Section 2: Solutions

1. Big-$\mathcal{O}$-No

Put these functions in increasing order. That is, if f comes before g in the list, it must be the case that $f(n)$ is $\mathcal{O}(g(n))$. Additionally, if there are any pairs such that $f(n)$ is $\Theta(g(n))$, mark those pairs.

- $2^{\log(n)}$
- $2^{n \log n}$
- $\log(\log(n))$
- $2^{\sqrt{n}}$
- $3^{\sqrt{n}}$
- $\log(n)$
- $\log(n^2)$
- $\sqrt{n}$
- $(\log(n))^2$

Hint: A useful trick in these problems is to know that since $\log(\cdot)$ is an increasing function, if $f(n)$ is $O(g(n))$, then $\log(f(n))$ is $O(\log(g(n))$. But be careful! Since $\log(\cdot)$ makes functions much smaller it can obscure differences between functions. For example, even though n^3 is less than n^4, $\log(n^3)$ and $\log(n^4)$ are big-Θ of each other.

Solution:

(a) $\log(\log(n))$

(b) $\log(n)$

(c) $\log(n^2)$ This function is $\Theta(\log(n))$. They differ only by the constant factor 2.

(d) $(\log(n))^2$

(e) $\sqrt{n}$

(f) $2^{\log(n)}$ Note that this is just n.

(g) $2^{\sqrt{n}}$

(h) $3^{\sqrt{n}}$

(i) $2^{n \log n}$

2. Write it Slicker: A Proof by Contradiction

Claim: For every directed graph G, if every node of G has out-degree at least 1, then G has a directed cycle.

(a) Prove the claim using proof by contradiction. **Solution:**

An unclear proof. Suppose, for the sake of contradiction, there is a directed graph G such that every node of G has out-degree at least 1, but G has no directed cycle.

We construct long and longer simple paths in G: Start at some node v_0 of G. Its out-degree is at least 1, so we can move from v_0 along an out-edge. If this leads to v_0 that would be a directed cycle of length 1, a contradiction! Therefore we have a new node v_1, which (like every other node) has out-degree at least 1, so there is an out-edge pointing to v_2. If v_2 is a vertex we have already visited (i.e., v_0 or v_1), then we

have found a directed cycle, a contradiction! Otherwise, from v_2, we may repeat the same argument and continue finding v_3, v_4, ... Since the graph is finite, so we cannot continue this process forever. Eventually we find a repeated vertex, which means we have a directed cycle, a contradiction! □

(b) Rewrite the proof, using the proof by contradiction with extremality technique.

Solution:

Proof. Suppose, for the sake of contradiction, that there is a directed graph G such that every node of G has out-degree at least 1, but G has no directed cycle. Let $P = v_0, v_1, \ldots, v_k$ be a longest simple directed path in G.

Since v_k has out-degree at-least 1, it must have an out-edge to a vertex W.

Since P is a longest simple directed path, w must be a repeat of a vertex among $v_0, v_1, \ldots, v_k$ (otherwise $v_0, v_1, \ldots, v_k, w$ would be a longer simple directed path). Let v_j be the vertex $= w$..

But then adding the edge (v_k, v_j) to the part of P from v_j to v_k gives a directed cycle in G:

$$v_j, \ldots, v_k, v_j.$$

But that contradicts the assumption that G has no directed cycle! □

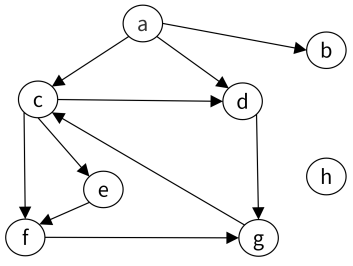
(c) There's another style yellow-flag in the version of this proof from part (a). We're proving an implication and our contradiction was the negation of one of the two things we supposed at the start. That usually means that proof by contrapositive would be clearer. Try writing this proof by contrapositive. **Solution:**

Proof. We argue by contrapositive. That is, we will show for all graphs, G, that if G does not have a directed cycle, then it has a node of out-degree 0.

Let G be an arbitrary directed acyclic graph, and consider a maximal path $v_0, v_1, ..., v_k$. Since the graph is acyclic, there cannot be an edge from v_k to any vertex in $v_0, v_1, \ldots v_k$. Since the path is maximal, v_k also can't point to anything else so it must have out-degree 0. □

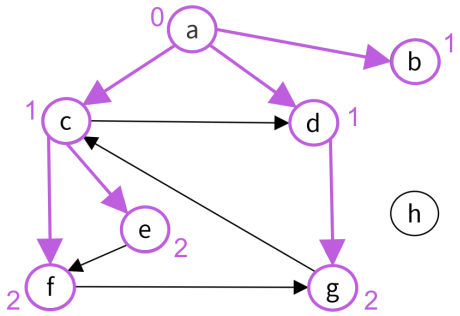
3. Mechanical: BFS and DFS

Consider the graph below.



(a) Run Breadth-First-Search on the graph below and number the layer for each node. Start with layer 0.

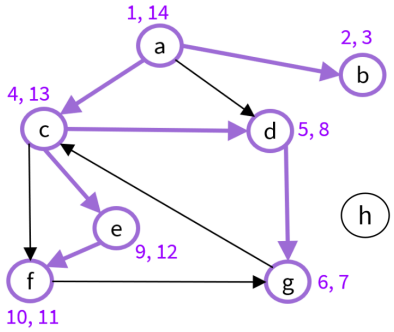
Solution:



Layer 0: a
Layer 1: b, c, d
Layer 2: e, f, g

(b) Run Depth-First-Search on the graph below to classify the edges. Mark the start and end times for each vertex.

Solution:



Tree Edges: (a,b), (a,c), (c,d), (d,g), (c,e), (f,g)
Back Edge: (g,c)
Forward Edges: (a,d), (c,f)
Cross Edges: (f,g)

4. Graph Modeling

In this problem we're going to solve a classic riddle.

(a) First, you should solve the classic riddle yourself to get a feel for the problem.

You are on the beach with a jug that holds exactly 5 gallons, a jug that holds exactly 3 gallons, and a large bucket. Your goal is to put **exactly** 4 gallons of water into the bucket. Unfortunately, the jugs don't have markings to tell how full they are (e.g., you can't just fill the larger jug $4/5$ full). What you can do are the following operations.

- Completely fill any of your jugs.
- Pour from one of your containers into another until the first container is empty or the second is full.
- Pour out all the remaining water in a container.

How do you get 4 gallons of water into the bucket?

Solution:

Fill the 5 gallon jug, pour into the 3 gallon jug until it is full (the jugs now contain 2 and 3 gallons respectively). Pour from the larger jug into the large bucket (it now has 2 gallons). Empty the 3 gallon jug, and repeat all these steps to get the desired 4 gallons.

Alternatively, Fill the 3 gallon jug, pour into the 5 gallon jug. Fill the 3 gallon jug again, pour until the 5 gallon jug is full (they now contain 5 gallons and 1 gallon respectively) . Pour the 1 gallon from the 3 gallon jug into the bucket. Refill the 3 gallon jug and pour into the bucket to bring the total to 4 gallons.

There may be other solutions.

(b) Now, let's write an algorithm to solve any instance of this puzzle. You are given a list of 10 jugs with (positive integer) capacities $c_1, ..., c_{10}$, ranging from 1 to C. Your goal is to determine whether it is possible to get exactly t gallons into a bucket with capacity that is at least t and at most B.

Hint: Think about how you can relate the possible "states" of the puzzle to the nodes of some graph. What would be a good way to define edges for this graph? How could you think of solutions to the puzzle in terms of the graph?

Solution:

Intuition
The "state" of the puzzle can be represented as the number of gallons in each of the jugs and the bucket. We encode the rules of the puzzle such that each possible step is an edge.

Algorithm
Let S be the set of all 11-tuples, where for the first 10 entries, the entry is an integer between 0 and c_i, and the final entry is an integer between 0 and $10C$ (upper bound on the total capacity of all jugs). There are $(C+1)^1 0 \cdot (B+1)$ such states.

We make a graph with a vertex for every element of S. And add an edge from u to v if and only if the states meet one of these conditions.

From a given state $(j_1, j_2, ..., j_{11})$, you can move to another state $(j'_1, j'_2, .., j'_{10}, j'_{11})$ if and only if one of the following hold:

- There is only one index, k, where the tuples differ, and $j'_k = 0$. (we emptied a jug or the bucket)
- There is only one index, k, ($k \leq 10$) where the tuples differ, and $j_k = c_k$. (we filled a jug)
- There are two indices, k, ℓ where the tuples differ:
  - $j'_k = 0$ or $j'_\ell = c_\ell$
  - $j_k + j_{ell} = j'_k + j'_\ell$.

Finally, we add a target vertex z, to the graph. Add an edge from every tuple where $j_{11} = t$ to z.

We then use [B/D]FS, starting from the all 0's tuple, and searching to see if z is reachable. If it is, we can return true (and predecessor edges will show the steps to take). If z is not reachable, then return false.

Correctness
Suppose our algorithm returns true. Then [B/D]FS found a walk from all 0's to z. By construction of the graph, each edge corresponds to a valid rule we can apply in the original puzzle. Since the only edges going into t are from states where $j_{11} = t$, the walk must reach such a state, thus the puzzle can be solved by doing the steps on the edges of that walk.

Conversely, suppose the puzzle is solvable. Then there is a series of steps that can be taken to put t gallons into the bucket. Each legal step has a corresponding edge in the graph to the next state by construction, so there is a path to a valid stopping state (i.e., a tuple where $r = t$), we added an edge from all such vertices to z, so there is a path from all 0's to z. [B/D]FS will discover this path, so we will return true.

Running Time
Let $n = (C+1)^{10}(B+1)+1$. Note that this is the number of vertices in our graph. Each vertex has a constant number of edges leaving it (as you are performing one of three operations (dump, pour, fill) among a constant number of jugs. So the graph can be has $\mathcal{O}(n)$ edges and can be constructed in $O(n)$ time. Running [B/D]FS in a graph with $\mathcal{O}(n)$ vertices and edges takes $\mathcal{O}(n)$ time, so the overall running time is $\mathcal{O}(n)$.

5. Judging Books by Their Covers

You have a large collection of books, and just got a new bookshelf. For aesthetic reasons, you're going to arrange your books by the color of their covers (not by author or subject). You wish to put only books of a single color on any given shelf. You have a list of pairs of books which you know to be the same color. This list might be only partial (it's possible that $u, v,$ and w are all the same color, but your list might only have "u and v are the same color. w and v are the same color.", for example). You should assume that the "same color" relation is transitive.

Given your list, your job is to give an upper-bound on the number of shelves you need so that no shelf has more than one color of book. Describe an algorithm to give the best bound you can on the number of shelves needed.

You do not need a full proof of correctness, but you should describe the running time in terms of (whichever subset is appropriate): b, the number of books; p the number of pairs listed; s the number of shelves required (i.e., your final answer). **Solution:**

We'll make a graph as follows: have a vertex for every book and an edge between u and v if they are listed as a "same color" pair. Since "same color" is transitive, if there is a walk from u to v then u and v are the same color. Thus any subset of a connected component can be on its own shelf (and we can't put books from separate connected components on the same shelf).

Our algorithm can run [B/D]FS to find connected components, and then iterate through the vertices to count the number. The graph will have b vertices and p edges, so the running time is $\mathcal{O}(b+p)$.