box is width 300px and height 2px and nothing fits?

On-site 5
Forgot this guy's name and I didn't get a chance to ask him what team he was on. He gave off an aura of "really soft-spoken, but knows his shit."

Question:
Given a black and white image which is represented as a one-dimensional array of bytes, write a function to flip the image so that it's a mirror image of itself. For example if the original image is just the letter d, after calling the function on it, it should be an image of the letter b. Assume the lengths and widths fit into bytes even though 1px is represented by 1 bit.

Answer:
Yeah, this is probably one of those interview questions where it would have been better not to do Python. The straightforward way of doing this is just to go down each row, from the outermost bytes to the innermost bytes, and swap/mirror the bytes. If there is an odd number of bytes, then go through the center column and mirror all of those bytes. Wrote the swapping function (I think the idea was to check that you know how to work with matrices) and the mirror function (can you do bit manipulation).

Follow-up: How can you make this faster? Assuming that you can't use parallelization, I guess you can just do calculations in larger chunks. In particular, if you know you're going to be calling mirror(byte) a lot of times, you can pre-calculate the 128 possible mirrorings and just do a lookup.

Groupon

1. Reverse a linked list
2. Tell me about priority queue, hashtable, and binary search tree
3. Write a recursive function, with only 1 input i, generate the i'th prime number.

Guidebook

1. Can you explain to me what a dictionary is in Python? alternatively, explain a hashmap in java.

2. Explain atomic operations to me.

3. Can you explain how TCP/IP works?

4. Code a method in any language that finds if argument A is a substring of argument B.
str1.toLowerCase().contains(str2.toLowerCase());

5. **Exercise 1:**
input: [1,1,1,1,2,3,5,5,7]
output: [1,2,3,5,7]

Provide a solution that the output contains no duplicates.
```
public int[] nodup(int[] array) {
    ArrayList<int> result = new ArrayList();
    int last = array[0];
    for (int i = 0; i < array.length; i++) {
        if (!(array[i + 1] == last)) {
            result.append(last);
            last = array[i];
        }
    }
    return result;
```

If(we)
1. Resume questions, What are you looking for
2. Given two strings, check if the are anagrams
3. Given a dictionary of words (represented as an array of strings), check how many pairs of anagrams I have

1. Fast way to multiply a # by 7 in a computer
2. Factorial iterative and recursive
3. Check if a # is a power of 2

lmo.lm

1. Given array of unsorted integers (the array has size N) find the largest element in the array.
For loop, iterate and find largest. O(N). Nothing Faster.

Find the set of the largest K (K<N) elements in the array.
If your array is of size K, then that's it. Now say it's size K+1. We know first K elements are largest. When we look at K+1, search for minimum in K-size array, if min < K+1, replace. Repeat this process until end of array. O(N)

Now, what data structure can we use to optimize this search of K largest elements? Use Binary Search Tree, O(N*lgK)

2. Given large file containing only integers ~ 128 GB stored on disk and computer that has just 2 GB of free memory. Sort this file. You have extra disk space if you need it.

The interviewer also mentioned the stuff inside the files are non radix sortable. So use merge sort. Break the files down into 2GB, giving 64 files total. Upload one at a time, mergesort, and once finished, remove from RAM. Then upload next onto and so on till all 64 are sorted. Then this is where we stopped. There's something beyond this...CALLED external sorting

3. Given two words A and B (|A| = |B| < 10).
bool is_in_dictionary(string word) O(1)

One Step Transformation between two words X and Y exists iff
|X| = |Y| and X and Y differ in exactly 1 character.

Find the shortest path of OST between A and B
where each intermediate word is in the dictionary.

'a'-'z'

A=dog, B=let
D={dog, let, log, leg}
returns: ["dog", "log", "leg", "let"]

A=small, B=smile

D={small, smile, stile, stall, still}
returns: ["small", "stall", "still", "stile", "smile"]
Think of this as a graph problem. Each node represents a word and its neighbors are words with a one letter difference. Use Dijstkra's algorithm to find shortest path.

Which searching algorithm would you use to find solution? Breadth First Search since it goes layer by layer and we want to find shallowest layer possible.

How to return the path that gives us the shortest path? Look at each node's parent nodes and you shall find the desired path.

Infer
input list
YYV -> SFO
MDT -> JFK
JFK -> YYV
ORD -> MDT

return ORD -> MDT -> JFK -> YYV -> SFO

```
# take input list of pairs, and construct src->dest dict
# find start airport
# curr airport = start airport
# while curr airport != dest airport
#  add curr airport to ret list
#  curr airport = dest [ curr airport ] (from src->dest dict)
# return ret list
```

Informatica
1 hour pair-coding phone interview
March 24, 2014

1) Code a preorder traversal for a binary tree. Do this recursively and iteratively.

2) Given a String, find its first non-repeating character.
Example: "bumble bee" => 'u'

Possible solutions:
a) Assume the String is ASCII encoded. Then there are only 256 possible characters in the String,

so we can create an int[] array of 256 elements, iterate through the String and "count" the number of times we have seen each character. With the array filled out, we iterate through the String again and check if we have seen this character exactly once. Return the first character we find. This is pretty much modified counting sort.

b) Create two sets, and iterate through the String. The first time we see a character, we put it in the first set. If we have seen a character two or more times, we put it in the second set. For the final answer, iterate through the String again and return the first character not in the second set.

c) The interviewer prodded me to reduce it from two sets to one set, but I couldn't get it so she explained the answer. Create a Map that maps each character in the string to a true or false value if we have seen the character exactly once, or if we have seen it two or more times.

3) Design a top-10 favorite list for your browser.

Ans: For each website in your history, use a series of metrics such as "time on site", "frequency", "in bookmarks", and "Page rank" and assign each a weight w_1, w_2, w_3, w_4, Then compute a heuristic h(website) = "time on site" × w_1 + "frequency" × w_2 + "in bookmarks" × w_3 + "Page rank" × w_4 …
for every website, and find the top 10. You only need to sort the top 10, so do not sort the the thousands of websites below rank 10. While iterating through the heuristic array, use a min-heap to maintain a rolling list of the 10 best heuristics. Update this list on browser close or when there is free CPU time.
4) Asked by hiring manager: What is a virus?

Intentional Software

My interviewer was pretty chill and just had me describe how I would program these instead of reading him code through the phone.

1. Say you have one arbitrary String variable. How would you find the number of words? (iterate through and find the number of spaces + 1; need to account for edge cases though like no spaces, only spaces, blank String)
2. Let's say you're a Google developer, and you have a giant list of results, say a million of them. They are unsorted. How do you find the 10 most relevant results to display on the front page? (use min heap)
3. Implement a Tree data structure. (use objects and variables and constructors)
4. Implement breadth first traversal and depth first traversal. (queue and stack)
5. When considering graphs, can you use breadth first and depth first traversals like you do for trees? Explain. (I talked about how you can definitely use breadth first, but depth first might be a problem because the maximum depth for a tree is pretty defined, but for a graph it might be

arbitrary. also a tree has a defined end, and graphs may not, so you don't want to be caught in an infinite loop)

6. Your input is a list of the format {A before C, B before C, C before D…}. All of these are tasks, each of which must be completed before an arbitrary number of other tasks (there may be more than one order that works). Tell me how you would find an order of tasks that works. (use directed graph that is formed by iterating through the list. find the # of unique nodes. then use breadth first traversal. also need an integer weight along the way to keep track of the current level away from the first node as you progress)

(He gave me the last question with less than 5 min. to go and said that it was pretty open-ended. He also gave me the answer right after I said use graph, find #, and use BFT. So I think he wasn't expecting anybody to get that question.)

Intuit

Mike's questions were hella dumb just fyi, but mine was:
figure out if a binary tree's left side is a mirror of its right side. just google the answer, my interviwer was dumb so he didn't understand what i was doing, but you can do it via conditionals of recursive calls for example

Here are Mine (Mike):
- What is local scope vs global scope
- What does static mean
- Write a function to detect a duplicate in an array

IXL Learning

- Given arandom function random(a, b) that returns a random integer in the range of a to b, implement a randomOdd(a, b) function that returns a random odd integer in the range of a to b
- Given an array of ints, find two numbers that sum to a target and return those two numbers
- Given an array of ints, find three numbers that sum to a target and return those 3 numbers
- Do this generally, in a recursive manner

Jane Street
- Write a function to mirror a binary tree
- Implement Tetris
- Create a text/console game library
- Client/Server tree diff algorithms to minimize RPC calls
- Implementing an abstraction for lazy and iterative execution of calculating and comparing

lengths of collections of lists in parallel efficiently

JP Morgan

1. Explain to me what abstract and interface are in Java.

2. What does a constructor do? What about a destructor(Garbage Collection in Java)?

3. Can you explain to me how to remove duplicates in a SQL query? Use distinct in Select Statement

4. Difference between Outer and Inner Join? Think of Venn Diagram. Inner Join is the intersection of records in both tables. Full Outer join is the entire Venn Diagram. Left Outer Join is Table A + The intersection of A and B. Right Outer join is Table B + intersection of A and B.

5. What are tuples in Python? Immutable Data structures

6. How to reverse a string? What about reversing a phrase? Say for example, "My name is Michael". I want "Michael is name My". How to do that?

To reverse a string do the following code.
```
public static String reverse(String word) {
    int i, len = word.length;
    StringBuffer dest = new StringBuffer(len);
    for (i = (len - 1); i >= 0; i--) {
        dest.append(word.charAt(i));
    }
    return dest.toString();
```

To reverse a sentence, do the following.
```
public static String revIt(String sentence) {
    String[] words = sentence.split(" ");
    String rev = "";
    for (int i = words.length - 1; i >= 0; i--) {
        rev = rev + words[i] + " ";
    }
    return rev;
}
```

7. Say you have two coins. One of radius R and One of radius S. Coin R and Coin S are making contact at a single point. If I were to roll Coin S around Coin R until Coin S is back at its starting

point, how many rotations will Coin S go through? 1 + R/S
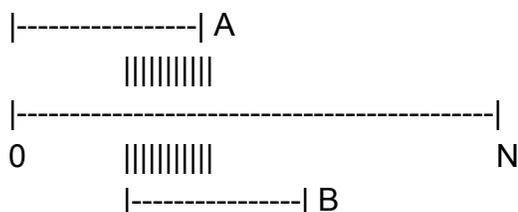
8. I have this code right here.
Arrays A and B are millions of integers. The arrays are monotonically increasing. Duplicates allowed.
ForEach(Elem in A)
   ForEach(Elem in B)
     If Elem.A == Elem.B
       Result.add(A)

What's run time and what's it doing? $O(N^2)$ and returning a list of intersections.
What can we do to make this faster?
Think of it like this. The line represents a line of numbers.

```
|----------------| A
          ||||||||||||
|-------------------------------------------|
0          |||||||||||||                     N
          |----------------| B
```

The shaded part represents the intersection. Thus you can use binary search to find the smallest # and the largest #, and return that. $O(lgN)$f

9. Say I have this gambling situation. This game allows me to roll a die and whatever number I roll, I get that amount in $. What's a fair fee to pay for this game? ⅙ * (1 + 2 + 3 + 4 + 5 + 6) = $3.5

Now say we have the option of re-rolling if we don't like our number. What's the fair fee now?

10. What is a singleton?

11. What is SDLC?

LinkedIn

1. Given a string input, determine if it's a valid floating number. The interviewer wouldn't let me use regex, so you have to iterate through it and have flags for decimal and negative signs.

2. Find the maximum subarray sum. Now find maximum subarray product.

Site Reliability Engineer (new grad) - HR screen questions
The usual "why LinkedIn, why this job, are you willing to relocate to Mtn View...." Also asked for GPA.
There are a couple of basic UNIX / networking trivia questions, including:

- in IPv4, what is the address to broadcast to all networks
- in UNIX, what is the command to see the most recent command (i confused myself on this one but I hope they just meant the up-arrow)
- what is the system call to start a new process

SRE Phone Screen: FizzBuzz and parsing log files (regex or lots of python string.split())

LiveRamp

Pre-interview online test:
A couple simple linked list questions.
Which of these is an asymptotically tight upper bound for the best case running time of an optimal algorithm that finds the maximum element in an arbitrary array of integers of size n?
What is the upper bound of the fastest possible algorithm to compute the sum of all possible pairs in an array?
How would you design a program to find the shortest path in the Six Degrees of Kevin Bacon game between two people?
Why are you interested in working at LiveRamp? <-- hardest question right here

Phone screen
30-minute phone screen with an engineer. This was strangely short and curt compared with other interviews I've had. Might have just gotten a socially awkward interviewer, but the inflection and tempo of his speech made it feel like I was talking to an old school automated voice system. Not rude or anything, just kind of weird. Asked how I was liking Berkeley, what classes I was taking that I enjoyed, asked some questions about what I was learning in said classes. One question, no code sharing, just talking. He started sounding a little bit more like a human at the end of the interview. Maybe he just wasn't awake yet.

Q: Given two strings of the same length, how would you find the shortest path to change one word into another, using only valid words? Assume you have a function to check if words are valid, and don't worry too much about efficiency.
A: Conduct what is essentially a BFS of the word space by generating all possible words you could get to from a state, and store those in a queue. To keep track of the path, you could store the words in a dictionary with they value being the word used to generate the key. Since you're using BFS you know that when you do find the end state, you found a shortest path.

Didn't ask about anything on my resume or what previous experience I had....kindaweirdyou'dthinkthey'dbeinterestedinthat. I asked a couple of questions about what LiveRamp actually does and how their engineering teams are structured, since it's still really confusing what *exactly* they do even after reading the website. From what I gather, some company says "we have some

data, we need it to be in format X so that we can feed it into platform Y for analytics things," and LiveRamp does the transformations using something involving Cascading and Hadoop. They also have front-end teams for customer-facing and internal UIs. They don't have any sort of required rotation program but it's fairly common and easy for engineers to move around within the company if they want to try different things. The interviewer actually said that when he started there even he wasn't sure what exactly the company did so it's perfectly understandable to be confused about what they do_{wut...ok}

MathWorks

Math Questions
What is pigeonhole principle?
Simplify (A intersect B) union C)
P union (P intersect Q) <=> Q is a tautology?
What is P vs NP?
Definition of countable vs uncountable sets?
What is diagonalization?

Programming Questions
Local vs global scope
Why are global scope variables bad?

What are static variables? static means that the variable or method marked as such is available at the class level. In other words, you don't need to create an instance of the class to access it.

What's a recursive function? Give me an example
What is OOP, in your own words?
Advantages of OOP? vs Functional Programming?
What is a linkedlist, any advantages?
What's a hash function?
Good properties of a hash function?
Differences between directed acyclic graph vs tree?
What is $O(n)$ notation? What does it mean?
If $f = O(n)$, $g = O(n!)$, then $f + g$ is…? $O(n!)$. $f(x) + g(x) = max(f(x), g(x))$
$f = O(n)$, $g = O(n^2)$, $f * g$? $O(n^3)$. $f(x) * g(x) = f(x) * g(x)$

Java Questions
What is JDK? JRE?
Platforms and JVM?
What java library function allows you to detect spaces in a string? Trim method
what's an abstract class?
what is a default constructor in java?

Meltwater

Initial code screen: Codility test with 5 questions. You can choose the language [Python, Java, C#, Ruby] Solutions are correct (though kind of ugly and the tabbing got screwed up) and meet efficiency requirements unless otherwise stated.

Q1

Find the length of a linked list. (You should be able to do this in your sleep.)

Q2

SQL query so easy that you don't really even need to know SQL to do it.

Q3

A non-empty zero-indexed array A consisting of N integers is given. A 'pit' in this array is any triplet of integers (P, Q, R) such that:
  * 0 ≤ P < Q < R < N;
  * sequence [A[P] > A[P+1], ..., A[Q]] is strictly decreasing,
  i.e. A[P] > A[P+1] > ... > A[Q];
  * sequence A[Q], A[Q+1], ..., A[R] is strictly increasing,
  i.e. A[Q] < A[Q+1] < ... < A[R].

The 'depth' of a pit (P, Q, R) is the number min{A[P] - A[Q], A[R] - A[Q]}.

For example, consider the array A consisting of 10 elements such that:

A = [0, 1, 3, -2, 0, 1, 0, -3, 2, 3]

Triple (2, 3, 4) is one of three pits in this array, because sequence [A[2], A[3]] is strictly decreasing (3 > -2) and sequence [A[3], A[4]] is strictly increasing (-2 < 0). Its depth is min{A[2] - A[3], A[4] - A[3]} = 2. Triple (2, 3, 5) is another pit with depth 3. Triplet (5, 7, 8) is yet another pit with depth 4. There is no pit in this array deeper (i.e. having depth greater) than 4.

Write a function:

        def solution(A)

that, given a non-empty zero-indexed array A consisting of N integers, returns the depth of the deepest pit in array A. The function should return -1 if there are no pits in array A.

For example, consider array A defined earlier. The function should return 4, as explained above.

Assume that:
        * N is an integer within the range [1..1,000,000];
        * each element of array A is an integer within the range [-100,000,000..100,000,000].

Complexity:

* expected worst-case time complexity is O(N);

* expected worst-case space complexity is O(N), beyond input storage (not counting the storage required for input arguments).

Elements of input arrays can be modified.

```python
def debug_helper(i):
    if i==0:
    return "maybe in pit"
    elif i==1:
    return "in pit"
    else:
    return "not in pit"


def calc_depth(l,b,r):
    # print l, b, r
    return min(l-b, r-b)


def solution(A):
    # write your code in Python 2.7
    """
    Model as FSM where edges are whether value increases, decreases, or remains level
    nodes are whether we are in a pit (had been descending, now ascending), not in a
pit,
    or might be a pit (descending).
    """
    depth = -1
    # -1 = decreasing, 0 = level, 1 = increasing
    prev = None
    # 0 = maybe, 1 = in pit, 2 = not in pit
    state = None
    prev_val = A[0]

    left_edge = A[0]
    bottom = None
    right_edge = None

    for i in xrange(1,len(A)):
    # print "* at index {0}".format(i)
    # print "From state: {0}".format(debug_helper(state))
    if A[i] < prev_val: # descending
            # print "descending {0} -> {1}".format(prev_val, A[i])
            if state == 1:
            new_depth = calc_depth(left_edge, bottom, right_edge)
            if new_depth > depth:
                    depth = new_depth
            left_edge = A[i-1]
            if state == 2:
            left_edge = A[i-1]
            state = 0
            prev = -1
    elif A[i] > prev_val: # ascending
```

```
            # print "ascending {0} -> {1}".format(prev_val, A[i])
            if state == 0:
            bottom = A[i-1]
            if state == 0 or state == 1:
            right_edge = A[i]
            state = 1
            elif state == None or state == 2:
            state = 2
            prev = 1
        else: # if level, then not in pit
            # print "level {0} -> {1}".format(prev_val, A[i])
            if state == 1:
            new_depth = calc_depth(left_edge, bottom, right_edge)
            if new_depth > depth:
                depth = new_depth
            left_edge = A[i] # right_edge = None; bottom = None
            state = 2
            prev = 0
        prev_val = A[i]
        # print "To state: {0}".format(debug_helper(state))

        if state == 1:
        right_edge = prev_val
        new_depth = calc_depth(left_edge, bottom, right_edge)
        if new_depth > depth:
            depth = new_depth
        return depth
```

Q4

A non-empty zero-indexed array A of N integers is given. A pair of integers (P, Q), such that $0 \leq P \leq Q < N$, is called a 'slice' of array A. The sum of a slice (P, Q) is the total of A[P] + A[P+1] + ... + A[Q].

A 'min abs slice' is a slice whose absolute sum is minimal.

For example, array A such that:

A = [2, -4, 6, -3, 9]

contains the following slices, among others:

    * (0,1), whose absolute sum = |2 + (-4)| = 2
    * (0,2), whose absolute sum = |2 + (-4) + 6| = 4
    * (0,3), whose absolute sum = |2 + (-4) + 6 + (-3)| = 1
    * (1,3), whose absolute sum = |(-4) + 6 + (-3)| = 1
    * (1,4), whose absolute sum = |(-4) + 6 + (-3) + 9| = 8
    * (4,4), whose absolute sum = |9| = 9

Both slices (0,3) and (1,3) are min abs slices and their absolute sum equals 1.

Write a function:

> def solution(A)

that, given a non-empty zero-indexed array A consisting of N integers, returns the absolute sum of min as slice.

For example, given the array A as defined above, the function should return 1.

Assume that:
> * N is an integer within the range [1..100,000]
> * each element of array A is an integer within the range [-10,000..10,000]

Complexity:
  * expected worst-case time complexity is O(N*log(N));
  * expected worst-case space complexity is O(N), beyond input storage (not counting the storage required for input arguments).

Elements of input arrays can be modified.

*FYI This solution got 80% on correctness and 0% on efficiency.

```
from itertools import combinations
def solution(A):
    # write your code in Python 2.7
    slices = combinations(xrange(len(A)+1), 2)
    min_sum = abs(A[0])
    for slice in slices:
    print slice, min_sum, abs(sum(A[slice[0]:slice[1]]))
    min_sum = min(min_sum, abs(sum(A[slice[0]:slice[1]])))
        return min_sum
```

Q5
A non-empty zero-indexed array A consisting of N integers and sorted in a non-decreasing order is given. The 'leader' of this array is the value that occurs in more than half of the elements of A.

You are given an implementation of a function:

> def solution(A)

that, given a non-empty zero-indexed array A consisting of N integers, sorted in a non-decreasing order, returns the leader of array A. The function should return -1 if array A does not contain a leader.

For example, given the array A consisting of ten elements such that:

> A = [2, 2, 2, 2, 2, 3, 4, 4, 4, 6]

the function should return -1, because the value that occurs most frequently in the array, 2, occurs 5 times, and 5 is not more than half of 10.

Given an array A consisting of 5 elements such that:

A = [1, 1, 1, 1, 50]

the function should return 1.

Unfortunately, despite the fact that the function may return expected result for the example input, there is a bug in the implementation, which may produce incorrect results for other inputs. Find the bug and correct it. You should modify at most THREE lines of code.

Assume that:
* N is an integer within range [1..100,000];
* each element of array A is an integer within the range [0..2,147,483,647];
* array A is sorted in non-decreasing order.

Complexity:
  * expected worst-case time complexity is O(N);
  * expected worst-case space complexity is O(N), beyond input storage (not counting the storage required for input arguments).

Elements of input arrays can be modified.
```
Example test:   [2, 2, 2, 2, 2, 3, 4, 4, 4, 6]
OK

Example test:   [1, 1, 1, 1, 50]
OK
```
```
def solution(A):
      n = len(A)
      L = [-1] + A
      count = 0
      pos = (n + 1) // 2
      candidate = L[pos]
      for i in xrange(1, n + 1):
      if (L[i] == candidate):
            count = count + 1
      # if (count > pos):
      if ((count > pos) and n%2==0) or ((count >= pos) and n%2==1): # changed frm above
      return candidate
      return -1
```

Microsoft

Given a 2D array of integers, find each zero in the array and zero out its row and column.

Write pseudocode to solve the tower of hanoi.

Write a generic hashtable in Java

Reverse strings and linked-lists

Write a Run-Length encoding function for strings in C

Given a two nodes in a non-pathological tree, find their their closest common ancestor in O(log(n)) time.

Given a list of integers, and an integer contained one or more times in the list, return a random index for a copy of that integer in O(n) time

Check if a BT is a BST

Given a timeline for a stock (A chronologically ordered list of day objects containing the highest and lowest price for the stock on that day) find the date to buy and the date to sell that would result in the highest profit.

Minted
Making change: Given
1. Currency system
2. Amount of money

write me a method that can return the combination of coins and bills (given a currency system) that will represent the amount of money inputted. This must return a combination that uses the fewest amount of coins and bills

amount money = 0, 1

Find highest currency that can represent the amount
from the difference, repeat process until amount is = 0

N: amount
K: len(currencysystem)

```
public static ArrayList<Integer> returnSmallestChange(HashMap<Integer,
Boolean> currencysystem, int amount) {
    if (amount < 0) {
        return some kind of warning that amount was invalid
    }
    ArrayList<Integer> answer = new ArrayList<Integer>();
    int change = amount;
```

```
    while (change != 0) {
        int biggestChange = biggestCointSmalllerThan(currencysystem, change);
        answer.add(biggestChange);
        change = change - biggestChange;
    }
}


// Assume you already have this:
public int biggestCoinSmallerThan(HashMap<Integer, Boolean> currencysystem,
int amount)


Amount  - BiggestCoin = difference
if (difference > factor * biggestCoin)
    factor++
```

Newfield Wireless

(Python Quality Engineer)
HR Phone Screen
- Standard sort of "tell me about yourself" and "what questions do you have about the company/role".
- You should demonstrate that you know at least at a high level what their main product does.

Hackerrank Code Test (Must use Python. Should be easy.)
1. Check if an item is in a binary search tree.
2. Given a string, return whether it has balanced parentheses (use a stack)
3. A flag is represented as a 2D matrix. A valid flag is composed of stripes, which means that every row only has pixels of a single color, and adjacent rows are not the same color. Write a function that will return whether given flags are valid or invalid.

Phone screen with engineer from QE
They go the "ask many specific questions" rather than "ask one general question and do it on collabedit" route. Questions include things like
- Linux - How do you find the process that's using the most memory and kill it?
- Python - Given a very long list of objects with some quality, how do you remove all the objects with quality X.
- Python - How do you find the mode of a list of things?
- C++ - What is `const int * const`
- C++ - What is a reference variable?
- QA - What is whitebox vs blackbox testing? Pros and cons of each.
- QA - Engineer describes a product / scenario (complicated enough that you'll need to draw a diagram for it) with emphasis on you not having all the information on its usage, you pull things out of your ass on what information would be most helpful and how you would test it for functionality and compatibility.
- QA - What if there's too many things to test within some module and you need to find what causes a crash? Assume you have no info about usage.
- QA - How could you improve an optimization project that keeps getting stuck at local maxima? (Note that this is the same idea as the previous question.)

- Probability - If you've learned Bayes rule in at least one class you should be able to do this. If anything the hardest part is not overcomplicating it.

Nodeprime (tj)

30 minute phone interview

February 25, 2015

1. Describe what is an anagram.
2. Given two strings, how would you check if they are anagrams of each other?
3. How would you do this on $O(n \log n)$?
4. How might you do this in $O(n^2)$? O(exponential)?

3 hour onsite interview

March 4, 2015

1. Given an array, return true if two values in it sum to 1000
2. Explain how hashing works. How might you implement a hash function?
3. Given three hash functions h1, h2, h3; implement a bloom filter.
4. Write a parser for Javascript floating point numbers
5. Write a JSON parser
6. How might you not implement this recursively?

OpenTable (PM)

1. How would you define the PM role and responsibilities?
2. Tell about your favorite product and why you like it
3. I said Quora, so follow up questions include Why do you think Quora became successful as a Q&A platform compared to others, like Yahoo Answers?
4. If you were the CEO of Quora, what are some important metrics you would want to look at? (Vanity vs non-vanity metrics)
5. How would you determine what actions to take based on those metrics?
6. If you were the CEO of OpenTable, what would be some things keeping you up at night?
7. I said Competition, so discussion of Yelp vs OpenTable.
8. What would you do to ensure OpenTable is better than Yelp's SeatMe?
9. How can we go about making the content on OpenTable much easier for users to understand?
10. How would you go about determining the amount of OpenTable points if users posted a review about the restaurant or photo?
11. How would you go about making users go to OpenTable's reviews of restaurants over Yelp's?

Optimizely

1. What are you looking for and what do you want out of your experience?

2. Tech interview question: Write a function that determines the longest palindrome in an input String and returns the palindrome.

I spent the entire interview solving and optimizing this thing. The interviewer said I had the optimal solution at the end.

Palantir

1. Tell me the differences between waterfall and agile development.

2. Let's design a database for Tim. Tim has some toys and he wants to be able to keep track of what toys he currently has, what toys he has lent out to his friends, and which friends have his toy. Design this database.

3. (Added by Jeff Huang) Implement Snake.

```
* Given a string of '(' and ')' characters, determine whether the string is balanced.
*
* Examples:

*   '()'   : balanced. return true.
*   '(())' : balanced. return true.
*   ')('   : NOT balanced. return false.
*   '('    : NOT balanced. return false.
*   ')(()' : NOT balanced
*   '())------'
*   ')'
*/
boolean isBalanced(String parens) {
    int leftCount = 0;
    int rightCount = 0;
    for (int i = 0; i < parens.length(); i++) {
        if (rightCount > leftCount) {
            return false;
        }

        Char currentChar = parens.charAt(i);
        if (currentChar == '(') {
            leftCount += 1;
        } else if (currentChar == ')') {
            rightCount += 1;
        }
    }
    if (leftCount == rightCount) {
        return true;
```

```
        }
        return false;
    }

    /*
     * Given a string of '(' and ')' characters, flip the minimum number of
     * parentheses required to balance the string, if it is not already balanced.
     * Then, print out the balanced string, and return the number of flips that
     * were required. If the string is impossible to balance, print nothing, and
     * return -1.
     *
     * Examples:


     *   '()'   : already balanced. print '()' and return 0.
     *   '))()' : NOT balanced. print '()()' and return 1.
     *   '(()('
     *    '()()

     *   '('    : NOT balanced. return -1, because it's impossible to balance.
     *   '(()' : -1 //Odd length, lc diff from rc -> -1

        ')(()'
         **
        '()()'
        '(())'

        SB = '()()'
     */
    int fixBalance(String parens) {
      // use ifBalanced and if it's true, return 0
      // check length to see if even or odd
      // if odd return -1
      //
        if (isBalanced) {
            return 0;
        }

        int length = parens.length();
        if (length % 2 == 1) {
            return -1
        }
```

```java
        int numberofChanges = 0;
        int leftCount = 0;
        int rightCount = 0;
        StringBuilder correctlyBalancedString = new StringBuilder();
        for (int i = 0; i < length; i++) {
            Char currentChar = parens.charAt(i);
            if (currentChar == '(') {
                leftCount += 1;
            } else if (currentChar == ')') {
                rightCount += 1;
            }
            if (rightCount > leftCount) {
                rightCount -= 1;
                leftCount += 1;
                correctlyBalancedString.append('(');
                numberofChanges += 1;
            } else if (leftCount > rightCount && parens.charAt(i+1) == '(') {
                correctlyBalancedString.append(')');
                leftCount -= 1;
                rightCount += 1;
            } else {
                correctlyBalancedString.append(currentChar);
            }

        }
        System.out.println(correctlyBalancedString.toString());
        return numberofChanges;

}
```

RealNetWorks

1. What are normal forms 1, 2, and 3?

Redfin

1. Resume questions
2. Given two linked lists, check if they are merged (ie intersect and end at the same null tail)
3. Given a char array, rotate the array given a certain index

RGM Advisors

1. Say you have an array of numbers and a target number. You want to check if there is a sum of

two numbers in the list that will equal the target number and return that pair.
Have a pointer point to the 1st element. Then, do target number - pointer number = complement number. Now, when you look at the first element, check if it is in the HashMap. If not, Hash so Key is index and value is number. Then check if the value number is equal to the complement number and Key number doesn't equal current index. Repeat until you reach a desired number or return null at the end if not found. Return pair. O(N)

2. What are your languages and how would you rate yourself? What are the benefits and downsides to each language for you personally and when would you use one over the other?

3. In Java, what's the difference between final, finally, and finalize?

4. What's the difference between int and Integer in Java?
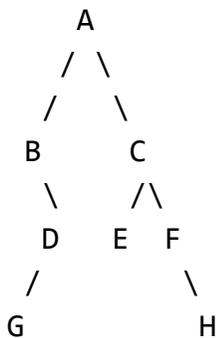
Sift Science

You're handed the root node of a binary tree.
Each node has 3 fields : left, right, next
left and right are the normal tree links
next has the property that, for every node which is not the rightmost node at its depth
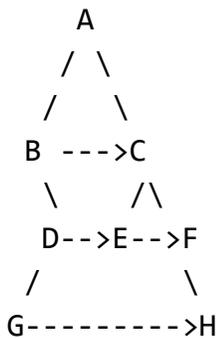node.next points to the node on its right at the same depth.

Given input:

```
        A
       / \
      /   \
     B     C
      \    /\
       D  E  F
      /       \
     G         H
```

Where the binary tree is:
- not assured to be balanced (there may be deep paths in some places, and shallow paths in others)
- nodes have 3 pointers:
  - left
  - right
  - next (which starts null)
- note that nodes do *not* have pointers back to the parent.

Fill in next pointers to produce this output (where the horizontal pointers are the "next" pointers):

```
        A
       / \
      /   \
    B --->C
      \     /\
     D-->E-->F
    /          \
  G--------->H
```

Rule:
- If Node n is the right-most node at its depth, n.next remains null/empty/nil
- otherwise, n.next points to the node immediately to its right at the same depth

```
class Node {
    private Node left;
    private Node right;
    private Node next; // you're setting these values
    // there no pointer back to the parent
}
```

Smarking
```
# def say_hello():
#    print 'Hello, World'

# for i in xrange(5):
#    say_hello()



# input -> clients.csv

# client, garage, spaces
# abc parking, sf garage, 1000
# parkxyz, central park garage, 600
# abc parking, unlimited garage, 500
# ...



# client, spaces
# abc parking, 1500
```

```
# parkxyz, 600
# …


# dict = {} #key = client name, value = []
# {abc parking: 1500, parkxyz : 600}

def aggregateClientParkingSpace(csv):
    dict = {}
    for row in csv:
        sentence = row.split(",")
        if sentence[0] not in dict.keys():
            dict[sentence[0]] = sentence[2]
        else:
            dict[sentence[0]] += sentence[2]


#   C1   C2
# [0,4],[4,0]
# C3     C4
# [2,2],[3,3]
# import Math

def isOverLap(coord1, coord2, coord3, coord4):
    xMin = min(coord1[0], coord2[0])
    xMax = max(coord1[0], coord2[0])
    yMin = min(coord1[1], coord2[1])
    yMax = max(coord1[1], coord2[1])
    if (coord3[0] >= xMin and coord3[0] <= xMax) or (coord4[0] >= xMin and coord4[0] <= xMax):
        if (coord3[1] >= yMin and coord3[1] <= yMax) or (coord4[1] >= yMin and coord4[1] <= yMax):
            return True
    return False


print (isOverLap([0,4],[4,4],[2,2],[3,3]))
```

rashmi.malani@smarking.net

a.csv

10001, 2015-08-01 08:21, 2015-08-01 17:20
10002, 2015-08-01 13:03, 2015-08-01 21:31
10001, 2015-08-01 13:03, 2015-08-01 21:31

10001 -> b.csv

Assume there are 5 different price tiers 1 - 5. Write a SQL query to find the number of transactions in each price tier, and order by the number of transactions, from high to low.

transactions_table

tiers|transaction_id

1|tx234
2|tx453
2|tx333
3|tx000

1 : 1
2: 2

SELECT sum(case 1 when tiers = 1 else 0 end), sum(case 1 when tiers = 2, else 0), ...
FROM transactions_table

1, 2, 1

Rotated Array Binary Search
Merge two sorted singly linked list into one without building another data structure

Simply Hired

1) Given an array of 3 distinctly colored marbles (ie. Red, White, Blue), sort them in the order Red, White, Blue. [Dutch National Flag problem for hint/answer]   [ONSITE]
2) Writing a SQL query that involved three different tables and having to reference each table for info(don't remember full details)   [ONSITE]
3) Questions about JS, mainly testing to see if I understood "this" and what "this" refers to in particular contexts. [ONSITE]
4) What are some ways to hide an element with just CSS? Explain the different between visibility and display, when "hiding" an element. [ONSITE]
5) Write the power function. (key thing here was to talk it out, consider edge cases, and write a recursive, then iterative method. Then handle negative powers.   [PHONE]

SOASTA

1. What is encapsulation?
2. What is a deadlock?
3. In Java, what can I use to make sure multiple threads don't conflict with one another?
4. What is a join in databases?
5. What is a prepared SQL statement?
6. What is the fastest sorting algorithm to sort a list of numbers?

Splunk

1st round questions:

1. from blah import foo: what are blah and foo?
2. say you just declare a class like this:
class blah{ }
What is the accessibility of the class? *package private*

2nd part is a coding test with the manager, but he required me to give the answer in Java and Python. I was too lazy to do it.

Square

```
ROWS = 40
COLS = 40

canvas = [[' '] * COLS for r in range(ROWS)]

#render_circle(20, 20, 5)

def render_circle(x_center, y_center, radius):
    if x_center > ROWS or x_center > COLS or y_center > ROWS or y_center > COLS
or x_center < 0 or y_center < 0 or radius < 0:
        print("Error with origin coordinates or radius")
    canvas[x_center][y_center] = "."


    list = find_coordinates_within_circle(x_center, y_center, radius)
    for i in range(0, len(list)):
        coords = list[i]
```

```
            canvas[coords[0]][coords[1]] = "."


# Given an origin and radius, find all coordinates within the canvas that would
be inside of the circle
# [[1, 0], [2,0]]
# ((x1 - x2)^2 + (y1 - y2)^2)^1/2
def find_coordinates_within_circle(x, y, r):
    result = []
    for i in range(ROWS):
        for j in range(COLS):
            coordinates = [i, j]
            eucledian_distance = ((x - i)**2 + (y - j)**2)**(0.5)
            if (eucledian_distance <= r):
                result.append(coordinates)
    return result


# Find the y = ax + b form of the line, find coordinates that fit with that
line, and print out those coordinates
def render_line(x1, y1, x2, y2):
    if x1 > ROWS or x1 < 0 or y1 > ROWS or y1 < 0 or x2 > ROWS or x2 < 0 or y2
> ROWS or y2 < 0:
        print("Error with coordinates")
    run = x2 - x1
    rise = y2 - y1
    slope = rise / run

    # (x1, y1)
    # (0, y_int)

    # y = slope * 0 + y_int
    #
    y_int =

    canvas[x1][y1] = "."
    canvas[x2][y2] = "."

    point1 = [x1, y1]
    point2 = [x2, y2]
    list = []
    while point1 != point2:
        point1[0] += run
```

```
        point1[1] += rise
        list.append(point1)


    for i in range(0, len(list)):

# Test Cases
render_circle(20, 20, 5)
render_circle(0, 0, 0)
render_circle(-1, 0, 3)



for row in canvas:
    print ''.join(row)
```

Swift Navigation

Notes: I originally applied to "Software Engineer: Infrastructure" but after the first phone screen they decided that I would be a better fit for "Software Engineer: Receiver Test/QA Infrastructure" so instead of one initial phone screen and a second phone screen with a technical question, I just re-did the initial phone screen with the manager for the testing team.

I feel like I asked a normal amount of questions during the phone screen, but afterwards I emailed the interviewer a bunch of questions about the role / engineering questions about the product / company culture / business strategy that I came up with after a bit more thinking. Lol apparently asking hella questions is a big positive if you're interviewing for a job developing testing software.

Application Code Sanity-check
A variation of fizzbuzz. The actual question is super easy, just demonstrate that you can write nice code with unit tests and documentation.

Phone Screen
The usual talking about projects on your resume (decisions, challenges, technology used, etc), what you do at your current role, what are you looking for in your next job, generally explaining what the company does and answering questions.

On-site
Probably the most comprehensive yet practical on-site interview process I've been through. You have to give a slide deck presentation to a group consisting of the team you're interviewing for + anyone else who will be interviewing you about your background and give a technical walkthrough of a project you've done before (10-15 minutes total for presenting, then Q&A afterwards). Then 4 1-on-1 technical interviews, plus lunch (which is basically like "informal culture fit interview") and then a final session with one of the founders to just talk about more business-related things.

Systems design — Interviewers outline what they would want in a simplified version of the testing pipeline, and with a decent amount of prodding I had to design it (at a really high level). What information you would

need to have/keep at each of the stages they described, how you would organize it in a database, different ways to handle queueing jobs, that kind of thing.

Data analysis — You get a spreadsheet of actual data put out by a stationary test session of their GPS receivers, and then you have to derive how to calculate one of the metrics they use to measure accuracy. It's relatively hand-holdy and you don't actually have to know how to work with spreadsheets or pandas, you can just tell the interviewer what you want to do e.g. "subtract that column from that column and square it" and they'll handle it. Mostly to see how you reason through data analysis stuff and whether you have a basic grasp of statistics and linear algebra.

Web programming — Implement a super simple REST API in the framework of your choice. Basically just checking that you can actually do irl programming. You do this on your laptop and go through all the debugging and stackoverflow-ing. Then explain what other sorts of things you would need to consider in an expanded/irl version of it (tests, security, scaling).

ISO Standard programming interview — Warm-up: Given a binary tree, print the leaves left-to-right. Full question: Given a binary tree, write a function to return whether a left-to-right traversal of the leaves is ordered. Basically do some sort of iterative DFS since you don't want to search the whole tree if, say, you could identify that the (second leaf < first leaf).

Where this notation is (Root LeftNode RightNode) only because I'm too lazy to write these to look like trees.
(A (B (C D)) (E Null)) --> Leaves are [C D E] --> True
(A (B (C E)) (D Null)) --> Leaves are [C E D] --> False

Tealeaf

Technical Interview 1:
1. What is a reflection in Java?
2. What are the different OS layers?
3. Can a class extend multiple classes?
4. How do you make a class extend more than one thing?
5. What is the difference between blackbox and whitebox?
6. Does a class that extends an abstract class need to implement all the abstract methods?
7. What are generics?

2nd part is onsite, but I'm studying abroad, so meh.

TE Connectivity

2 hour on-site interview (San Jose office)
March 12, 2014

Interview Breakdown:

- 5 minute chat with the recruiter
- 20 minute phone call with the hiring manager (in Minneapolis)
- 45 minute interview with site manager & one engineer + campus tour
- 45 minute interview with two other engineers

This was a super casual interview. I don't think any of the interviewers really cared about the interview so we spent most of the time chatting about the company, the product, unix, C++, and the stuff on my resume. They didn't have anything prepared and they only asked one technical question. Here's that question (and its followups):

What is a makefile? What do you use a makefile for? How does make know which files it needs to compile every time?

TubeMogul

Super chill guys. 30 minute interview consisting of 15 min. general behaviorial stuff and 10-15 min. of technical questions relating to finding duplicates.

1. What are your long term plans? (Apparently I was the only interviewee the entire day who mentioned startups)
2. What kind of products do you want to work on? Give an example of one that has value. (I gave a personal example with Facebook and then we dissed FB messenger lol)
3. What technologies/stack do you want to do at TubeMogul?
4. Write an algorithm to find duplicates in an array. Output the duplicate values. What is the runtime?
5. Write an algorithm to find duplicates in two arrays (overlap but with one element only). What is the runtime?
6. Write an algorithm to find all sets of 1, 2, or 3 values in an array that sum to 3 and return them. No negative numbers. What is the runtime?

final round was 3 30 minute interviews with different ranked engineers

1. You're given either a file or a directory as an input. Write a function that goes through the input and any subdirectories if they exist and returns a 2-d structure that contains all the files that have the same content grouped together. (key points are dfs, hashing, and optimizations via file size and maybe checksums)
2. Implement a queue in every way you can think of. Give the advantages and disadvantages of each implementation. (singly/doubly linked list, array)
3. Given a String url input, write a function that gets all possible links on that page and all subpages. (use dfs/bfs with visited set since some links will link back to home page or other previous pages, DOM model with jQuery command)

More On-site questions:
1. Convert a sorted list of numbers into a BST.
2. Write a multipurpose data structure that can be used as a stack or a queue.

Tumblr (Android Internship)

Phone questions:
1. What's the difference between a hashset, hashtable and hashmap
2. Design chess (using classes/methods)
3. What data structures do you mostly use when programming and why would you use them

Shared code screen question:
1. Implement findViewById
   Answer: Do BFS on the root Viewgroup until you find the id you're   looking for
....crap, what happened to the rest of this??

Twitch
Interviewing for Pipeline Engineer

Phonescreen 1:
Asked about my background and what I did, etc.
online code editor: Implement a simple Fibbonaci number finder in both iterative and recursive fashion.
online code editor: given a root to a binary tree, and 2 numbers known to be in the binary tree, print all values in the binary tree including and between those values. (very straight forward if you don't overthink it)

Phonescreen 2:
Online code editor: write a fairly optimized (nth) prime number finder.
Online code editor: Write the pseudocode on how you would write Map and Reduce functions for a continuous stream of incoming requests modeled by ID and session. Multiple sessions can map to the same ID and the reduce should return an ID with a list of all the sessions associated with that ID.
look into mr. Job, a python framework for mapreduce (very simple, but led into…)

Followup into: Given X requests per second, Y Bytes per request and A Gbps IO limit and B Gbps Network saturation limit per node, how many nodes would you need to spin up in a cluster to not have a queue of requests building up? Open ended Question. Can talk about HDFS, Map Reduce data paradigms, Optimizing your prior MR functions, etc.

 Onsite 1 (and Final)
Interview 1: Talking with data analyst. Spent all the time talking about my resume and the projects I did. Technologies I used and the design methodology i followed, along with my experience in certain fields gained from said fields.

Interview 2: Onsite, but through a computer to an offsite employee (worked on the team that I was applying a position for). asking if I was familiar with C, which I am, but prefered python for interview questions. Asked me to print a checkerboard made of 1's and 0's where you can pass in, num of rows, columns, and height and length of each square. Some python string manipulation made that easy.
Followed up with a simple fibbonacci question asked in a prior phone screen

Two Sigma

1. You are to write a program that takes a list of strings containing integers and words and returns a sorted version of the list

The goal is to sort this list in such a way that all words are in alphabetical order and all integers are in numerical order. Furthermore, if the nth element in the list is an integer it must remain an integer, and if it is a word it must remain a word.

Input:

The input will contain a single, possibly empty, line containing a space-separated list of strings to be sorted. Words will not contain spaces, will contain only the lower-case letters a-z. Integers will be in the range -999999 to 999999, inclusive. The line will be at most 1000 characters long.

Output:

The program must output the list of strings, sorted per the requirements above. Strings must be separated by a single space, with no leading space at the beginning of the line or trailing space at the end of the line.

Constraints:

The code you submit must take input from stdin and produce output to stdout as specified above. No other output is permitted. You can assume the input will be valid. Feel free to use standard libraries to assist in sorting.

In the examples below, the text "Input:" and "Output:" are not part of the output, and neither are any blank lines.

Input:

1

Output:

1

Example 2:

Input:

car truck bus

Output:

bus car truck

Example 3:

Input:

8 4 6 1 -2 9 5

Output:

-2 1 4 5 6 8 9

Example 4:

Input:

car truck 8 4 bus 6 1

Output:

bus car 1 4 truck 6 8

2. Oh no! Disaster has struck some of ACME's redundant data centers. The administrators have managed to restore backups, but some data sets are still missing from some data centers. Fortunately, every data set can be found at least once in one or more of the data centers. However, before ACME can resume normal operations, it needs to ensure that each data center has a copy of every data set.

Your goal is to help ACME resume normal operations by writing a program to synchronize data sets between data centers using as few copies as possible.

Input:

The first line of input will contain an integer between 0 and 999999 inclusive, representing the number of data centers.

Following that will be one line of input for each data center. Each line will contain an integer from 0 to 299 representing the number of data sets at that data center, followed by a space and space-separated list of data set ids currently present at that data center. Data set ids are each an integer between 1 and 999999, inclusive. Each line will be at most 999 characters long.

Data set ids are not necessarily consecutive. The list of data sets will not be in any particular order.

Output:

The program must output an optimal data set copy strategy to ensure that every data center has a copy of every data set. Output one line for every copy instruction.

A copy instruction is of the form <data-set-id> <from> <to>, where <data-set-id> is the data set id, <from> is the index of the data center the data set will be copied from (1 = the first data center), and <to> is the index of the data center to copy the data set to.

When there are no more copy instructions, the program must output the word "done" on a line by itself.

There is often more than one correct output with minimal number of operations for a given input, and any output that satisfies the requirements is valid.

Constraints:

The code you submit must take input from stdin and produce output to stdout as specified above. No other output is permitted. You can assume the input will be valid.

Example 1:

Input:

4

3 1 3 4

3 1 2 3

2 1 3

3 1 4 2

One Possible Correct Output:

2 2 1

4 1 2

2 2 3

4 4 3

3 1 4

done

Example 2:

Input:

2

2 1 2

2 2 1

Output:

done

Example 3:

Input:

3

5 1 3 4 5 7

2 1 3

1 2

One Possible Correct Output:

2 3 2

2 3 1

1 1 3

4 1 2

5 1 3

5 3 2

4 2 3

3 1 3

7 1 2

7 1 3

done

Uber

Someone Else's Interview

```
import java.util.HashMap;

/*
In the game of Boggle, you are given a grid of letters (except Qu which counts as 1
letter). Find all the available words in the grid.

You are given a Dictionary util
    Dict.isWord(String);
    Dict.isPrefix(String)
isPrefix is inclusive isWord where if isWord is true, isPrefix is true
Words can be formed by letters that are adjacent to each grid character, This includes
sideways and diagonal.
e.g. SIP is a valid word.

  ---------------
|Qu | V | C | P |
|S  | D | K | I  |
|A  | E | S | L  |
|B  | L | O | T  |
  ---------------

SLI prefix
SLIP Word

*/
```

Software Engineer - Supply Team - Backend (full time)

Technical Phone Screen

Was referred by a friend who worked there, so I wasn't applying to a particular team. Before this call there was a chill HR phone screen where they ask all the usual things (what do you do, what are you looking for, why are you leaving, are you a US citizen, etc) and they decided which team I would be interviewing with.

I was interviewed by a guy who said he's been working on this team, but basically they handle the parts of the app/site that tell the drivers where they should go) for about four and a half months. Few minutes where he talked about what his team does, a few minutes to discuss my background (their hiring/jobs platform was apparently down so he hadn't been given a copy of my resume so it wasn't terribly in-depth), short technical question (can do it in language of your choice) on hackerrank's codepair platform, then time for you to ask the interviewer some stuff.

Probably would be good for you to have some suggestions of things that you're interested in or think could be improved about the app so you don't have awkward "ummm"-ing there like me. Also asked what technical challenges I've encountered in my current job and how that was solved, what I'm looking for in my next job, why I'm leaving my current job. Basically all the expected nontechnical questions.

Also be able to explain the run time of your solution, and also how you could improve the is_a_word function. I said that it would be more efficient to store things in a set so that the lookup time would be constant (then he asked me to explain why it would be constant, so I talked about the high-level concept of hash functions being constant time calculations vs iterating through everything in the collection), or barring that store everything in a trie (worst-case depends on the longest word in your collection of words).

```python
# Write a function that, given an input string and the function is_a_word below,
# will return a list of all substrings of its input that are valid words. We want
# no repeats in this output list.

# "thisisaword" => ["this", "hi", "his", "is", "word", "a"]
# "isisis" => ["is"]


def is_a_word(test_string):
    """Check if test_string is a word."""
    return test_string in ["this", "hi", "his", "is", "word", "a"]

# My answer
def get_substring(s, n):
    substrings = []
    for i in range(0, len(s)-n+1):
        substrings.append(s[i:i+n])
    # or just return [s[i:i+n] for i in range(0, len(s)-n+1)]
    return substrings

def get_all_substrings(s):
    substrings = []
    for i in range(1, len(s)+1):
        substrings += get_substring(s, i)
    return set(substrings)

def find_words(s):
    return [substr for substr in get_all_substrings(s) if is_a_word(substr)]

# print(get_all_substrings("word"))
# print(get_substring("thisisaword", 1))
# print(get_substring("thisisaword", 2))
```

```
# print(get_substring("thisisaword", len("thisisaword")))
print(find_words("thisisaword"))
print(find_words("isisis"))
```

Typical soft questions — talk about a project you've worked on and are proud of / what are some challenges you had with that, what is REST, talk about a time when you disagreed with someone about something technical, why here / what are you looking for, etc etc. Hiring Manager engineer was literally just reading off a list of questions. Overall everyone was really friendly and seemed to have their shit together.

First interview: Convert a binary search tree into a doubly-linked list (basically node.left becomes the previous item in the list, and node.right should point to the next thing).
Second interview: lunchtime interview with hiring manager, so talked about experience and general job stuff but no technical question. Food was pretty good.

Third interview: Same sorts of general questions. Tech question was given a file that has words delimited by spaces and a word that you want to check is in that file, how would you implement this? Eventually you get around to suggesting that if the input file is sorted, you can use binary search with the added complication of having to find the entire word at each iteration of binary search.

Fourth interview: Jk the rest of the interviews (five were scheduled) didn't happen. After the third one the recruiter came in and said they had already decided not to give me an offer because I didn't have the experience they were looking for. I'm a little bit confused about this because they had my resume, which stated that I had <1 year irl software engineering experience and I'm not sure what they were hoping I would reveal during the interview that would pass their 'enough experience' threshold. They told me that they were interested in keeping in touch I should definitely apply in several months or ~1 year.

Veeva Systems

1st technical round:

A lot of Java-related questions, let's see if I remember:
1. What is the difference between a set and a list? (a list is sorted)
2. Can you explain what static means, both with methods and variables? Give an example.
3. What is the difference between String, StringBuffer, and StringBuilder? (you only need to know one of StringBuffer or StringBuilder). (if you add together strings, another instance of the object is created every time, but if you use StringBuffer, all strings are stored in an array and then added together at once at the end to create the object).
4. StringBuffer is synchronized and StringBuilder isn't. What is synchronization? (synchronized keyword in Java means that something is thread-safe)
5. Given 2 arguments, each an array of integers, return a list with the duplicates. (use 2 pointers and a while loop to iterate through in O(n) time)
6. Given 2 strings, find if one is an anagram of the other. (use hashing. sorting is slower but uses less memory I believe...)

2nd technical round:

1. You have 2 lists. Find if one list has the exact same contents as the other.

First check if the length of both lists is the same. Then use hashing to place the contents of one list in a HashMap. Then go through the other list - if you ever find a unique key, return false. Also return false if the value of the key-value pair you're accessing in that iteration goes below 0. Then if you can manage to go through the entire list without returning false, return true.

2. The hardest question of all time - he blocked access to the question after the interview so let's see if I can still remember…

The following 2 things are implemented already for you: a HashMap of type <String, List<String>> that contains as its key a column letter, A or B for example. The value is a list of its dependencies, or other columns that depend on it. There were 4 key-value pairs in his example, but I only remember 3: A | B, C; B | A C D; E | A; and one other. You're also given a setValue(String column, Integer value) function that has already been implemented, though you can't make assumptions about its exact details. You should only use this function once on each column.

Write a RECURSIVE function that takes a String column and Integer value and sets that column to the the input Integer value. Then the function should go through and set that column's dependencies to the value + 1. It should recursively go through and keep setting the column's dependencies' dependencies to the value + 1 + 1, but be sure to do this only for non-encountered dependencies to avoid an infinite-loop.

You have to use two global Lists, which he didn't tell me I could use until I asked for one and he suggested the other. One List stores the columns you have already visited, while the other stores the columns you need to visit in the subsequent recursive call.

You also need 2 for loops, the first will go through and take the column's dependencies and set them to value + 1 if that particular dependency is not in the first List as described above. It will also add the dependency to the first and second Lists if it is new. The second will go through the same dependencies and will make recursive calls if they are contained in the second List, removing them right before making the recursive call.

Took me almost an hour of interview time to solve with 3+ hints from the guy. Afterward he went into a passionate explanation about how I just implemented a very similar algorithm to BFS without a queue and how tons of people fall into the traps of infinite recursion, getting the wrong value + 1 values set for each column, etc. My brain just hurt the entire time...

Veeva Systems (tj)

1-hour phone interview
February 4, 2015

Note: I heard on the onsite afterward that apparently they do a 30-minute phone screen first, however Paul waived it because I talked with him for like 15 mins at the career fair. In the phone call, spent the first 15 minutes talking about random stuff (Brian: they want to check you're not an asshole and have social skills) and where my interviewer worked previously. Then while chatting he asked two questions.

Q: In Java, what is a List and a Set? When would you use them, and what are the differences between them?

A: Sets and Lists are both interfaces. Lists are used for storing an ordered list of objects (allowing repeats), while Sets store unordered objects without repeats. Also talk about the interface similarities/differences.

Q: Programming question. Given the following class definition, implement `getOldestSisters()`.

```
public class Person {
  Person mother;
  Person father;
  Gender gender;
  Integer age;
  List<Person> children;

  enum Gender {
      MALE,
      FEMALE;
  };

  /* Returns a List of this Person's oldest sisters. A Person is
   * considered a sister if she is FEMALE and also shares a mother or
   * father. If there is no oldest sister, return an empty list.
   * Do not return the calling person. */
  List<Person> getOldestSisters() {}
}
```

A. There are surprisingly many cases to handle.

```
List<Person> getOldestSisters() {
    HashSet<Person> sisterSet = new HashSet<Person>();
    int oldestAge = Integer.MIN_VALUE;

    if (this.mother != null && this.mother.children != null) {
      for(Person per : this.mother.children) {
          if (per.age != null && per != this && per.gender == Gender.FEMALE &&
              && per.age.getInt() > oldestAge) {
            oldestAge = per.age.getInt();
            sisterSet.add(per);
          }
```

```
        }
    }

    if (this.father != null && this.father.children != null) {
      for(Person per : this.father.children) {
          if (per.age != null && per != this && per.gender == Gender.FEMALE &&
              && per.age.getInt() > oldestAge) {
            oldestAge = per.age.getInt();
            sisterSet.add(per);
          }
      }
    }

    LinkedList<Person> sisterList = new LinkedList<Person>();
    for(Person per : sisterSet) {
      if (per.age.getInt() == oldestAge) {
          sisterList.add(per);
      }
    }
    return sisterList;
}
```

Thoughts: Interviewer gave huge hints on how to do this before the interview (by asking questions about what is a List and a Set). Also, the big thing where I messed up was I forgot to read the question. I forgot the `per.gender == Gender.FEMALE` and `per != this` lines, which without don't give the right answer. Also, this probably should have been done with Comparables ... oh well.

Veeva Systems (tj)

5-hour onsite interview (Pleasanton, CA)
February 11, 2015

Took the Dublin/Pleasanton BART to end-of-line and then walked like 15 mins to the Veeva office. Entered the reception on the top floor and spent 30-min with Paul O'Flynn: Senior Manager. He gave me a tour of the office and talked about the three divisions of the company. Completely nontechnical, no soft-questions so I think this was mostly a culture fit. Afterward, met Chris Rink: Software Engineer who gave me a tour of the Vault software product. This chat was also completely nontechnical, with no soft questions. Then I had lunch with Lunch with Kent Nguyen: Associate Software Engineer and Sayaka So: Software Engineer at nearby Sendo Sushi. They didn't have a set script and it was really quiet at times so TBH I don't think they were prepared, probably Ashley (the recruiter) told them to go out. After lunch, I had an interview with Tudo Nguyen: Principal Software Engineer. He asked two technical questions:
Q1: Program Fizz-Buzz in the programming language of your choice
A1: Should be fairly easy to write this out .. but the reason it's asked is because it's super easy to make a hard to find mistake. For example, the order of the if/else statements matters a lot, and you

have to understand how range works in python.

```python
def fizzbuzzGenerator():
  for i in range(1,101):
    print(fizzbuzz(i))
def fizzbuzz(val):
  assert(1 <= val <= 100)
  if val % 15:
    return "fizzbuzz"
  elif val % 3:
    return "fizz"
  elif val % 5:
    return "buzz"
  else:
    return str(val)
```

Q1Followup: How would you write tests for this code?
A1Followup: Create an instance of the test code and have a bunch of
`assert_equal(fizzbuzz(57), "fizz"); assert_exception(fizzbuzz(0), AssertionError)`. Make sure to handle every possible case. For the `for` loop, I wasn't sure how to test it without examining stdout but he suggested mocking it then checking it's called 100 times.

Q2: In 15 minutes, talk about how would you design a database schema for an IMDB for TV-Shows.
A2: I haven't taken CS186 (the database class) so I let him know but gave a shot at it:
TV-Show: list(Season)
Season: list(Episodes), title, dynamically generated/cached list of Actors/Producers/Writers
Episode: length, title, summary, list(Actors), list(Producers), list(Writers)
Person: Could be either a Actor, Producer, Writer; has a name, DoB, height, bio

I had an issue trying to describe how to dynamically generate a cached list of Actors/Producers that is recalculated every time an Episode is added but I don't know the database term for this so I let him know and he seemed fine with this.

Afterwards, I had an interview Dan Soble: Vice President, Vault Engineering and he asked me a lot of questions about my resume and what I was looking for at a workplace. He asked my expected compensation then pitched the company for 5 minutes. After that, I had a technical interview with Jonathan Huang: Director of Engineering, who asked the same question as Brian:
Q: Given a `HashMap<String, List<Integer>> dependencyMap` that has the following values:
A -> [B, C]
B-> [A, C, D]
C-> [E]
E -> [A]

Write a RECURSIVE function called `setValue(String col, int val)` that sets (1) col to val (2) the dependents of col to val+1 and (3) recursively calls this function on the dependents of col
For example, `setValue("A", 1)` should set:
A -> 1
B -> 2
C -> 2
D -> 3
E -> 3

A: This is just BFS value propagation done recursively instead of iteratively. The dependencyMap may have cycles, but (assuredly) no self cycles so all you have to do is make a global HashMap<String, Integer> that saves the current value of the BFS propagation and traverse the graph in a DFS way but stop if you already have explored a node previously. However, if you decrease the value of a node, this node needs to be reexplored. This question is very confusing, however if you stare at it long enough and keep thinking of a Dijkstra's algorithm using BFS then it should be doable. Also, make sure you understand how Generics work in Java and how autoboxing of ints/Integers is handled in Java. I took over the allocated time but Jonathan wanted to keep talking so he pushed the next interview back.

Finally, I had a talk with Ashley Chou: College Recruiting Manager, and we talked about logistics (where I currently lived, and if I would move to Pleasanton). I made a mistake: do not bring up things that could compromise your ability to join the company at the time requested (I brought up a Masters), so she brought in Jonathan and he talked about his time at Berkeley; IMO he did a better job than Dan at pitching the company. I ended up over time so the onsite ended after this talk.

Thoughts: Every single person asked "Do you have any questions for me?" at the end of each conversation but I asked questions during the interview so most of my questions were already answered? So I'm not sure what I was supposed to expect (advice appreciated)

WealthFront

backend

Given:
a list of numbers: [3, 1, 2, 300, 1000, 400, 200, 750, -150, 300]
a goal number: 600
find all pairs (2 numbers per pair, it's not the knapsack problem) of numbers that sum to the goal number.

```
static class Pair<A, B> {
    public final A left;
```

```java
    public final B right;

    Pair(A left, B right) {
        this.left = left;
        this.right = right;
    }
}


static List<Pair<Long, Long>> findSummingPairs(List<Long> numbers, Long goal) {
    List<Pair<Long, Long>> answer = new List<Pair<Long, Long>>();
    for (int i = 0; i < numbers.length(); i++) {
    for (int j = 0; j < numbers.length(); j++) {
        if (numbers[i] + numbers[j] == goal & j != i) {
            answer.append(Pair(numbers[i], numbers[j]));
        }
    }
    }
    return answer;
}


static List<Pair<Long, Long>> findSummingPairs(List<Long> numbers, Long goal) {
    List<Pair<Long, Long>> answer = new List<Pair<Long, Long>>();
    HashMap<Long, int> map = new HashMap<Long, int>();
    for (int i = 0; i < numbers.length(); i++) {
    map.put(numbers[i], i);
    }
    for (int j = 0; j < numbers.length(); j++) {
    long difference = goal - numbers[j];
    if (map.contains(difference) & j != map.get(difference)) {
        answer.append(numbers[j], difference);
    }
    }
    return answer;
}
```

[200,400], 600 => [(200,400), (400,200)]
[750, -150], 600 => [(750, -150), (-150, 750)]
[300, 300], 600 => [(300, 300), (300, 300)]

Test case to deal with overflow with adding long numbers
[300], 600 => []
[200, 200, 400, 400], 600 => [(200, 400), (200, 400), (200, 400), (200, 400), (400, 200), (400, 200)
, (400, 200) , (400, 200)]


```
class Employee {

    private final String[] titles;

    Employee(String[] titles) {
        this.titles = Arrays.copyOf(titles, titles.length);
    this.titles = titles;
    }

    String[] getTitles() {
        return titles;
    }

}

String[] t = new String[] { "Foo" };
final Employee e = new Employee(t);
t[0] = "Bar";
```


Front End questions
```
function getElementsByClass(klass, node) { //
  var node = node || document.body;  //1
  var found = [];  //2
  if (hasClass(node, klass)) { //
      found.push(node);
  }
  for (var i=0; i<node.children.length; i++) {
      found = found.concat(getElementsByClass(klass, node.children[i]));  //
  }
```

```
    return found;
}


function hasClass(node, klass) {
  for(var i=0; i<node.classList.length; i++) {
    if (node.classList[i] === klass) {
      return true;
      }
  }
  return false;
}




root = document.body


node.children returns a collection of child nodes of the given node
node.classList - returns a token list of the class attribute of the node




div.classList = ['this-is-a-class', 'this-is-too']

<div class='this-is-a-class this-is-too'> //b.classList =
      <p>hello</p>
</div>
[ , ]


              a //this
              b     c    // these ?    [b, c]
      c     d    s     a
```

2. In HTTPS, what are put and delete?

3. In CSS, what's the box model? Describe the border, padding, and margins. How do you calculate the width of an element?

4. In the context of positioning, what are inline, relative positioning, and absolute positioning?

5. In JavaScript, what does scope mean? What types of scope are there?

6. What is a closure in JavaScript? What are some concerns behind a closure?

7. What is the DOM?

Workday - Data Scientist

Explain to me what logistics regression is and what assumptions there are in the model.
What is supervised vs unsupervised learning?
Derive least squares regression and logistics regression (MLE)
What is p-value, null hypothesis, critical region, etc

Yelp

APM Questions

If you were a PM on Yelp's team that gets business owners to claim businesses on Yelp, how would you get owners to sign up?

What happens when you enter yelp.com on your browser?
How does caching relate to this process?

SQL question

table: User

- id (auto-inc, pk)

- name (varchar)

- is_elite (bool)

table: review

- id (auto-inc, pk)

- user_id (fk)

- biz_id (fk)

- time_posted

table: biz

- id (auto-inc, pk)

- name (varchar)

Give me the username of the user who wrote the first review for a business named "foo"

id of biz == foo
look in review, grab all reviews that have biz_id
Sort by time_posted
Select top 1 review
grab that review's user_id

Yelp - Data Mining Engineer
January 2015
tl;dr *I hope you paid attention in CS188.*

Pulse-check
Online test asking random programming trivia, like "what UNIX command would you use to search for a term within a file". Should be really easy, don't mind if you can't get 100%.

Phone screen (shared code editor)
Given a function check_guess() that contains a secret_number, which takes a guess and returns -1, 0, or 1 depending if it is too low, correct, or too high, write a function that will return the secret_number. Range is whatever the limit of floats is.
Answer: General idea is to exponentially increase (or decrease. Can guess 0 first to see which way to go.) guess until check_guess says you went over (e.g. if the number is 1234, guessing 1,000 is too low, then going to 10,000 is too high), then do binary search between those two points.

On-site
Tour of building from recruiter, then 4 interviews with people from different data mining teams (ads, spam, search, etc). Some of them asked 2 questions.

Talk about anything Python vs. Java. (I'm not sure what they were looking for on this one)

Working with large files. You have a file API that has the methods:
f = File("filename")
f.seek()
f.getCharAt(int) # this will explode if you try and get a char out of range
f.size()

- Given a file (too big to load into memory) that is a dump of tweets (just the text) separated by newline characters, write a program that will return a random tweet
    - you can use rand(), and just print text to sysout.
    - Make sure that you give the first tweet a chance, and that you won't go over the file if you end up on the last tweet.
- Follow-ups: Run time, what could happen if it's other types of docs, and not just tweets? (Chance of returning a doc is directly proportional to how big it is. Not really a problem with tweets since they're all about the same size.)

Return the largest substring that does not have repeating characters.

Decision Trees
- Implement the classify() method for a trained (binary) decision tree. Define any other classes or methods that you need.
- Implement the split() method used in building a decision tree, where the inputs are a matrix of floats, where each row is a data point and each column is a feature, and a matrix of labels for each of the data points. Assume you have a method for calculating entropy/information. The interviewer went over the general implementation of a decision tree and what exactly entropy and information gain are.

How would you implement a spam detector for Yelp's forums? (Read as: explain how you would implement a Naive Bayes/LG/etc classifier.)
- What input should it take in? You can talk about things like whether the user that posted is active or if the post contains similar words to the OP, but I think the main thing the interviewer wanted to hear was "the wordcount vector for the message" since the typical NB example is on bag-of-words doc representation.
- How would you test/tune it? Basically get all the keywords (and explain what they are / how they affect each other): accuracy, precision, recall, f-score, ROC/AUC, cross-validation. Explain what kind of scores you would want for these, e.g. that accuracy is pretty useless here because most posts will not be spam.
- What should you do if the data grows too big to evaluate the features for the whole document? Only calculate the features previously determined to have the most impact (can't remember what the term for this is).

Given a matrix of 0's and 1's, where each row is sorted so that all the 0's come before the 1's, find the row that has the most 1's. (Could just sum the across each row, but looking for O(m+n) answer).
Example input:
0 0 0 1
0 1 1 1
0 0 0 0
0 0 0 1

General idea: start in upper right corner of matrix. Scan left until you hit a zero or the 0th entry of the row (if you hit the beginning you can just stop and return that row since that's the max). Go down a row. If that column in the new row has a 1, scan left until you hit zero, update the longest row. Do this for all the rows.

Yelp - Software Engineer (tj)
20 Minute Phone Screen
February 20, 2015
Met the Engineering Recruiter Michael Jaffe at the Career Fair and he reached out to me. Simple 20 minute phone conversation about the process, including a 30 minute HackerRank coding screen afterwards. It was 15 one-word answer questions (like: what port is HTTPS, what is 3 in octal, what does TLS stand for) followed by a programming question "given a list of words, determine whether they are all an anagrams of each other". The one line solution is:
```
len(set().union(["".join(sorted(x)) for x in foo])) == 1
```

45 Minute Skype Interview
March 18, 2015

Q1: How much do you know about the web?
Q2: Do you like Javascript?
Q3: Here's an issue Yelp ran into. Yelp pushes code to production several times a day but sometimes there would be issues not caught by testing. Clients would try a new feature and run into Javascript exceptions in their browser and they would not be discovered. Server side exceptions produce logs so they are easy to detect and fix. How might you design a system to detect client Javascript errors and report them to the server?
A3: Didn't really know Javascript so I gave an answer that worked but didn't use all the features of Javascript. Important to understand HTTP, like when to use a GET vs POST request and how servers handle incoming connections.

Impressions: I think they understood I could program but didn't know Javascript so they scheduled a second interview with more generic interview questions.

Another 45 Minute Skype Interview
March 24, 2015
Since the last interview was Javascript specific they scheduled another that was all programming.
Q1: Given two strings, determine if the first string is an in-order substring of the second. Example:
```
# str1 = "BAD"
# str2 = "Buy the stuff.cAn't believe it.maDness."
# str3 = "Buy Donuts A"
# is_bad_string(str1, str2) => True
# is_bad_string(str1, str3) => False
```

A1: Couldn't find this question online. Quite surprised. A couple edge cases, namely if str1 or str2 is empty. If str1 is empty it should match all strings.

```python
def is_bad_string(str1, str2):
    if str1 is None or str2 is None:
        return True

    i, j = 0, 0
    while i < len(str1) and j < len(str2):
        if str2[j] == str1[i]:
            i += 1
        j += 1

    return i != len(str1)
```

Q2: Instead of returning True or False, return the number of matching strings.

str1 = "ab"
str2 = "aabb"
count_creation(str1, str2) => 4, because we can match "a_b_", "a__b", "_ab_", "_a_b"

Just like count change. Note that the order of the order of the if statements is very important since we assume the empty string always matches.

```python
def count_creation(str1, str2):
    if str1 is None or str2 is None:
        return 0
    if len(str1) == 0:
        return 1
    if len(str2) == 0:
        return 0

    if str1[0] == str2[0]:
        return count_creation(str1[1:], str2[1:]) + count_creation(str1, str2[1:])
    else:
        return count_creation(str1, str2[1:])
```

Q3: Implement sqrt(x)
A3: Straight out of leetcode, google how to do it. use binary search

Yelp Onsite
April 10, 2015
Four 45-minute interviews

First interview with the manager of search
Q0: My interview was for Software Engineering New Grad so it was fairly open on what I would work on. In a pretty informal way they asked what kind of software position I was looking for. Do you prefer backend, frontend, data analysis, machine learning, etc. If you are already interviewing for a specific team I don't think they'll ask this.
A0: Try to think about this beforehand. I said no strong preference but preferred backend, so I think

he changed his interview question to reflect this.

Q1: In unix there is a command called fortune that prints out a random fortune. Fortunes are stored newline-delimited in a text file saved at /etc/fortunes. Explain how you would print a random fortune from this file.
A1: Read the file into memory, split() by newline, store the file in a list, then return a random element from that list.
Q2: Now instead of printing fortunes, assume we have a very large file storing a bunch of tweets and we want to print a random tweet. The file is very large (several terabytes) and tweets are stored newline-delimited. We can only access the file with the following methods:
- `char fgetc()` => get character at current file pos, increment file pos
- `void fsetpos(long pos)` => set file pointer to pos
- `long flen()` => returns file length

Write code to return a random tweet.
A2: Since tweets are at most 140 characters, I asked how the tweets are stored. He said that they are newline delimited. I asked does the output need to be truly random. He said it's fine if some tweets appear more often than others. Basically, seek to a random index and proceed until you find a newline. Return the tweet following that newline.

```
def randomtweet(file):
    startindex = random.randint(file.flen())
    file.fsetpos(startindex)
    while file.fgetc() != "\n":
        startindex += 1
    endindex = startindex
    outstring = ""
    while endindex < file.flen():
        ch = file.fgetc()
        if ch == "\n":
            break
        else:
            outstring += ch
            endindex += 1
    return outstring
```

Q3: The code above has two bugs. 1) It may fail in a specific case and 2) Not every string in the array has a chance of appearing. How can we 1) fix this and 2) make it so every string has a chance of appearing?
A3: It will crash if the random pointer lands in the last tweet of the file. Also, In this code, the first string in the file will never appear.

Q4: Rewrite the code so this fixes both issues.

A4: If we reach the end of the file, wrap around. Not too hard.

Q5: Interviewer went on a spiel on why he decided to work for Yelp. Then he asked: Why do you want to work for Yelp?
A5: This is probably the most important question. If you do not give a good answer, you won't get the job. You want to talk about growth and how Yelp is in a really great position to do so. Ask me in person for more info.

Second Interview with someone from product
Quick backstory: the Yelp backend was written in Python, so their interview questions are Python related.
Q1: Write a function "nextfib" so that each time you call it, it prints out the next fibonacci number. Write this in two different ways. It should work as follows:

```
>>> foo = fibgen()
>>> foo.nextfib()
0
>>> foo.nextfib()
1
>>> foo.nextfib()
1
>>> foo.nextfib()
2
>>> foo.nextfib()
3
>>> foo.nextfib()
5
```

A1: Use a generator function (python specific thing). Also use a class.

Third Interview with someone from revenue
Q1: What's the slowest possible sort you can design?
A1: First I said sort it, then sleep() for a year, and return. He then asked how would you make the most computationally complex sort. Didn't really know what he wanted so I described bogosort, he said make it slower, then I said bogobogosort. He seemed ok with this answer.
Q2: So bogobogosort.. can you estimate the runtime in big O notation?
A2: (This is a dumb question, since it's a randomized algorithm you can't prove an upper bound. You can prove an expected upper bound, but you cannot prove it will always sort in a bounded amount of time). I said expected $O(n!^{n!})$ but likely lower than that. I don't think he knew either.
Q3: Coding question, straight off leetcode. Given a list of anagrams, group like anagrams together and return as a list.
A4: The key to understanding this question is how to group anagrams. The trick is as follows: if string `str1` and `str2` are anagrams, then `sorted(str1) == sorted(str2)`. If you did not know

this before starting the question, you will have a very hard time. Here is the condensed version of how to do this. I did not write this in the interview, I tried making this as short as possible afterwards for fun.

```
from itertools import groupby
def groupAnagrams(self, strs):
        sortedAnagrams = sorted(sorted(strs), key=lambda a: sorted(a))
        return [list(v) for k,v in groupby(sortedAnagrams, key=lambda a:
sorted(a))]
```

Impressions: I think I did ok. Talked to him afterwards and it turns out he was a new grad, graduated last year from Cal. It makes sense why he was asking random questions.

Fourth Interview, someone from revenue

Q1: A user complains that the website is slow. What's happening?

A1: Mentioned everything that could go wrong: User's computer is old, user is running other programs making it slow, browser is doing other things at the time, DNS lookups are slow, Javascript is slow, Internet connection is dropping packets, Yelp's CDN is having trouble, or Yelp's servers are overloaded.

Q2: How does TCP rate limiting work?

A2: He saw on my resume I interned at a networking company and took a class on networking so he asked to see if I remember anything. You probably won't get asked this.

Q3: (mock situation). Yelp is storing business IDs as a 20-character case-sensitive alphanumeric string. That means it has the characters [0-9], [a-z], and [A-Z]. Imagine you are given a list of 100 of these IDs all converted to lowercase. Can you match each lowercase-ID back to its proper business ID?

A4: Recursive answer was fine.

Zazzle

1. Implement a HashMap
2. Say I have a multitude of lists, that contain cross streets of bus stops, from start to finish of one bus line. Now, I want to be able to return all the combinations of lists that are reverses of each other. Write a function to do that.
3. Node {
        Node l;
        Node r;
        int v; // value stored in the node
   }

// Example tree; note the binary tree doesn't have to be complete
   8   <-- level 0 (root)
  2 3   <-- level 1
4 5   1 <-- level 2

```
// Examples:
PrintLevel(root, 0); -> 8
PrintLevel(root, 1); -> 2 3
PrintLevel(root, 2); -> 4 5 1

// Complete this function
void PrintLevel(Node root, int level)  {


}
```