CSM 61B Fall 2020

Special Topics

Mentoring 14: November 23, 2020

1 Asymptotics

1.1 Give a tight asymptotic bound for `mystery`.

```
void mystery(int N) {
    for (int i = 1; i <= N * N; i *= 2) {
        for (int j = 0; j < i; j += 1) {
            System.out.println("moo");
        }
    }
}
```

$\Theta(N^2)$

1.2 Give a tight asymptotic runtime bound for `mysterySearch` as a function of N, the size of the array, in the *best case, worst case, and overall*. Assume the array is sorted.

```
public static boolean mysterySearch(int[] a, int value) {
    if (Math.random() < 0.5) {
        return linearSearch(a, value, 0);
    } else {
        return binarySearch(a, value, 0, a.length);
    }
}
```

$\Theta(1)$ in the best case, $\Theta(N)$ in the worst case, and $O(N)$ overall.

2 T,F,G,V,E

2.1 State if the following statements are True or False, and justify. For all graphs, assume that edge weights are positive and distinct, unless otherwise stated.

(a) If a graph G has a unique MST, it must have unique edge weights.

False.
The converse, however, is True: If a graph has unique edge weights, then the graph has a unique MST.

(b) Adding some positive constant k to every edge weight does not change the minimum spanning tree.

True.

(c) Doubling every edge weight does not change the minimum spanning tree.

True.
For the four parts above, we can consider when graph transformations affect the two algorithms:
MST algorithms depend on the relative order of *edge weights*. Hence, adding a constant, or doubling the edge weights does not alter the MST. (More broadly, any monotonically increasing function can be applied, such as squaring the edge weights, assuming they are all positive.)

Shortest path algorithms depend on the relative order of *sums of edge weights*. More specifically, we are concerned about sums of edge weights that represent paths to vertices in the graphs. We can see then that adding a constant k to all edge weights does alter the relative order of these sums. In fact, as k increases, the algorithm becomes more biased towards paths that are shorter in *hop-length*, i.e. number of vertices in the path. One intuitive way to think about this would be to make k a very large number, tending towards infinity. Then all edge weights are approximately the same length, and shortest path algorithms will find the shortest path by hop-length, just like BFS. On the other hand, if we double every edge weight, the relative order of sums does not change. $2w_1+2w_2+2w_3 = 2(w_1+w_2+w_3)$. We see that we can factorize out the multiplier, and the ordering is still dependent on the original sums of edge weights. More broadly, we can consider any positive multiplication of edge weights to not affect shortest path trees.

(d) Let $(S, V-S)$ be a specific cut of the graph. If an edge e is not the lightest edge across this cut, it cannot be a part of any MST.

False. Consider the graph $\{(A, B, 1), (B, C, 2)\}$. Even though edge (B, C) is not the lightest edge across the cut $\{A\}, \{B, C\}$, it is necessarily still a part of all MSTs (since this graph is a tree).

3 Redundant Connections

3.1 For any connected graph with N vertices and N edges, there is at least 1 edge whose removal will keep the graph connected (graph would become a tree).

Given a graph represented with an **edge set**, design an algorithm that would find the K redundant edges of minimum weight. If there are not K redundant edges, return all redundant edges.

Hint: Consider how Kruskal's Algorithm checks for a cycle.

This problem is adapted from Leetcode: https://leetcode.com/problems/redundant-connection/solution/

Intuition

We know from Kruskal's Algorithm that we can find a redundant edge in the graph. If we use that as inspiration, there is a method to find redundant edges. Simply find all the redundant edges and place them into a Priority Queue.

Algorithm

Since the graph is represented as an edge set, simply iterate through the edges and unioning those vertices with a Disjoint Sets object, checking beforehand that they are not connected. If we find that two vertices are connected, the current edge we are looking at is a redundant edge! Add this edge to a PriorityQueue *smallestRedundantEdges* .

Repeat this process but make sure to add the edges in *smallestRedundantEdges* first, to make sure that we are always finding a new redundant edge. If we run through this process without finding a redundant edge, we have found all the redundant edges.

Once we have found all the redundant edges, simply pop from *smallestRedundantEdges* K times if possible.

4 Threads

4.1 For most of the programming assignments in CS 61B, we would write some code in Java, compile it with the `javac` command, then execute it with the `java` command. To truly understand threads, let's dive deeper into the compilation and execution stages.

The Java compiler is invoked by the `javac` command and turns human-written Java code into an *executable* file written in Java bytecode. The executable is the `sample.class` file that results from running `javac sample.java` in the command line.

This executable is a Java program. When we later run this executable (with `java sample.class`), our operating system creates a new process to run the program. By definition, a *program* is the result of compiling code and a *process* is a program currently in execution.

A process that executes a Java program consists of multiple threads. A thread is an independent execution sequence of code. If you have taken CS 61A, one way to think about it is that each thread has its own environment diagram. Another way to think about it is that each thread executes its own chunk of code.

One of the reasons why Java is considered a "high-level" programming language is because each of these threads (within a running Java program) has its own specialized task. For instance, one thread executes the code that we have written (i.e. the `main()` function), another thread frees up unused memory (i.e. garbage collection), another thread may update the display, etc.

Threads within the same process share the same memory. This is extremely useful in the age of parallel computing since it allows us to take advantage of multi-core processors. However, there lies danger in the concurrent access of shared memory (i.e. race conditions). With great power comes great responsibility, and CS 61C and CS 162 will teach you methods to write code that is *thread-safe*.

4.2 Let's explore how we can write multi-threaded Java programs. Due to COVID-19, CSM would like to start a mask donation and distribution program for the needy. Here is the single-threaded version:

```
public class MaskDonationDrive {
    public static int mask_count = 0;
    public static int getCount() { return mask_count; }
    public static void changeCount(int m) { mask_count += m; }

    public static void main(String[] args) {
        while (true) {
            Scanner scanner = new Scanner(System.in);
            System.out.println("How many masks would you like to donate?");

            // Suspends execution until user specifies an integer at the command-line.
            int donated_masks = scanner.nextInt();

            System.out.println("Thank you for donating " + donated_masks + " masks!");
            MaskDonationDrive.changeCount(donated_masks);

            int mask_count = MaskDonationDrive.getCount();
            if (mask_count > 0) {
                System.out.println("Donated " + mask_count + " masks to the needy!");
                MaskDonationDrive.changeCount(-1 * mask_count);
            }
        }
    }
}
```

Fill in the skeleton code below for the multi-threaded version of the program.
One thread will constantly receive donations and update the mask count.
The other thread will distribute donations to the needy every 5 seconds.

```
public class MaskDonationDrive {
    public static int mask_count = 0;

    public static int getCount() { return mask_count; }

    public static void changeCount(int m) { mask_count += m; }

    public static void main(String[] args) {
        Thread donator = new Thread(new Donator());
        Thread distributor = new Thread(new Distributor());

        donator.start();
        distributor.start();
    }
```

```
}

public class Donator extends MaskDonationDrive implements Runnable {
    public void run() {
        Scanner scanner = new Scanner(System.in);
        while (true) {
            System.out.println("How many masks would you like to donate?");

            // Suspends execution until user specifies an integer at the command-line.
            int donated_masks = scanner.nextInt();

            System.out.println("Thank you for donating " + donated_masks + " masks!");
            MaskDonationDrive.changeCount(donated_masks);
        }
    }
}

public class Distributor extends MaskDonationDrive implements Runnable {
    public void run() {
        while (true) {
            Thread.sleep(5000); // Suspends execution of the thread for 5 seconds.
            int mask_count = MaskDonationDrive.getCount();
            if (mask_count > 0) {
                System.out.println("Donated " + mask_count + " masks to the needy!");
                MaskDonationDrive.changeCount(-1 * mask_count);
            }
        }
    }
}
```

Multi-threaded programs can take advantage of processors with multiple cores. In this example, the Donator thread and Distributor thread can run simultaneously on different cores. For heavy computational workloads, parallel computing may potentially save a lot of time. However, access to shared memory must be synchronized correctly to avoid incorrect behavior.