CS 161 Fall 2017: Section 5 Solutions

Farmville

Suppose there is a set of farms that have entered into a cooperative agreement. They built a series of irrigation pipes connecting the farms, with each pipe having a direction of flow (i.e. water can only travel in one, fixed, direction in each pipe). Each farm uses some of the water that passes through the pipes that go through the farm, and each farm also has a well that can contribute water to any irrigation pipes leaving the farm. [Think of this as a way of distributing the risk that certain wells might go dry one year, and others go dry a different year...some years Farmer Joe takes more water from the network than he contributes, some years he supplies more water from his well than he takes.]

(a) After designing the network of (directed) pipes, the farmers want to ensure that water from every farm can get to every other farm. (For example, if there is only one farm whose well is still working, every farmer should still be able to get some of that water via the irrigation pipe network.) The only algorithm that the farmers know is BFS, which runs in time $O(|V| + |E|)$ and takes as input a directed graph $G = (V, E)$ and a source node $s \in V$, and outputs the set of nodes that can be reached from node s. Design a way for the farmers to check whether the desired condition holds for a given proposed pipe network, that uses their BFS algorithm only a *constant* number of times.

- Pick an arbitrary node $v \in V$.
- Run BFS in G from v and check that all of V is reachable from v; if not, return "no".
- Construct the reverse graph $G' = (V, E')$, where $E' = \{(v, u) : (u, v) \in E\}$.
- Run BFS in G' from v and check that all of V is reachable from v; if not, return "no". Otherwise return "yes".

Running time: It takes $O(1)$ time to pick some node v, $O(m + n)$ time to run BFS on G, $O(m + n)$ to build G' from G (or just reverse the direction of existing edges in G), and $O(m+n)$ time to run BFS again on G'. This is $O(m + n)$ in total.

Correctness: Running BFS from v in G tells us whether v can reach all other nodes; running BFS from v in the reverse graph G' tells us whether all other nodes can reach v in G. This is clearly a necessary condition if we want all nodes to be able to reach one another; is it also sufficient? It is, since to get from any node u to any other node w, we can take the path $u \to \cdots \to v \to \cdots \to w$. Thus our algorithm outputs "yes" if and only if every node can reach every node, which corresponds to every farm being able to send water to every other farm.

(b) Thanks to your solution, the farmers built a successful network, and enjoy many years of steady crops. One year, however, a turnip gets stuck in an irrigation pipe, clogging that pipe; in the week it takes to fix that pipe, Farmer Joe's potatoes all perish from lack of water. The farmer coop forms a new emergency planning committee to design a new, improved irrigation plan. This time, they want a network such that the guarantees of the above part (namely that water from any farm's well can flow to any other farm via the irrigation pipes) will continue to hold even if any single pipe becomes clogged. Describe

an algorithm for this part that uses your algorithm from the previous part. Justify the correctness and runtime of the algorithm.

For each edge $e \in E$, run the algorithm of part (a) on $G \setminus \{e\}$, the graph obtained by removing e. Return "no" if any of these returns "no"; otherwise return "yes".
Running time: We run the algorithm of part (a) m times, giving a running time of $O(m^2 + mn)$

(c) What is the minimum number of irrigation pipes, as a function of the number of farms, n, that are necessary in any network that satisfies the desired property described in part (b)? Prove your answer– both that the claimed number can be achieved, and that no smaller number is possible.

The minimum number of pipes/edges is $2n$. To see this, note that:

- It is necessary that for every node $v \in V$, indegree$(v) \geq 2$; that is, at least 2 edges pointing towards it. If not, then if we removed the sole edge (u, v) from G, then v cannot reach any other nodes. The same argument shows that we require outdegree$(v) \geq 2$. If this is the case, then the number of edges is
$$m = \frac{1}{2} \sum_{v \in V} \deg(v) \geq \frac{1}{2} \sum_{v \in V} 4 \geq 2n$$
where the $1/2$ comes from the fact that each edge contributes to the degrees of two nodes.
- The condition above is also sufficient. To see this, number the vertices $v_1, \cdots, v_n$ and consider the "double-cycle" graph containing all edges (v_i, v_{i+1}) and (v_{i+1}, v_i). This graph remains strongly connected if any single edge is removed, and it has exactly $2n$ edges.

2SAT-SCC

In 2SAT, you have a set of boolean clauses, each containing two variables. Your goal is to set all of these clauses to true, or report that there is no way to accomplish this goal. For example,

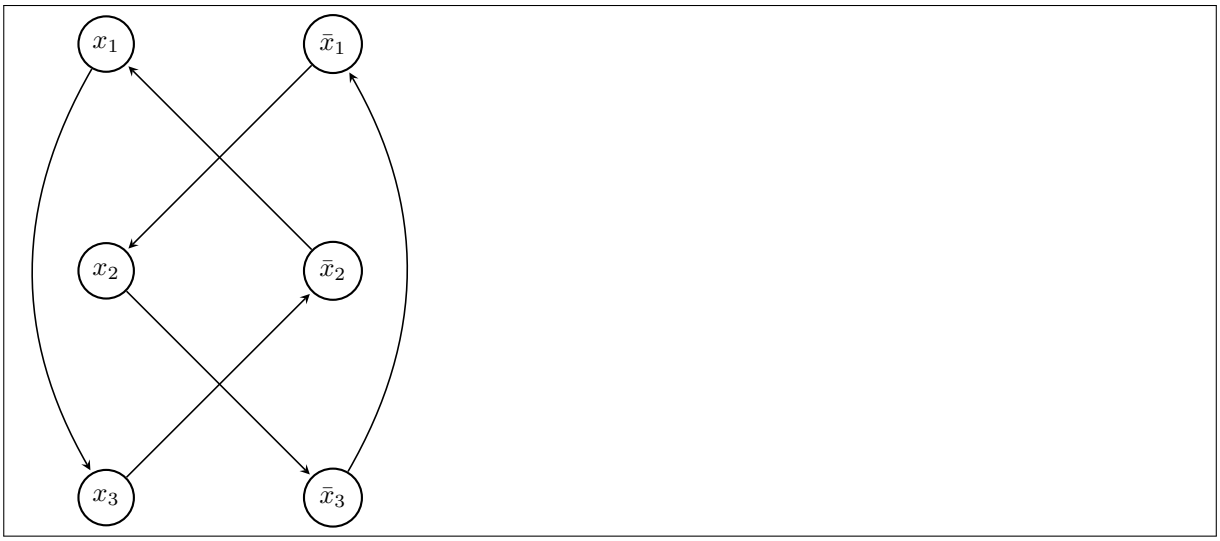
$$(x_1 \vee x_2) \wedge (\bar{x}_1 \vee x_3) \wedge (\bar{x}_2 \vee \bar{x}_3)$$

can be satisfied by setting x_1 and x_3 to true and x_2 to false.

(a) We first decompose each clause into its implications: this means that a clause represented by $a \vee b$ would be represented by $\bar{a} \implies b$ and $\bar{b} \implies a$. What are the implications in the clauses above?

$\bar{x}_1 \implies x_2$, $\bar{x}_2 \implies x_1$
$x_1 \implies x_3$, $\bar{x}_3 \implies \bar{x}_1$
$x_2 \implies \bar{x}_3$, $x_3 \implies \bar{x}_2$

(b) We build a graph by creating two nodes for each variable x: x and $\bar{x}$. We then add directed edges based in the implications from each clause. What is the graph representation of the above boolean clauses?



(c) What do SCCs in these 2SAT graphs represent? What if x and $\bar{x}$ were inside of the same SCC?

SCCs represent nodes that share the same state. If one node in an SCC is set to true, the implication graph forces all variables in that SCC to be set to true.
If x and $\bar{x}$ were inside of the same SCC, this means that the original 2SAT problem has no solution because it is logically impossible for x to be both true and false, but the implication graph would require that.

(d) Assume we have a sink strongly-connected component C from the graph, and that no variable and its negation both appear inside of C. If C has nodes of the form $a, b \dots z$, what corresponding component do we know must exist in the graph? Is there an edge between these two components?

First, to show that there is a corresponding component $\bar{C}$, we observe that if there is a cycle $a \implies b, b \implies c \dots y \implies z$ in C, we must have a corresponding cycle $\bar{z} \implies \bar{y} \dots \bar{b} \implies \bar{a}$ in $\bar{C}$. Therefore, any cycle our component has must also exist somewhere else in the graph, which defines a corresponding component.

To show that there are no edges between the component and its corresponding component, we use a proof by contradiction. If we assume that there is an edge between C and $\bar{C}$, C must contain some node m that has an outward edge to $\bar{C}$, $m \to n$, where n is within $\bar{C}$. However, this also means that there is an edge $\bar{n} \to \bar{m}$, where $\bar{n} \in C$ and $\bar{m} \in \bar{C}$. Therefore, there are edges to go from $C \to \bar{C}$, and from $\bar{C} \to C$, which means that they are the same connected components, meaning that variables and their negations are within the same SCC, breaking our original assumption.

(e) Given this graph structure and the properties above, develop an algorithm to solve 2SAT.

We go through the components in reverse-topological order. We set the nodes in the component to true, and also set the nodes in the corresponding component to false.