Example sheet 5

Graph traversal, path finding

Algorithms—DJW[*]—2019/2020

Questions labelled $\circ$ are warmup questions. Questions labelled $*$ involve more thinking and you are not expected to tackle them all.

Question 1°. Read the definitions of *directed acyclic graph* and *tree* from lecture notes. Draw an example of each of the following, or explain why no example exists:
(i) A directed acyclic graph with 8 vertices and 10 edges
(ii) An undirected tree with 8 vertices and 10 edges
(iii) A graph without cycles that is *not* a tree

Question 2°. Show that `dfs_recurse` has running time $O(V + E)$.

Question 3. The algorithms `dfs` and `dfs_recurse` (as given in lecture notes) do not always visit vertices in the same order. Modify `dfs` so that they do. Give pseudocode. *[Assume that there is an ordering on vertices, and that v*`.neighbours` *returns a sorted list of v's neighbouring vertices.]*

Question 4. Give pseudocode for a function `dfs_recurse_path(`g, s, t`)` based on `dfs_recurse`, that returns a path from s to t. Modify your function so that it does not visit any further vertices once it has reached the destination vertex t.

Question 5. Consider a directed graph in which every vertex v is labelled with a real number x_v. For each vertex v, define m_v to be the minimum value of x_u among all vertices u such that either $u = v$ or u is reachable via some path from v. Give an $O(E + V \log V)$-time algorithm that computes m_v for all vertices v.

Question 6. Modify `bfs_path(`g, s, t`)` to find *all* shortest paths from s to t.

Question 7. The breadth-first search algorithm from lecture notes uses $O(1)$ storage within each vertex object (to store the `seen` flag), plus extra memory for `toexplore`. What is the worst case memory requirement of `toexplore`? Give your answer using Ω notation, in terms of V and E. Modify the algorithm to use $O(1)$ storage within each vertex object, plus $O(1)$ extra memory.

Question 8 (FS)°. In a directed graph with edge weights, give a formal proof of the triangle inequality

$$d(u,v) \le d(u,w) + c(w \to v) \quad \text{for all vertices } u, v, w \text{ with } w \to v$$

where $d(u,v)$ is the minimum weight of all paths from u to v (or ∞ if there are no such paths) and $c(w \to v)$ is the weight of edge $w \to v$. Make sure your proof covers the cases where no path exists.

Question 9. There is a missing step in the proof of Dijkstra's algorithm, as given in lecture notes. The proof looks at the state of program execution at the point in time at which some vertex v fails the assertion on line 9. Call this time t. It uses the equation

$$u_{i-1}\texttt{.distance} = \text{distance}(s \text{ to } u_{i-1}),$$

and justifies it by saying "the assertion on line 9 didn't fail when u_{i-1} was popped." Let t' be the time when u_{i-1} was popped, and explain carefully why u_{i-1}`.distance` cannot change between t' and t.

[*]Questions labelled FS are from Prof Stajano. Questions labelled CLRS are from *Introduction to Algorithms, 3rd ed.* by Cormen, Leiserson, Rivest and Stein.

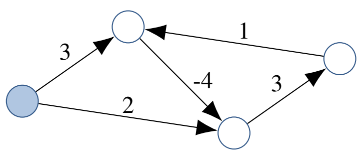
Question 10 (CLRS-24.3-4). A contractor has written a program that she claims solves the shortest path problem, on directed graphs with edge weights ≥ 0. The program produces v`.distance` and v`.come_from` for every vertex v in a graph. Give an $O(V+E)$-time algorithm to check the output of the contractor's program. *[Hint. Pay attention to the case where some edge weights are zero.]*

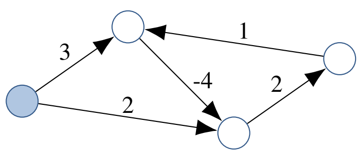
Question 11. Suppose we have implemented Dijkstra's algorithm using a priority queue for which `push` and `decreaskey` have running time $\Theta(1)$, and `popmin` has running time $\Theta(\log n)$ where n is the number of items in the queue. Construct a sequence of graphs indexed by k, where the kth graph has $|V| = k$ and $|E| = \Theta(k^\alpha)$, such that Dijkstra's algorithm has running time $\Omega(k^\alpha + k\log k)$. Here α is a constant, $1 \leq \alpha \leq 2$.

Question 12. We are given a directed graph where each edge is labelled with a weight, and where the vertices are numbered $1, \dots, n$. Assume it contains no negative weight cycles. Define $F_{ij}(k)$ to be the minimum weight path from i to j, such that every intermediate vertex is in the set $\{1, \dots, k\}$. Give a dynamic programming equation for $F_{ij}(k)$, and a suitable definition for $F_{ij}(0)$.

Question 13°. By hand, run both Dijkstra's algorithm and the Bellman-Ford algorithm on each of the graphs below, starting from the shaded vertex. The labels indicate edge costs, and one is negative. Does Dijkstra's algorithm correctly compute minimum weights?

[Run `dijkstra` *as described in lectures, which loops until* `toexplore` *is empty. Some textbooks use different versions of Dijkstra's algorithm, that terminate once a destination vertex has been popped, or that don't allow popped vertices to re-enter* `toexplore`*.]*





Question 14. In the course of running the Bellman-Ford algorithm, is the following assertion true? "Pick some vertex v, and consider the first time at which the algorithm reaches line 7 with v`.minweight` correct i.e. equal to the true minimum weight from the start vertex to v. After one pass of relaxing all the edges, u`.minweight` is correct for all $u \in \text{neighbours}(v)$."

If it is true, prove it. If not, provide a counterexample.

Question 15 (CLRS-25.3-4). An engineer friend tells you there is a simpler way to reweight edges than the method used in Johnson's algorithm. Let w^* be the minimum weight of all edges in the graph, and just define $w'(u \to v) = w(u \to v) - w^*$ for all edges $u \to v$. What is wrong with your friend's idea?

Question 16*. You've learned what $O(n)$ means when there's a single variable n that increases. How would you define $O(E + V\log V)$?

Question 17*. Give a $O(V+E)$ algorithm that solves Question 5. *[This goes beyond what is taught in lectures. Credit to Georgiev (matriculated 2016) for spotting this.]*

Question 18*. Consider a graph without edge weights, and write $d(u,v)$ for the length of the shortest path from u to v. The *diameter* of the graph is defined to be $\max_{u,v\in V} d(u,v)$. Give an efficient algorithm to compute the diameter of an undirected tree, and analyse its running time. *[Hint. One way to solve this is by using breadth-first search, twice.]*

Question 19*. The Bellman-Ford code given in lecture notes will report "Negative cycle detected" if there is a negative-weight cycle reachable from the start vertex. Modify the code so that, in such cases, it returns a negative-weight cycle, rather than just reporting that one exists.