Constrained Reachability is NP-complete

Miki Hermann

March 27, 1998

Given an oriented graph $G = (V, A)$ with two specified nodes $s, t \in V$, the REACHABILITY problem is to determine if there exists a path from s to t. This problem is known to be **NL**-complete, i.e., complete for the nondeterministic logarithmic space [Pap94]. In constraint solving we encounter a similar problem for an oriented multigraph $G = (V, A)$ with two specified nodes $s, t \in V$, where each arc $a \in A$ is labelled by a constraint $c(a)$. The problem is to find a path p from s to t, such that the conjunction of the constraints $c(a_i)$ labelling the arcs a_i in the path p is satisfiable. The constraints $c(a)$ can be of different nature. We consider in this paper the simplest case where the constraints are represented by clauses in propositional logic. Recall that a clause is a disjunction of literals, and a literal is a Boolean variable or its negation.

Hence, we consider the complexity of the following problem of finding a constrained path in an oriented multigraph.

CONSTRAINED REACHABILITY
Instance: Directed multigraph $G = (V, A)$ with two specified nodes $s, t \in V$, and clauses $c(a_i)$ labelling the arcs $a_i \in A$, respectively for each $i = 1, \ldots, |A|$.
Question: Is there a path $p = a_{i1} \cdots a_{ik}$ from s to t, such that the conjunction $c(a_{i1}) \wedge \cdots \wedge c(a_{ik})$ of the clauses labelling the arcs of the path p is satisfiable?

Theorem 1 *The problem* CONSTRAINED REACHABILITY *is* **NP**-*complete.*

Proof: The lower bound is easily shown by reduction from SAT, the satisfiability problem for Boolean CNF formulas. Let $\phi = c_1 \wedge \cdots \wedge c_k$ be a Boolean CNF formula. We construct the oriented graph $G = (V, A)$: $s = v_0 \xrightarrow{c_1} v_1 \xrightarrow{c_2} v_2 \cdots v_{k-1} \xrightarrow{c_k} v_k = t$, with the nodes $V = \{v_0, v_1, \ldots, v_k\}$ and the arcs $A = \{v_{i-1} \longrightarrow v_i \mid i = 1, \ldots, k\}$. We label each arc $v_{i-1} \longrightarrow v_i$ by the clause c_i, and set $s = v_0$ and $t = v_k$. It is straightforward that there exists a constrained path from s to t in G if and only if the formula ϕ is satisfiable.

For the upper bound, we show that the problem of finding a constrained path can be solved by reduction in polynomial time to the satisfiability problem of a Boolean formula. Given an oriented multigraph $G = (V, A)$, let $in(v)$ be the set of incoming arcs pointing to the node v. For each node $v \in V$, write the equation $e(v)$: $v = (c(a_1) \wedge v_1) \vee \cdots \vee (c(a_n) \wedge v_n)$ if $in(v) = \{a_1, \ldots, a_n\}$ is the set of incoming arcs pointing to the node v and $a_i = v_i \longrightarrow v$ for all $i = 1, \ldots, n$. For the nodes $v \in V \setminus \{s\}$ with no incoming arcs ($in(v) = \emptyset$) put $v = 0$. For the starting node s write $s = 1$. The system S is the conjunction of equations $e(v)$ for each node $v \in V$ of the multigraph G.

Now, solve the system S by mutual substitution eliminating the node variables $v_i \in V$. This means that we transform a conjunction of equations

$$(v = (\phi_1 \wedge w) \vee \psi) \wedge (w = \phi_2) \quad \text{to} \quad (v = (\phi_1 \wedge \phi_2) \vee \psi) \wedge (w = \phi_2).$$

where v and w are node variables. This operation resembles the merge rule in syntactic unification. The exhaustive application of substitution derives the solved system S^*. If the original oriented multigraph G was acyclic then S^* contains the equation $t = \phi$, where ϕ is a Boolean formula upon the clauses $c(a_i)$ for $a_i \in A$. If there is a cycle in G, say $v_0 \xrightarrow{c_1} v_1 \xrightarrow{c_2} \cdots \xrightarrow{c_k} v_k \xrightarrow{c_{k+1}} v_0$, then we end up with the equation $e(v_0)$ equivalent to $v_0 = (c_1 \wedge \cdots \wedge c_{k+1} \wedge v_0) \vee \psi$. If we recursively substitute v_0, we get an equation equivalent to $v_0 = (c_1 \wedge \cdots \wedge c_{k+1} \wedge v_0) \vee (c_1 \wedge \cdots \wedge c_{k+1} \wedge \psi) \vee \psi$. Hence, each disjunct will contain the formula ψ. Therefore, by absorption, the equation $e(v_0)$ is equivalent to $v_0 = \psi$. The exhaustive application of absorption to recursive equations derives a solved system S^* containing the equation $t = \phi$ for a Boolean formula ϕ without node variables v_i.

We show that there exists a constrained path from s to t if and only if the Boolean formula ϕ in the equation $t = \phi$ from the solved system S^* is satisfiable. Suppose that there exists a constrained path p: $s = v_0 \xrightarrow{c_1} v_1 \xrightarrow{c_2} \cdots \xrightarrow{c_k} v_k = t$. Then the solved system S^* contains the equations equivalent to $s = v_0 = 1$, $v_1 = c_1 \vee \phi_1$, $v_2 = (c_1 \wedge c_2) \vee \phi_2$, $\cdots$, $t = v_k = (c_1 \wedge \cdots \wedge c_k) \vee \phi_k$ for some Boolean formulas ϕ_i, $i = 1, \ldots, k$. From the existence of the constrained path p follows that the conjunction $c_1 \wedge \cdots \wedge c_k$ is satisfiable. Hence, also the formula $\phi = (c_1 \wedge \cdots \wedge c_k) \vee \phi_k$ is satisfiable.

Conversely, suppose that the solved system S^* constrains the equation $t = \phi$ with the satisfiable Boolean formula ϕ. Suppose that the incoming arcs are $in(t) = \{a_i = w_i \longrightarrow t \mid w_i \in W\}$ for some $W \subseteq V$. If the solved system S^* contains the equations $w_i = \phi_i$ then the Boolean formula ϕ is equivalent to $(c(a_1) \wedge \phi_1) \vee \cdots \vee (c(a_k) \wedge \phi_k)$, where $c(a_i)$ are the clauses labelling the arcs a_i, respectively. There must be a node $w_i \in W$ with a satisfiable formula ϕ_i and a satisfiable clause $c(a_i)$, since ϕ is satisfiable. We perform the proof by induction on the length of the constrained path. If the path is of length 1, then the set W must contain the starting node s. If $s = w_i$ then $\phi_i = 1$ and there exists the constrained path $s = w_i \xrightarrow{c(a_i)} t$. Now assume that there exists a constrained path p of length n from s to w_i, where $w_i = \phi_i$ is the equation in S^*, such that ϕ_i and $c(a_i)$ are satisfiable. Hence, there exists the constrained path $p \cdot a_i$ from s to t of the length $n + 1$, since the formula ϕ is satisfiable as it contains the satisfiable disjunct $c(a_i) \wedge \phi_i$. $\Box$

References

[Pap94] C.H. Papadimitriou. *Computational complexity.* Addison-Wesley, 1994.