CS3000: Algorithms & Data — Summer I '21 — Drew van der Poel

Homework 5
Due Friday, June 18 at 11:59pm via Gradescope

Name: Gabriel Peter
Collaborators:

- Make sure to put your name on the first page. If you are using the LaTeX template we provided, then you can make sure it appears by filling in the `yourname` command.

- This assignment is due Friday, June 18 at 11:59pm via Gradescope. No late assignments will be accepted. Make sure to submit something before the deadline.

- Solutions must be typeset in LaTeX. If you need to draw any diagrams, you may draw them by hand as long as they are embedded in the PDF. I recommend using the source file for this assignment to get started.

- I encourage you to work with your classmates on the homework problems. *If you do collaborate, you must write all solutions by yourself, in your own words.* Do not submit anything you cannot explain. Please list all your collaborators in your solution for each problem by filling in the `yourcollaborators` command.

- Finding solutions to homework problems on the web, or by asking students not enrolled in the class is strictly forbidden.

Problem 1. *No one man should have all that power* (26 points)

You work for a public relations firm and have acquired complete information on the post-election Washington DC lobbying network. In particular, you have a list V of n persons (congressmen, bureaucrats, lobbyists, businessmen, etc.) and for each person i in V, a list of persons in V that i can *influence*. (Note that it is possible that i can influence j but j cannot influence i.)

Having just studied graph algorithms, you immediately capture the above information by a directed graph G with V as the set of vertices and the set E of edges defined as follows.

$$E = \{(i,j) : i \text{ can influence } j\}.$$

Call a person i a *powerbroker* if for every other person $j \in V$, there is a path from i to j in G.

(a) **(4 points)** Give an example lobbying network with five persons in which there are exactly three powerbrokers.

   Solution:
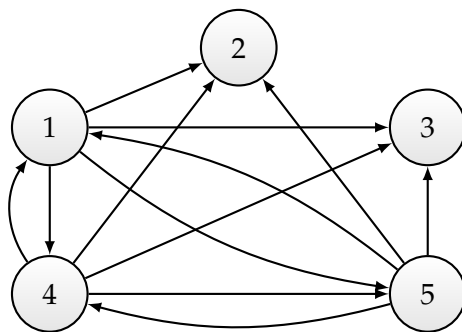
(b) Design and analyze a polynomial-time algorithm to determine *all* powerbrokers in the given lobbying network. If there are no powerbrokers, then your algorithm must indicate so.

   (i) **(2 points)** Describe precisely what your algorithm is given as input and what it needs to output.[1]

   Solution:

   The algorithm simply needs a provided graph in the form of adjacency lists. It will then output a list of the vertices that compose the list of powerbrokers (aka the source SCC, which will be explained later.)

   (ii) **(8 points)** Describe your algorithm. You may invoke or modify any of the graph traversal algorithms studied in class.

   Solution:

---

[1]**Check:** Make sure you have this right, before you move on to designing the algorithm.

```
def findBrokers(graph):

        # Simplified DFS that only
        # counts nodes visited: O(n+m)
        def DFS(node, graph):
                global count
                count = 0
                explored = {}

                def DFShelper(node, graph):
                        global count
                        explored[node] = 1
                        count += 1
                        for v in graph[node]:
                                if explored[v] == 0:
                                        DFShelper(v, graph)


                DFShelper(node, graph)
                return count

        # Runs for each run node: O(n)
        for node in graph:
                influence_count = DFS(node, graph)
                if influence_count == len(graph):
                        print(f 'Node {node} is a powerbroker ')
```

This algorithm simply iterates a DFS at each node.

(iii) **(8 points)** Justify the correctness of your algorithm. Your justification does not need to be long or formal, just convincing.

Solution:

As mentioned before, we can be certain that all powerbrokers are in the same SCC because they must be connected to each in order to satisfy the requirement that:

"Call a person i a *powerbroker* if for every other person $j \in V$, there is a path from i to j in G."

Additionally, the SCC in question must be the source SCC because a source component, has a path to all other nodes which is necesarry for a power broker.

Therfore, powerbrokers $\leftrightarrow SCC_{source}$
This is ensured by the algorithm by printing the

(iv) **(4 points)** Analyze the running time of your algorithm in terms of the number of vertices n and number of edges m of G.

3

(Any correct polynomial-time algorithm will suffice to get full credit. See if you can design an algorithm with $O(n+m)$ running time using strongly connected components.)

Solution:

The runtime of this algorithm is identical to the SCC mentioned in class, but simply more things are cached for later reference:

   i. tranposing the graph: $O(m)$
  ii. DFS on the tranposed: $O(n+m)$
 iii. comp init: $O(n)$
 iv. Nested for-loops assigning SCC's to vertices: $O(n+m)$
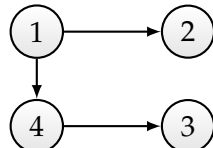
TOTAL: $O(m) + O(n+m) + O(n) + O(n+m) = O(n+m)$

Problem 2. *Stagecraft* (24+5 points)

You're in charge of assembling the stage for this year's *AlgoRhythms Music Festival (starring Master P).* Assembling the stage in time will require careful planning. You are given:

- A set V of n small tasks that are required to complete the stage.

- A set E of pairs of tasks in V. A pair (u, v) is in E if task u must be completed before task v is started. There are no cycles in E.

- For each task $u \in V$, the amount of time $t(u)$ required for the task.

You have a very large team, so you can work on any number of tasks in parallel, but you cannot start a task u until all of the prerequisite tasks v have been completed.

Design an algorithm that takes as input the graph $G = (V, E)$ (represented as an adjacency list) and the time for each task, and outputs a list consisting of the earliest possible time that each task can be completed. An example of a correct input-output is:

Input:

$t(1) = 10$
$t(2) = 11$
$t(3) = 8$
$t(4) = 5$

Output:

$c(1) = 10$
$c(2) = 21$
$c(3) = 23$
$c(4) = 15$

Your algorithm should run in $O(n + m)$ time.

(a) **(6 points)** Describe in 1-3 English sentences how you will determine the earliest possible completion times for the tasks.

Solution:
I will determine the the earliest possible completion of each node by running a modified DFS algorithm for topological sorting, which has an additional modification that labels each node's topology as their cumulative time weights such that we keep track of the current max accumulation of times for a DFS.

(b) **(6 points)** Describe your algorithm in pseudocode. You may make use of any algorithm we've seen in class without describing how it works, but you should clearly state what you are assuming about the algorithm.

Solution:

```
# Main algo method which returns mappings of
# min times needed to complete a node
def FastTop(graph, weights):
    times = {k:0 for k in graph.keys()}
    queue = []        #Initialize a queue
    in_deg = {k:0 for k in graph.keys()}
```

```
# Removes a node and adjusts values that are
# affected from its deletion (times, in_deg)
def deleteNode(node, time_acc):
    time_acc = max(time_acc, times[node]
                   + weights[int(node)-1])
    times[node] = time_acc
    for w in graph[node]:
        times[w] = max(times[node], times[w])
        in_deg[w] -= 1
    for w in graph[node]:
        if in_deg[w] == 0:
            deleteNode(w, time_acc)

# Marks all nodes
for node in graph.keys():
    neighbors = graph[node]
    for out in neighbors:
        in_deg[out] += 1

# Finds 0 in_degree nodes
for node, deg in in_deg.items():
    if deg == 0:
        queue.append(node)

#Queue starts.
while queue:
    u = queue.pop(0)
    deleteNode(u, 0)

return times
```

The algorithm works very similar to the FastTop() mentioned in lecture 16:

Essentially, we remove all nodes and then add their accumulated time to complete starting at the nodes with 0 in-degree edges. By deleting nodes, we cause others to now have 0 in-degrees and are then deleted as well. The topological times are a max of the current label or the new accumulation, which denotes the soonest this node (task) can be completed after all prerequisites are finished. The labelings are stored in a map: 'times' and from there can be easily acessed such that: $c(i) = times[i]$

(c) **(8 points)** Justify that your algorithm outputs the earliest possible completion time for each project. Your justification can take any form you like as long as the argument is clear and logical.

Solution:

Theorem: G has a topological ordering ↔ G is a DAG

As that is a provided contraint of the graph input, that it is a acyclic DAG. We know that we can get a topological ordering. In a normal topological sort, know that we can get a list which is a guaranteed precedence ordering of the nodes provided. I alter that algorithm ever-so-slighty to maintain its correctness of topological order, but labeling them through an accumulation of their $t(i)$ value. We know based on the alogrithm that the accumulation will only be passed to nodes that are in the precendence of the currently deleting node, as the next-to-be deleted are out_degrees of it. However, based on the configuration of the graph, we may have multiple precedences for a single node, and this causes a value conflict in out labeling. To solve this, we take the max(old accumulation, new accumulation) as the problem constraints mentions that all preceding nodes but be completed, even in parallel. Thus, the soonest time the node have be completed is after the longest prerequisite is.

(d) **(4 points)** Analyze the running time of your algorithm. If your algorithm uses some algorithm from class as a subroutine, don't forget to include this running time in your analysis.
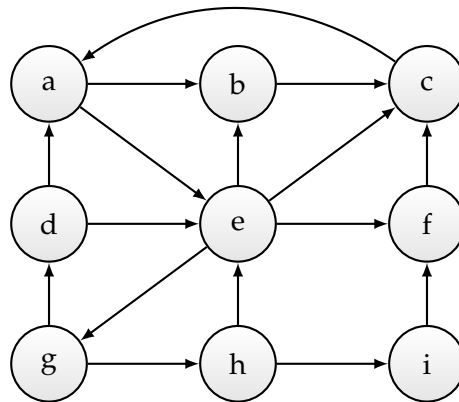
Solution:
The runtime of this algorithm is idential to the pseudocode's runtime of FastTop in lecture 16 as the modifications do not contribute to the asymptotic performance. Thus, $O(n + 3m) = O(n + m)$

(e) **(+5 points)** - In order to better your familiarity with graph algorithms, we have created a hackerrank challenge (https://www.hackerrank.com/CS3000-summer1-2021-programming-assignment-5). Please implement your algorithm and submit it to the challenge. Please list your username below. Because this problem is bonus, the course staff will generally refrain from providing assistance here.

Solution:
peter_g

Problem 3. *DFS and Topological Ordering* (15 points)

Consider running depth-first search on this graph starting from node *a*. When there are multiple choices for the next node to visit, go in alphabetical order.

(a) **(5 points)** Label every edge as either tree, forward, backward, or cross.

Solution:

Lecture 15...

(a,b): forward
(a,e): tree

(b,c): cross

(c,a): cross

(d,a):
(d,e):

(e,b): tree
(e,c):
(e,f):
(e,g): tree

(f,c): tree

(g,h): tree
(g,d): tree

(h,e):
(h,i): tree

(i,f): tree

(b) **(5 points)** Give the discovery and finish times of all vertices

Solution:

Vertex	Discovery	Finish
a	1	18
b	15	16
c	9	10
d	4	5
e	2	17
f	8	11
g	3	14
h	6	13
i	7	12

(c) **(5 points)** Is this graph a DAG? If so, give a topological ordering.

Solution:
No, the exist no node that has $in\_degree(v) = 0$. Therefore, a cycle is in this graph, which a DAG cannot have being acyclic.

Problem 4. *Shortest Paths Practice* (10 points)

Use Dijkstra's algorithm the single-source shortest path problem on the following graph, using node 1 as the source node s. Write the distance from s to each node and the parent of each node in the shortest path tree in the skeleton table provided. Show your work in order to receive partial credit.
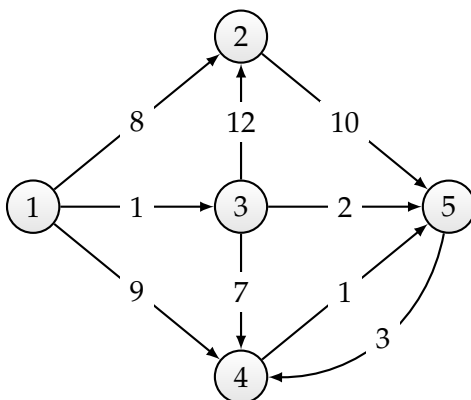
Solution:

Node	1	2	3	4	5
Distance	0	8	1	6	3
Parent	$\emptyset$	1	1	5	3

ATTACHED SCRATCH WORK

Node	1	2	3	4	5
d_0	0	inf	inf	inf	inf
d_1	0	8^1	1^1	9^1	inf
d_2	0	8	1	8^3	3^3
d_3	0	8	1	6^5	3
d_4	0	8	1	6	3

$U = \{1, 3, 5, 4, 2\}$