Complexity of Canadian Traveler Problem Variants

Dror Fried[a,*], Solomon Eyal Shimony[a], Amit Benbassat[a], Cenny Wenner[b]

[a]*Department of Computer Science, Ben-Gurion University of the Negev P.O.B. 653 Beer-Sheva 84105 Israel*
[b]*KTH - Royal Institute of Technology and Stockholm University, NADA, Att: Cenny Wenner SE-10044, Stockholm*

Abstract

The Canadian traveler problem (CTP) is the problem of traversing a given graph, where some of the edges may be blocked – a state which is revealed only upon reaching an incident vertex. Originally stated by Papadimitriou and Yannakakis (1991), the adversarial version of CTP was shown to be PSPACE-complete, with the stochastic version shown to be #P-hard.

We show that stochastic CTP is also PSPACE-complete: initially proving PSPACE-hardness for the dependent version of stochastic CTP, and proceeding with gadgets that allow us to extend the proof to the independent case.

Since for disjoint-path graphs, CTP can be solved in polynomial time, we examine the complexity of the more general remote-sensing CTP, and show that it is NP-hard even for disjoint-path graphs.

Keywords: Canadian Traveler Problem, Complexity of Navigation under Uncertainty, Stochastic Shortest Path with Recourse

1. Introduction

In the *stochastic Canadian traveler problem* (CTP) [1] we are given an undirected connected weighted graph $G = (V, E)$, a source vertex ($s \in V$), and a target vertex ($t \in V$). Any edge $e \in E$ may be *blocked* with a known

---

*Corresponding author
Email addresses: `friedd@cs.bgu.ac.il` (Dror Fried), `shimony@cs.bgu.ac.il` (Solomon Eyal Shimony), `amitbenb@cs.bgu.ac.il` (Amit Benbassat), `cenny@cwenner.net` (Cenny Wenner)

probability $p(e)$. The actual state of each edge $e \in E$ becomes known only upon reaching a vertex incident on e. Traversing an unblocked edge e incurs a non-negative cost equal to the weight of e. The problem is to find a policy π that minimizes the expected traversal cost $C(\pi)$ from s to t.

CTP formalizes a basic question of navigating in a partially known environment, which is a fundamental task for transportation, autonomous robotic systems, computer games, and more. Other variants of CTP have been introduced and analyzed in the research literature [2, 3, 4]. There has been a strong recent resurgence of interest in CTP, both theoretical [5, 6] and empirical [7, 8, 9]. A preliminary alternative proof of Theorem 2 appears in an unpublished work by one of the authors [10].

When originally introduced in [1], two variants were examined: the adversarial variant and the stochastic variant. The adversarial variant was shown to be PSPACE-complete by reduction from QSAT. For the stochastic version, membership in PSPACE was shown, however only #P-hardness was established by reduction from the *st-reliability problem*, leaving the question of PSPACE-hardness open. Apparently proving the stronger result requires some form of dependency between the edges, achieved "through the back door" in the adversarial variant. This paper settles the question, showing that CTP is indeed PSPACE-complete.

Since the size of an optimal policy is potentially exponential in the size of the problem description, we in fact show that it is PSPACE-hard to find even the optimal first action at s.

We begin with a variant of CTP with dependent directed edges, *CTP-Dep*, which allows for a simple proof of PSPACE-hardness by reduction from QSAT, before proceeding with the proof for the "standard" stochastic CTP. Although the latter result subsumes the former, proving the dependent CTP result first greatly simplifies the intuition behind the proof of the standard case.

Another variant we explore is remote-sensing CTP, henceforth called *Sensing-CTP*, in which additional actions called *remote-sensing actions* are allowed. Each such action reveals, for a certain cost, the status of a non-incident edge. Recently it was shown [9] that stochastic CTP can be solved in low-order polynomial time on disjoint-path graphs. It was believed that generalizing CTP to allow remote-sensing actions makes the problem harder – indeed we show that allowing remote-sensing makes CTP NP-hard even on disjoint-path graphs.

2. Dependent directed CTP is PSPACE-hard

This general form of *dependent CTP* (called CTP-Dep) is a 5-tuple (G, w, s, t, B) with $G = (V, E)$ a directed graph, a weight function $w : E \to \Re^{\geq 0}$, $s, t \in V$ are the start and goal vertices respectively, and a distribution model B over binary random variables indexed by the edges E. We assume that B is specified as a Bayes network over these random variables E [11] as follows. Each random variable expresses the state (blocked, unblocked) of an edge in E (abusing notation we use the symbols indicating the edges to denote the respective random variables). The Bayes network (E, A, P) consists of a set of directed arcs A between the random variables E, so that (E, A) is a directed acyclic graph. P describes the conditional probability tables, one for each $e \in E$.

Theorem 1. *CTP-Dep is PSPACE-hard.*

PROOF. by reduction from QSAT [12]. Recall that QSAT is the language of all satisfiable quantified boolean formulas (QBF), $\Phi = \forall x_1 \exists x_2 ... \varphi(x_1, x_2, ..., x_n)$, where φ is a boolean formula in conjunctive normal form, with n variables and m clauses, which contain literals, each is consisting of either a variable or a negated variable. We assume that each clause has at most 3 literals. Given a QBF Φ, construct a CTP-Dep instance (G_Φ, w, s, t, B) as follows (see Fig. 1). G_Φ consists of a *variables section*, and an *exam section*. Vertices in the variables section have labels starting with v or o, and vertices of the exam section begin with r. An always unblocked edge (s, t), called the *default edge*, has a cost of h. All other edges, unless mentioned otherwise, are zero-cost edges known to be unblocked.

The variables section contains a subsection X_i for every variable x_i, which begins at v_i, and ends at v_i'. For every $i < n$, X_i is connected to X_{i+1} through an edge (v_i', v_{i+1}).

Every X_i contains a *true-path* $(v_i, v_{i1}, \cdots, v_{im}, v_i')$, and a *false-path* $(v_i, \bar{v}_{i1}, \cdots, \bar{v}_{im}, v_i')$. If x_i is a universal variable (resp. existential variable), the edges (v_i, v_{i1}), and $(v_i, \bar{v}_{i1})$ are called *universal edges* (resp. *existential edges*). While the existential edges are always unblocked, we set the universal edges to have have blocking probability $1/2$ and to be mutually exclusive: for each universal variable x_i, exactly one of (v_i, v_{i1}), and $(v_i, \bar{v}_{i1})$ is blocked.

In addition, for every $1 \leq i \leq n$, and $1 \leq l \leq m$, there are edges (o_{il}, v_{il}), and $(\bar{o}_{il}, \bar{v}_{il})$, called *observation edges*. These edges are only meant to be observed, as their source vertices are unreachable. Every observation edge
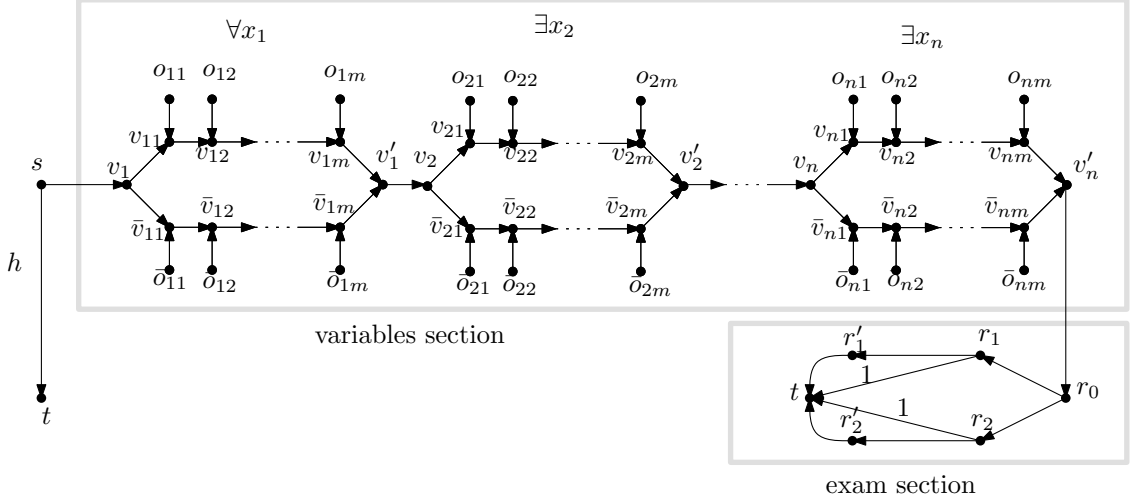
Figure 1: Reduction from QBF to CTP-Dep. Note that vertex t appears twice in order to simplify the physical layout.

is blocked with probability $1/2$, and the dependency of the observation edges is defined according to appearance of variables in the clauses of Φ, as follows: an observation edge (o_{il}, v_{il}) (resp. $(\bar{o}_{il}, \bar{v}_{il})$) is considered "in" a clause C_l if x_i appears unnegated (resp. negated) in clause C_l. All observation edges that are "in" the same clause C_l co-occur: they are either all traversable or all are blocked (with probability $1/2$, as stated above), independent of all other edges that are not "in" C_l.

The exam section consists of an *odd-path* (r_0, r_1, r'_1, t), and an *even-path* (r_0, r_2, r'_2, t). In addition construct edges (r_1, t), and (r_2, t) with cost 1. The edges (r_1, r'_1), and (r_2, r'_2) are called *choice edges*. The edge (r_1, r'_1) (resp. (r_2, r'_2)) is unblocked if and only if the observation edges are unblocked for an *odd* (resp. *even*) number of clauses. Hence exactly one of the choice edges is blocked [1]. If at least one observation edge in each clause is observed, the status of the choice edges can be determined with certainty. Otherwise the posterior blocking probability of each choice edge remains $1/2$. In order to prove the theorem, it is sufficient to prove the following claim:

Claim 1. *An optimal policy has expected cost 0 just when Φ is satisfiable*

---

[1] Note that as every clause has at most three literals, this dependency structure can be realized with a Bayes network of constant in-degree, a construction that has polynomial size.

4

(in which case the optimal first action is to traverse (s, v_1)). Otherwise (for any $h < 2^{-\frac{n}{2}-1}$) the optimal policy is to traverse (s, t) with a cost of h.

PROOF. Suppose first that Φ is satisfiable. Then there is a policy for assigning values to all the existential variables, each given every setting of the enclosing universal variables, such that φ is true. Following this policy for each existential variable x_i, i.e. traversing edge (v_i, v_{i1}) if x_i should be *true*, and $(v_i, \bar{v}_{i1})$ otherwise, leads (by construction) to following a path such that at least one observation edge is seen in every clause. Hence, the "exam" is passed (i.e. the 0-cost unblocked path in the exam section is chosen) with certainty.

Next, suppose Φ is not satisfiable. Then there is at least one setting of the universal variables for which some clause C_l is false under the same conditions, and thus no edge "in" clause C is observed. Since every setting of these variables occurs with probability $2^{-\frac{n}{2}}$ (assuming w.l.o.g. that n is even), in these cases the exam is "flunked" (picking the path where only the expensive edge is unblocked) with probability $1/2$, and thus the total expected cost of starting with (s, v_1) is at least $2^{-\frac{n}{2}-1}$. Hence, setting $h < 2^{-\frac{n}{2}-1}$, the optimal policy is to traverse (s, t) if and only if Φ is not satisfiable. $\square$

3. Complexity of CTP

Having shown that CTP-Dep is PSPACE-hard, we extend the proof to the "standard" stochastic independent undirected edges CTP:

Theorem 2. *CTP is PSPACE-complete.*

In order to prove Theorem 2, we use the same general outline of the reduction from QBF as in the proof of Theorem 1. However, in CTP-Dep, dependencies and directed edges restrict the available choices, thereby simplifying the proof. Here we introduce special gadgets that limit choice de facto, and show that any deviation from these limitations is necessarily suboptimal. Policies that obey these limitation are called *reasonable policies*. Each such gadget g has an *entry* terminal $Entry(g)$, and an *exit* terminal $Exit(g)$; an attempt to traverse g from $Entry(g)$ to $Exit(g)$ is henceforth called to *cross* g. The gadgets operate by allowing a potential *shortcut* to the target t; crossing these gadgets may either end up at $Exit(g)$, with some probability $q(g)$, or at t instead. The edges that allow direct access to t are called *shortcut* edges.

5

We introduce the gadgets in sections 3.1 and 3.2, and the CTP-graph construction in Section 3.3. The actual proof of Theorem 2 is in Section 3.4. In the description of the gadgets and CTP-graph, we sometimes add zero cost always traversable edges. These edges, which appear unlabeled in figures 2,3 and 4, were added solely in order to simplify the physical layout as a figure; any u, v connected by such an edge can be considered to be the same vertex.

3.1. Baiting Gadgets

A *baiting gadget* $g = BG(u,v)$ with parameter $L > 1$ is a three-terminal weighted graph (see Fig. 2): an entry terminal $u = Entry(g)$, an exit terminal $v = Exit(g)$, and a shortcut terminal which is always t. The latter terminal is henceforth omitted in external reference to g, for conciseness.

The baiting gadget consists of $N + 1$ uniform sections of an undirected path $(u, v_1, \cdots, v_N, v)$ with total weight L, each intermediate vertex has a 0-cost shortcut to t with a blocking probability $1/2$. In addition, there is a shortcut edge with cost L from the terminals u, v to t. Set $N = 2^{\lceil \log_2(4L) \rceil} - 1$. We assume that g is connected to the graph such that any policy executed at u, in which the edge (u, v_1) is not traversed, has an expected cost of at least 1. Later on we see that this assumption holds in the CTP-graph construction.

Let π be the following partial policy: *when at u for the first time, proceed along the path $(u, v_1, \cdots, v_N, v)$ to v, taking the 0-cost shortcut to t whenever possible, but never backtracking to u.*

It is easy to show that even if we need to take the cost L shortcut at v, the expected cost of executing π at u for the first time is less than 1. Because of the L cost shortcut edge (v,t), the expected cost of any optimal policy once at v (knowing all 0-cost shortcuts are blocked) is no more than L, hence under any reasonable policy, g is not retraced. A similar argument holds for retracing to u from other locations along the path $(u, v_1, \cdots, v_N, v)$. Hence we have:

Claim 2. *When at u for the first time, π is optimal for a baiting gadget $BG(u,v)$ with a parameter $L > 1$. After reaching v, it is suboptimal to backtrack to u in g.*

Note that g is actually symmetric w.r.t. u, v. However, since by construction of the CTP-graph, every reasonable policy always reaches one designated terminal u first, we treat g externally as if it were directional. A
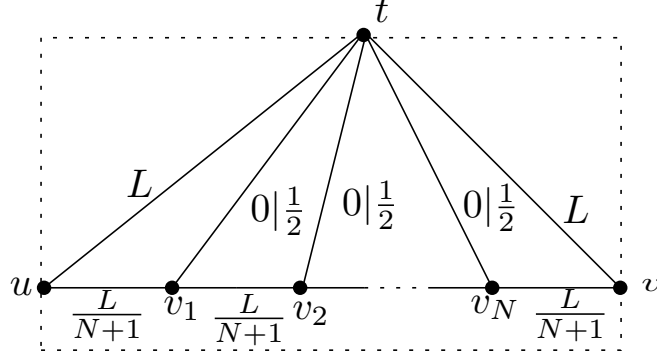
6

Figure 2: A Baiting Gadget $BG(u,v)$ with a parameter $L > 1$. Edge label $w|p$ denotes cost|blocking probability. The optimal policy at u is to cross the path $(u, v_1, \cdots, v_N, v)$, taking a shortcut edge to t whenever such an edge is found unblocked. After reaching v, retracing to u in g costs at least L.

precise derivation of the parameters of baiting gadgets appears in Appendix A.1.

3.2. Observation Gadgets

An *observation gadget* $g = OG(u, v, o)$ is a four-terminal weighted graph (see Fig. 3): an entry terminal $u = Entry(g)$, an exit terminal $v = Exit(g)$, an observation terminal o, and a shortcut terminal (again omitted in external references) which is always t. The observation gadget begins with a baiting gadget $BG_1 = BG(u, v_1)$ with parameter $L > 8$, which is connected to the "observation loop" beginning with a baiting gadget $BG_2 = BG(v_1, v_2)$ with parameter $3L/2$, a zero-cost edge (v_2, v_3) with blocking probability $3/4$, and a cost $L_1 = 5L/8$ unblocked edge to o. An always unblocked $3L/2$ shortcut edge (v_2, t) is assumed. The observation loop is closed by a cost L_1 unblocked edge to v_4 and a zero-cost edge (v_4, v_1) with blocking probability $3/4$. From v_1, a cost 1 unblocked edge (v_1, v') followed by a baiting gadget $BG_3 = BG(v_1', u)$ with parameter L completes the gadget.

We assume that o is either not directly connected to the rest of the graph, or connected through a path $(r_2, r_3, r_4, r_5, r_1', r_2')$ called the *exam section path* (o is identified with r_5) with the following properties: the edges (r_2, r_3), (r_1', r_2') and (r_4, r_5), have zero cost and blocking probability p_1, where $p_1 > 1 - 2/(3L + 1)$. (r_2, r_3) and (r_1', r_2') are called *guard edges*, (r_4, r_5) is called an *observation edge*. The edges (r_3, r_4) and (r_5, r_1') are always traversable edges with cost 1. The vertex o is allowed to coincide with observation terminals of other observation gadgets. The notations of
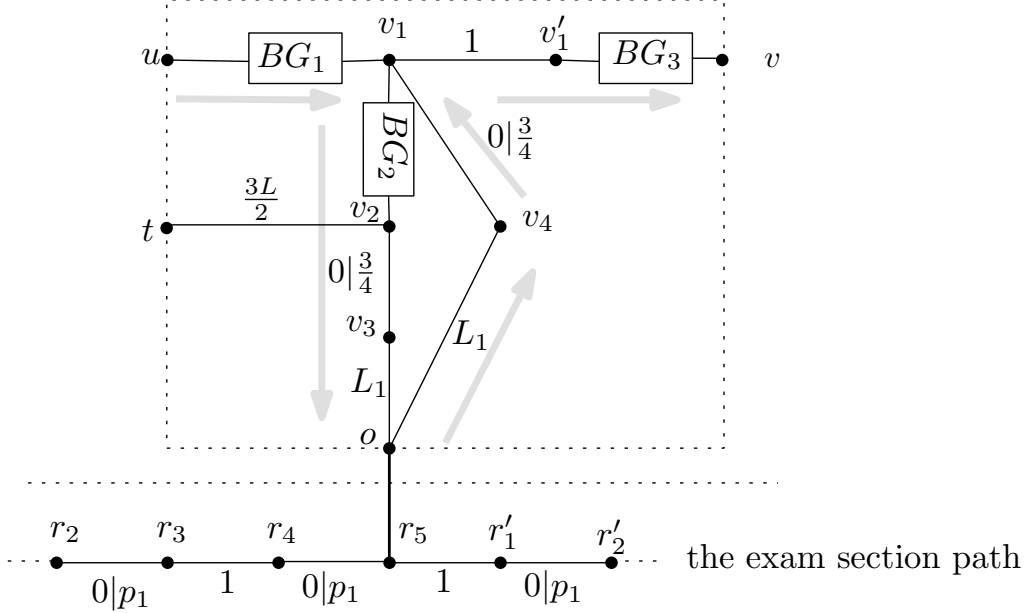
7

Figure 3: An Observation Gadget $OG(u, v, o)$. Light gray arrows indicate general traversal direction of the optimal policy π.

the exam section path are chosen to match the description of the CTP-graph construction in Section 3.3.

Let π_g be the following partial policy at g: *when at u, cross BG_1. Then (observing (v_1, v_4)), cross BG_2. If either (v_2, v_3), or (v_1, v_4) is found blocked, reach t by traversing the cost $3L/2$ shortcut edge (v_2, t). However, if both (v_2, v_3) and (v_1, v_4) are unblocked, traverse $(v_2, v_3, o, v_4, v_1, v_1')$ (observing any edges incident on o such as the observation edge (r_4, r_5)), and cross BG_3.*

Again, by construction of the CTP-graph (section 3.3), any policy at u other than crossing BG_1 results in a cost of at least 1.

Claim 3. *When at u for the first time, π_g is an optimal policy for g.*

PROOF OUTLINE. Properties of the baiting gadgets ensure that g is traversed in the correct order. The guard edges (r_2, r_3) and (r_1', r_2') ensure that it is suboptimal to "escape" from o by traversing edges in the exam section. The uncertain edges (v_4, v_1) and (v_2, v_3) ensure that it is suboptimal to enter a previously uncrossed observation gadget from o. Likewise for a previously crossed observation gadget g': entering g' through o is suboptimal because

8

all the baiting gadgets in g' have been crossed and observed to contain no unblocked zero-cost shortcuts.

Detailed derivation of the properties of observation gadgets appears in Appendix A.2.

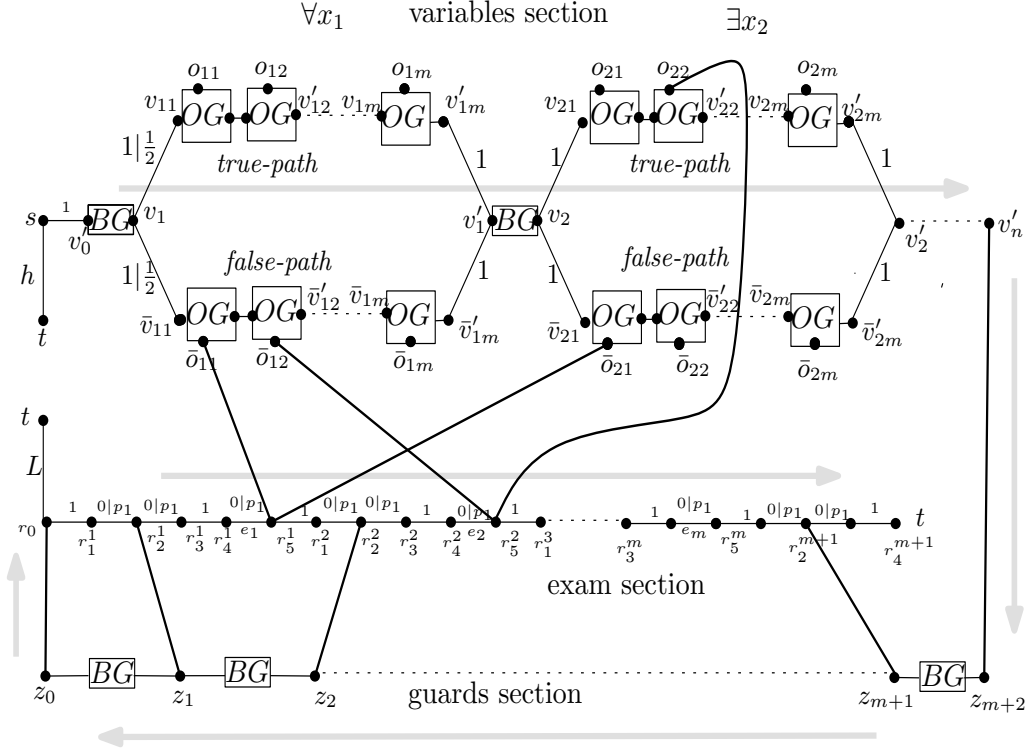### 3.3. CTP-graph Construction

Figure 4: CTP-graph construction for $\Phi = \forall x_1 \exists x_2 \cdots (\bar{x}_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee x_2) \cdots$. BG - a baiting gadget. OG - an observation gadget. Light gray arrows indicate the general traversal direction of the optimal policy when Φ is satisfiable.

Having shown the properties of the baiting and observation gadgets, we are ready to construct the CTP-graph: For a QBF Φ with n variables and m clauses, we construct G_Φ in the same general outline as the construction of the CTP-Dep graph (see section 2) with the following changes (see Fig. 4). The exam section is a path of $5(m+1)+1$ vertices $\{r_j^i | 1 \leq i \leq m+1, 1 \leq j \leq 5\}$ as follows. For every $0 < i \leq m+1$, (r_1^i, r_2^i), (r_2^i, r_3^i) and (r_4^i, r_5^i) have zero cost and blocking probability p_1, apart from (r_4^{m+1}, r_5^{m+1}) which has zero cost and is always traversable. (r_1^i, r_2^i), and (r_2^i, r_3^i) are called *guard edges*, and

9

(r_4^i, r_5^i) is called a *clause edge*. (r_3^i, r_4^i) and (r_5^i, r_1^{i+1}) are always traversable cost 1 edges. In addition, the exam section holds an additional vertex r_0, an always traversable cost 1 edge (r_0, r_1^1), and an always traversable cost L edge (r_0, t). In order to guarantee correct operation of the observation gadgets, we disallow reasonable policies to traverse exam edges too early while crossing the variable section. This is done by visiting the initially uncertain guard edges only later via a section called the *guards section*, which consists of a sequence of baiting gadgets $BG(z_i, z_{i+1})$ with parameter L that visit r_2^i for all $0 \leq i \leq m+1$.

The variables section is constructed as for CTP-Dep, except that the directed edges (v_i', v_{i+1}) are replaced by baiting gadgets $BG(v_i', v_{i+1})$ with parameter L. For each universal variable x_i the universal edges (v_i, v_{i1}), and $(v_i, \bar{v}_{i1})$ are cost 1 edges with blocking probability $1/2$. For each existential variable x_i, the existential edges (v_i, v_{i1}), and $(v_i, \bar{v}_{i1})$ are always traversable edges with cost 1. Inside each *true-path*, every (v_{ij}, o_{ij}), $(v_{ij}, v_{i(j+1)})$ pair is replaced by an observation gadget $g = OG(v_{ij}, v_{ij}', o_{ij})$. $(v_{ij}', v_{i(j+1)})$ are always unblocked zero-cost edges added for clarity. The observation vertex o_{ij} is identified with the vertex incident on the appropriate clause edge in the exam section. That is, if x_i appears unnegated in clause j, then o_{i_j} of the *true-path* is identified with r_5^j in the exam section. Likewise respectively for all the edges in the *false-paths*.

For example, Fig. 4 demonstrates the reduction for $\Phi = \forall x_1 \exists x_2 \cdots (\bar{x}_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee x_2) \cdots$. The variable x_1 appears negated in clause 2, so in G_Φ the vertex $\bar{o}_{12}$ at the section X_1, and the vertex r_5^2 of the exam section are connected by an unlabeled edge, hence the clause edge $e_2 = (r_4^2, r_5^2)$ can be observed from the observation gadget $OG(\bar{v}_{12}, \bar{v}_{12}', \bar{o}_{12})$ when traversing the *false* path of X_1. Likewise, the connection of other observation gadgets can be explained similarly.

3.4. Proof of Theorem 2

Given a QBF Φ with n variables and m clauses, we construct G_Φ as in Section 3.3. Set $L = 8m + 16$ and $p_1 = 1 - 2^{-\left\lceil \log_2(\frac{3L+1}{2}) \right\rceil}$. We show that it is optimal to traverse (s, v_0) if and only if Φ is satisfiable.

Unless stated otherwise, we henceforth consider only reasonable policies for G_Φ that do not begin with the default action of traversing (s, t). Due to properties of the gadgets (claims 2, 3) any reasonable policy π for G_Φ must follow the restrictions in Table 1, as any other action is suboptimal.

Most of these restrictions are immediate consequences of executing optimal policies at the baiting and observation gadgets (see Appendix A.1 and

Table 1: Reasonable policy actions in π

Location	Action
v_i', for $i < n$	cross $BG(v_i', v_{1+1})$
v_i, for $i \leq n$,	go to v_{i1} or $\bar{v}_{i1}$
v_{il}, for $i \leq n$,	cross $OG(v_{il}, v_{il}', o_{il})$
$\bar{v}_{il}$ for $i \leq n$,	cross $OG(\bar{v}_{il}, \bar{v}_{il}', \bar{o}_{il})$
z_l, for $0 < l \leq m + 2$	cross $BG(z_l, z_{l-1})$
r_0	pass exam or take shortcut

Appendix A.2 for detail). The following claim, proved in Appendix B.1, shows the actions of any reasonable policy for G_Φ at r_0.

Claim 4. *At r_0, any reasonable policy acts as follows:*

- *If all the edges in the exam section were observed to be unblocked, cross $(r_0, r_1^1, \cdots, r_4^{m+1}, t)$ until reaching t for a cost of $2(m+1)$.*

- *Otherwise, cross the cost L shortcut edge (r_0, t).*

Therefore, reasonable policies for G_Φ differ only in the choices made in the universal and existential edges, and in the choice at r_0 which is either to traverse the exam section if all clause edges were observed, or otherwise take the expensive shortcut (r_0, t).

Now let π be a reasonable policy for G_Φ, and denote the expected cost of π by $C(\pi)$. Define a *weather* to be an assignment of $\{traversable, blocked\}$ to the edges of G_Φ. Let W be the set of all possible weathers for G_Φ, and for $w \in W$ let p_w be the probability that weather w occurs. Define $C(\pi, w)$ to be the cost of executing π over a weather w. Then

$$C(\pi) = \sum_{w \in W} p_w C(\pi, w) \tag{1}$$

Next, partition W into *full-trip weathers* $W^f(\pi)$, in which r_0 is reached while executing π; and *shortcut weathers* $W^s(\pi)$ in which r_0 is not reached due to taking a shortcut edge to t before reaching r_0. Then:

$$C(\pi) = \sum_{w \in W^s(\pi)} p_w C(\pi, w) + \sum_{w \in W^f(\pi)} p_w C(\pi, w) \tag{2}$$

11

Let π^T be a policy for G_Φ such that in every subsection X_i of the variables section, whenever possible, the *true-path* is always chosen. Define:

$$D_{st} = \sum_{w \in W^s(\pi^T)} p_w C(\pi^T, w) \tag{3}$$

As all the *true-paths*, and *false-paths* of all the variables section are symmetric in the number of observation gadgets and other edges, there is a bijection $g_\pi : W^s(\pi) \to W^s(\pi^T)$ such that $p_w = p_{g_\pi(w)}$ and $C(\pi, w) = C(\pi^T, g_\pi(w))$ for every $w \in W^s(\pi)$. Hence we have:

$$D_{st} = \sum_{w \in W^s(\pi)} p_w C(\pi, w)$$

and

$$C(\pi) = D_{st} + \sum_{w \in W^f(\pi)} p_w C(\pi, w) \tag{4}$$

Again, due to symmetry, and the properties of the baiting and observation gadgets (claims 2, 3), the total cost from s to r_0 while executing π in any weather $w \in W^f(\pi)$ is independent of w. We denote this cost by D_{pt}, and can compute it simply by summing the cost of traversing from s to r_0 through the variable section and guard section, assuming that r_0 is reached. Then we get:

$$D_{pt} = 1 + (2 + \frac{19mL + 4}{4})n + (n + m + 1)L \tag{5}$$

Then from r_0 to t the cost is either $2(m + 1)$ (if the exam section is known to be completely unblocked), or $L > 2(m + 1)$ (taking the shortcut (r_0, t), if some edges in the exam section are known to be blocked, or some such unknown edges remain). Hence for any full-trip weather w, $C(\pi, w)$ is either $D_{pt} + L$, or $D_{pt} + 2(m + 1)$.

Let $P_\Phi^\pi \in [0, 1]$ be the probability that **not** all the clause edges of the exam section were observed in a full-trip weather by following π (this probability depends on the formula Φ). Then, with probability $(1 - P_\Phi^\pi)(1 - p_1)^{3m+2}$ all the edges of the exam section were observed and were found unblocked before reaching r_0. Let $P_{rt} = (1 - p_1)^{3m+2}$ be the probability that all the edges in the exam section are unblocked and denote by P_{r_0} the probability of reaching r_0 by executing π. Again, due to symmetry of the baiting and observation gadgets, P_{r_0} is independent of π. We get:

$$\sum_{w \in W^f(\pi)} p_w C(\pi, w) = P_{r_0}(D_{pt} + P_\Phi^\pi L + (1 - P_\Phi^\pi)(P_{rt}2(m+1) + (1 - P_{rt})L)) \quad (6)$$

And therefore

$$C(\pi) = D_{st} + P_{r_0}(D_{pt} + P_\Phi^\pi L + (1 - P_\Phi^\pi)(P_{rt}2(m + 1) + (1 - P_{rt})L)) \quad (7)$$

If Φ is satisfiable, then, as in the proof of Theorem 1, there is a reasonable policy π which follows the assignments that satisfy Φ, thus every clause edge is observed, and $P_\Phi^\pi = 0$. Define $B_0 = C(\pi)$ for such a policy π when Φ is satisfiable. Then

$$B_0 = D_{st} + P_{r_0}(D_{pt} + P_{rt}2(m + 1) + (1 - P_{rt})L) \quad (8)$$

If Φ is not satisfiable, then at some universal subsection X_i of the variables section, there is a probability of at least $1/4$ that a universal edge must be traversed, such that upon reaching r_0, not all the clause edges are visited. Hence, in total, there is a probability of at least $(\frac{1}{4})^{\frac{n}{2}}$ that not all the clause edges are visited. Note that as P_{r_0} already excludes events where both universal edges are blocked for some variable, if Φ is not satisfiable, then for every reasonable policy π, $P_\Phi^\pi > (\frac{1}{3})^{\frac{n}{2}}$. Hence define B_1 as follows.

$$B_1 = D_{st} + P_{r_0}(D_{pt} + (\frac{1}{3})^{\frac{n}{2}}L + (1 - (\frac{1}{3})^{\frac{n}{2}})(P_{rt}2(m + 1) + (1 - P_{rt})L)) \quad (9)$$

Then $B_1 > B_0$, and if Φ is not satisfiable, then $C(\pi) \geq B_1$. Now let $h = w((s,t)) = B_0 + (\frac{1}{4})^{\frac{n}{2}}mP_{r_0}$, so that $B_1 > h > B_0$. Thus the optimal action at s is to traverse (s,t) if and only if Φ is unsatisfiable. Since the CTP-graph construction used a polynomial number of vertices, and all the weights and probabilities by construction need only a polynomial number of bits (see Appendix B.2 for the technical computation of h), Theorem 2 follows. $\qquad\square$

Several corollaries follow due to the construction of G_Φ:

Corollary 1. *It is PSPACE-hard to determine the expected cost of the optimal policy in CTP.*

By replacing all the edges with appropriately directed edges, we get:

13

Corollary 2. *CTP with directed edges but no dependencies is PSPACE-complete.*

Finally, as every unknown edge in this construction of G_Φ has cost 0 and a probability which is a power of 2 of being unblocked (the universal edges, for example, can be split into a two-edge path), we can replace every unknown edge with a path of zero-cost, blocking probability $1/2$ edges and get:

Corollary 3. *CTP remains PSPACE-complete even if all the unknown edges have zero cost and blocking probability $1/2$.*

4. Complexity of CTP with remote sensing

A somewhat more general version of CTP is *Sensing-CTP*. In this variant, the state of graph edges can be remotely sensed at any time, paying a known cost. Formally, Sensing-CTP is defined exactly the same way as stochastic CTP, w.r.t. the graph, edge-blocking probabilities and weights, and source and target vertices (see Section 1). In addition, a sensing cost function $SC : V \times E \to \Re^+$ is given. An edge e, not necessarily incident on a vertex v, can be *sensed* for a cost $SC(v, e)$ and as a result, the true state of e is revealed. The problem in Sensing-CTP is to find a policy which minimizes the total expected sensing cost plus travel cost from s to t.

CTP is solvable in polynomial time when the graph consists of edge-disjoint paths which meet only at s and t [9]. This gives rise to the question whether Sensing-CTP is also tractable for disjoint paths. We show that this is not the case for Sensing-CTP unless P=NP. Again, since the size of the policy may be exponential in the size of the graph, we actually show that it is NP-hard to determine even the first action in an optimal policy.

Theorem 3. *Sensing-CTP is NP-hard even in disjoint-path graphs.*

PROOF. By reduction from the NP-complete problem vertex cover (VC) [12]. Let $G = (V, E)$ be an undirected graph, for which we need to decide if there is an $S \subseteq V$ of size at most k such that every edge in E is incident on a vertex in S (S is called a vertex cover of size k). The idea of the proof is, informally, for a policy to benefit if it 'has sensed' all of the edges in a given VC instance, where sensing of all neighbors of a vertex can be done at some constant cost. By tweaking constants, these actions will be beneficial

14

if and only if it is possible to sense all edges in the VC instance by sensing the neighbors of at most k vertices.

Construct the corresponding CTP-PATH graph $G' = (V', E')$ as follows (see Fig. 5). For each vertex $v \in V$, construct a "vertex" node $f(v) \in V'$, an edge $(s, f(v))$ with a cost C defined below, and an infinite-cost edge $(f(v), t)$. Also construct the "default edge" (s, t) with cost 4. The above edges are always traversable.

Construct the "sensing path", a path consisting of a

"leader" always traversable edge e_0 with cost L starting at s, and for each edge $e \in E$, a zero-cost edge $f(e)$ in sequence, and finally an infinite-cost edge (u, t). The probability that each of the latter edges is blocked is ϵ, a small positive number defined below, except for (u, t) which has infinite-cost, therefore is never traversable. Additionally, we have a two edge "uncertain" path $((s, x), (x, t))$ where (s, x) is always traversable and costs 2, and (x, t) is traversable with probability $1/2$ and costs 0. Note that the resulting graph G' consists only of edge-disjoint paths leading from s to t.
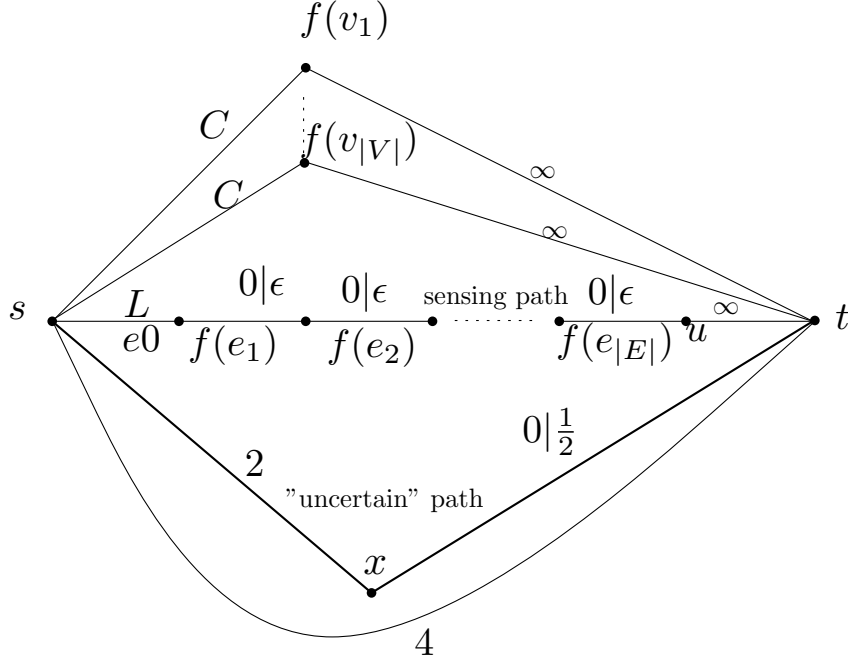

Figure 5: CTP-graph for reduction from vertex cover

The sensing cost from "vertex" node $f(v)$ on edge $f(e)$ is 0 for all e incident on v in G. Sensing (x, t) costs 0 from u. All other sensing costs

15

are infinite. We show that for parameter values defined below, the optimal policy is to immediately traverse (s, t) if and only if G does not have a vertex cover of size k.

Assuming no sensing, there are only two "reasonable" traversal policies: traversing (s, t) immediately for a cost of 4, or trying the uncertain path, which costs 2 if (x, t) is traversable, and 8 if not, for an expected cost of 5. Thus the optimal policy with no sensing is to traverse (s, t). However, if sensing is allowed, the "value of information" of knowing whether (x, t) is traversable is: $VOI(x, t) = 4 - (2 \times 1/2 + 4 \times 1/2) = 1$, due to the fact that if (x, t) is revealed as traversable (which happens with probability $1/2$) we gain 2 by taking the "'uncertain"' path instead of (s, t), and otherwise gain nothing.

We show below that for appropriate values of L, it is not beneficial to try to get to u in order to sense (x, t) unless sensing reveals that the path to u is unblocked. The probability that at least one edge on the sensing path is blocked is at least ϵ. However, all the edges in the sensing path can be sensed (stopping if any blocked edge is sensed) for a cost C_{sense} defined below. If G has a vertex cover of size k, then $C_{sense} \leq 2Ck$ by visiting k "vertex" nodes. However, if the smallest vertex cover is of size $k' \geq k + 1$, the expected sensing cost becomes:

$$C_{sense} = 2C \sum_{i=1}^{k'} \prod_{j=1}^{i-1} (1 - \epsilon)^{d_j} \geq 2Ck'(1 - \epsilon)^{|E|} \tag{10}$$

where d_j is the number of previously unsensed edges of the jth vertex in the (unknown) optimal sensing order. For any $0 < \alpha < 1$, set ϵ such that

$$\epsilon = 1 - \left(\frac{k + 1 - \alpha}{k + 1}\right)^{\frac{1}{|E|}}$$

Then, as $2Ck'(1 - \epsilon)^{|E|} \geq 2C(k + 1 - \alpha)$, we have that

$$C_{sense} \geq 2C(k + 1 - \alpha)$$

Now set $L = \frac{1}{2} - \frac{\epsilon}{4}$ and:

$$C = \frac{\epsilon(1 - \epsilon)^{|E|}}{2(2k + 1 - \alpha)}$$

To complete the proof, it is sufficient to prove the following claim:

Claim 5. *The optimal action at s is to traverse (s, t) if and only if G does not have a vertex cover of size k.*

PROOF. We show the following:

- If there is no cover in size k, the optimal policy is to traverse (s,t).

- Otherwise, the optimal policy is as follows: *visit the "vertex" nodes constituting the cover, doing the appropriate sensing actions; if the sensing path is unblocked, visit u in order to sense (x, t), and then take the path from s to t through x if (x, t) is found traversable.* (In this case, given the optimal policy, one can straightforwardly construct the vertex cover of size k.)

Note that it is suboptimal to try the sensing path unless assured that all edges leading to u are traversable, because then there is at least one edge that can be blocked with probability ϵ, in which case attempting this path results in no positive gain. To see this, note that the traversal cost $2L$ must be paid anyway, thus the total expected gain from trying the sensing path is:

$$g \leq (1 - \epsilon)VOI(x, t) - 2L = 1 - \epsilon - 2L = 1 - \epsilon - 2(\frac{1}{2} - \frac{\epsilon}{4}) = -\frac{\epsilon}{2} < 0$$

This also holds for all policies that attempt some sensing actions before trying the sensing path, but that do not make sure that u is reachable before trying the sensing path.

Now, if there is a vertex cover of size k, the expected cost of sensing all edges in the sensing path is at most $2Ck$. If u is found to be reachable (prior probability $(1 - \epsilon)^{|E|}$ of that happening) use the sensing path (which costs $2L$) to sense (x, t), gaining the expected $VOI(x, t)$ of 1. The total expected gain in this case is positive:

$$g' \geq (1-\epsilon)^{|E|}(VOI(x, t) - 2L) - 2Ck = (1-\epsilon)^{|E|}(1 - (1 - \frac{\epsilon}{2})) - 2k\frac{\epsilon(1 - \epsilon)^{|E|}}{2(2k + 1 - \alpha)}$$

$$= \epsilon(1 - \epsilon)^{|E|}(\frac{1}{2} - \frac{2k}{2(2k + 1 - \alpha)}) > 0$$

Finally, if there is no vertex cover of size k, then the expected gain for sensing in the best case is (assuming the policy of sensing all edges in the
17

sensing path from only $k+1$ "vertex" nodes, costing at least $2C(k+1-\alpha)$ in expectation) is at most:

$$g'' \leq (1-\epsilon)^{|E|}(VOI(x,t)-2L) - 2C(k+1-\alpha) =$$

$$(1-\epsilon)^{|E|}(1-(1-\frac{\epsilon}{2})) - 2(k+1-\alpha)\frac{\epsilon(1-\epsilon)^{|E|}}{2(2k+1-\alpha)}$$

$$= \epsilon(1-\epsilon)^{|E|}(\frac{1}{2} - \frac{2(k+1-\alpha)}{2(2k+1-\alpha)}) < 0$$

and thus the optimal policy here is not to perform sensing at all, but to traverse (s,t) immediately. $\qquad\square$

5. Discussion

Having shown that stochastic CTP is PSPACE-hard, several related questions on variants of CTP and CTP with restricted topologies arise. One issue of particular interest is the question of efficiently finding approximately optimal actions. The proofs in this paper make use of rather small gaps between expected values of two candidate actions, and thus leave open the possibility of efficient approximation algorithms.

Studies of the *competitive analysis* of the Canadian Traveler Problem reveal rudimentary bounds on approximability. Denoting by k the number of uncertain edges in an instance, there exists for the undirected case polynomial-time algorithms achieving competitive ratios of $2k+1$ [6]. As a consequence, stochastic CTP can be approximated within $2k+1$. With a slightly improved analysis, the same algorithm yields a $2n+1$-approximation. In the directed case, existing results from competitive analysis only yield approximations of $2^{k+1}+1$ and $2^{n+1}+1$, respectively [5].

These approximation algorithms forego entirely the stochastic nature of the problem and leave open considerable improvements. At the time of this writing, no notable hardness of approximation results are known.

Another issue is: what is the most general graph topology under which CTP is tractable or easy to approximate? An efficient algorithm was shown [9] for disjoint-path graphs, based on a lemma that there exists an optimal policy that is committing: (such a policy never returns to the source vertex unless its 'current' path to the target vertex is known to be blocked). But any departure from the disjoint-path structure (such as adding even one more edge that crosses between vertices in two of the paths) complicates things considerably by voiding the optimality of committing policies.

The *st*-reliability problem (finding the probability that an unblocked path from s to t exists) appears to be an essential building block in solving CTP. The reduction in [1] shows that CTP is intractable for almost any graph topology for which *st*-reliability is intractable as well. An open research question is whether CTP is tractable or easy to approximate for graph topologies in which *st*-reliability is tractable, such as "tree-structured" graphs, or the more general series-parallel graphs.

Sensing-CTP is generally harder than CTP for restricted topologies, as shown in Section 4. Other variants of CTP, such as CTP-Dep, are also harder than CTP, as dependencies can act like remote sensing. It was shown that CTP-Dep is NP-hard for disjoint path graphs [13]. However, when considering topological restrictions, one must also consider the topology of the dependency-graph, and the hardness proof in [13] used an essentially unrestricted topology Bayes network to represent the dependencies.

6. Acknowledgments

References

[1] C. H. Papadimitriou, M. Yannakakis, Shortest Paths Without a Map, Theor. Comput. Sci. 84 (1) (1991) 127–150.

[2] A. Bar-Noy, B. Schieber, The Canadian Traveller Problem, in: SODA, 1991, pp. 261–270.

[3] E. Nikolova, D. R. Karger, Route Planning under Uncertainty: The Canadian Traveller Problem, AAAI (2008) 969–974.

[4] G. H. Polychronopoulos, J. N. Tsitsiklis, Stochastic shortest path problems with recourse, Networks 27 (2) (1996) 133–143.

[5] Y. Xu, M. Hu, B. Su, B. Zhu, Z. Zhu, The canadian traveller problem and its competitive analysis, J. Comb. Optim. 18 (2) (2009) 195–205.

[6] S. Westphal, A note on the k-Canadian Traveller Problem, Inf. Process. Lett. 106 (3) (2008) 87–89.

[7] P. Eyerich, T. Keller, M. Helmert, High-Quality Policies for the Canadian Traveler's Problem, in: In Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2010), 2010.

[8] Z. Bnaya, A. Felner, D. Fried, O. Maksin, S. E. Shimony, Repeated-task canadian traveler problem, in: SOCS, 2011.

[9] Z. Bnaya, A. Felner, S. Shimony, Canadian Traveler Problem with Remote Sensing, IJCAI (2009) 437–442.

[10] C. Wenner, Hardness results for the shortest path problem under partial observability, Diploma Thesis, Lund University.

[11] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, CA, 1988.

[12] M. Garey, D. Johnson, Computers and Intractability, W.H. Freeman, 1979.

[13] D. Zarchy, Theoretical and experimental analysis of the canadian traveler problem, Master Thesis, Ben-Gurion University of the Negev.

Appendix A.

Appendix A.1. Baiting Gadgets

Let $g = BG(u, v)$ be a baiting gadget with a parameter $L > 1$, defined in Section 3.1 (see Fig. 2). Recall that π (as defined in Section 3.1) is the following policy for g: *At u, traverse (u, v_1). At v_i, for any $i < N$, do as follows: if (v_i, t) is unblocked, reach the destination through (v_i, t) for cost of zero. However, if (v_i, t) is blocked, traverse (v_i, v_{i+1}) for a cost of $L/(N+1)$. At v_N, if (v_N, t) is blocked, traverse (v_N, v).*

By construction of the CTP-graph, we assume that any policy other than traversing (u, v) results in a cost of at least 1 (see Section 3.1).

Apart from π, other policies at u that are not clearly suboptimal are:

- *choose not to traverse (u, v_1).*

- The following type of policies denoted by π_j, for $j \leq N$: *execute π until reaching v_j; if (v_j, t_j) is unblocked, reach the destination through (v_j, t_j); otherwise, retreat to u and execute an optimal policy with an expected cost of $M_j \geq 1$.*

Finally, we set $N = 2^{\lceil log_2(4L) \rceil} - 1$, implying $N + 1 \geq 4L$.

Claim 2. *When at u for the first time, π is optimal for a baiting gadget $BG(u, v)$ with a parameter $L > 1$. After reaching v, it is suboptimal to backtrack to u.*

PROOF. Denote by K the expected cost of the optimal policy executed once v is reached. As there is an L cost shortcut edge (v, t), it is clear that $K \leq L$ therefore it is always suboptimal to retrace g once v is reached. We first show that $C(\pi) < 1$, hence choosing not to traverse (u, v_1) is suboptimal.

Note that for every $i \leq N$, the probability that (v_i, v_{i+1}) is traversed in π is $(\frac{1}{2})^i$. Hence we have

$$C(\pi) = \frac{L}{N+1} \sum_{i=0}^{N} (\frac{1}{2})^i + (\frac{1}{2})^N K \qquad (A.1)$$

Thus

$$C(\pi) = \frac{2L}{N+1}(1 - 2^{-(N+1)}) + 2^{-N}K \qquad (A.2)$$

Then, as $K \leq L$, $N + 1 \geq 4L$ and $L > 1$, we have that

$$C(\pi) < \frac{2L}{4L} + 2^{-N}L < \frac{3}{4} < 1. \qquad (A.3)$$

As required.

Next we show that for every $j \leq N$, $C(\pi) < C(\pi_j)$, hence for every $j \leq N$, the policy π_j is suboptimal. We have that

$$C(\pi_j) = \frac{L}{N+1} \sum_{i=0}^{j-1} (\frac{1}{2})^i + (\frac{1}{2})^j (\frac{jL}{N+1} + M_j) \qquad (A.4)$$

Thus

$$C(\pi_j) = \frac{2L}{N+1}(1 - 2^{-j}) + \frac{2^{-j}jL}{N+1} + 2^{-j}M_j \qquad (A.5)$$

As $K \leq L$, and $1 < M_j$, it is sufficient from (A.2) to show that for every $0 < j \leq N$

$$\frac{2L}{N+1}(1 - 2^{-(N+1)}) + 2^{-N}L < \frac{2L}{N+1}(1 - 2^{-j}) + \frac{2^{-j}jL}{N+1} + 2^{-j} \qquad (A.6)$$

Therefore

$$\frac{2L}{N+1}(2^{-j} - 2^{-(N+1)} - 2^{-j-1}j) + 2^{-N}L < 2^{-j}$$

Hence we need to show that for every $0 < j \leq N$,

$$\frac{2L}{N+1}(1 - 2^{j-N-1} - \frac{j}{2}) + 2^{j-N}L < 1$$

As $N + 1 \geq 4L$ and $L > 1$, it is sufficient to show that for every $0 < j \leq N$,

$$\frac{1}{4}(2-j) + 2^{j-N}L < 1 \tag{A.7}$$

And inequality (A.7) follows since the function

$$f(x) = \frac{1}{4}(2-x) + 2^{x-N}L$$

over the reals, has only one extremum, $f(0) < 1$,$f(N) < 1$ and

$$\lim_{x \to \infty} f(x) = \lim_{x \to -\infty} f(x) = \infty$$

Appendix A.2. Observation Gadgets

Let $g = OG(u, v, o)$ be an observation gadget as defined in Section 3.2, and seen in Fig. 3. Recall that $L > 8$, and $L_1 = 5L/8$. π_g is the following partial policy for $OG(u, v, o)$: *At u, cross BG_1 (observe (v_1, v_4)). Then cross BG_2. If either (v_1, v_4), or (v_2, v_3) is found blocked, reach t by traversing the $3L/2$ cost shortcut edge (v_2, t). However, if both (v_1, v_4), and (v_2, v_3) are unblocked, traverse $(v_2, v_3, o, v_4, v_1, v_1')$ (at o, observe the edges incident on o, in case there are any), and cross BG_3.*

We again assume, by construction of the CTP-graph, that any policy at u other than crossing BG_1 results in a cost of at least 1.

Claim 3. *When at u for the first time, π_g is an optimal policy for g.*

PROOF. At u, as BG_1 is a baiting gadget, then by Claim 2, it is optimal to cross BG_1. When first arriving at v_1, after BG_1 is crossed, (v_1, v_4) is observed. As (o, v_4) has a cost of $L_1 > 1$, and (v_1, v_1') has a cost of 1, then by Claim 2, it is optimal to cross BG_2. Once at v_2, if (v_2, v_3) is blocked, it is optimal to take the shortcut (v_2, t) for a cost of $3L/2$.

It remains to show that if (v_2, v_3) is unblocked, the optimal policy at v_2 is:

1. if (v_1, v_4) is unblocked, traverse $(v_2, v_3, o, v_4, v_1, v_1')$, and cross BG_3.
2. otherwise, traverse the shortcut (v_2, t) for a cost of $3L/2$.

case 1: (v_1, v_4) is unblocked.

First note that arriving at v_1 a second time through (v_4, v_1), BG_1 and BG_2 are known not to have any blocked shortcut edges, hence by Claim 2,

22

the optimal policy when arriving at v_1 a second time is to traverse (v_1, v_1') and the baiting gadget BG_3.

Now, traversing $(v_2, v_3, o, v_4, v_1, v_1')$, and crossing BG_3 bears an expected cost of at most $2L_1 + 2$, while traversing (v_2, t) costs $3L/2$. Hence, as $2L_1 + 2 < 3L/2$, it is optimal at v_2 to traverse (v_2, v_3, o).

We now inspect the possible partial policies at o:

case 1.a: Traverse (o, v_4, v_1, v_1'), and cross BG_3 for an expected cost of at most $L_1 + 2$. We denote this partial policy by π'.

case 1.b: Traverse edges of another observation gadget $\widetilde{g}$, in case there exists such $\widetilde{g}$ incident on o. Suppose that $\widetilde{g}$ has not already been traversed (label the vertices of $\widetilde{g}$ as $\widetilde{v}_i$). Then traversing either $(o, \widetilde{v}_3)$ and trying to traverse $(\widetilde{v}_3, \widetilde{v}_2)$, or traversing $(o, \widetilde{v}_4)$ and trying to traverse $(\widetilde{v}_4, \widetilde{v}_1)$ results in an expected cost of at least $L_1 + 3L_1/4$. Hence, as $L_1 + 2 < L_1 + 3L_1/4$, we have that executing π' is cheaper than traversing any edges of $\widetilde{g}$.

Next suppose that $\widetilde{g}$ has already been traversed. Therefore we may assume that the policy $\pi_{\widetilde{g}}$ was executed in $\widetilde{u}$, thus the baiting gadgets of $\widetilde{g}$ are known not to contain any unblocked zero-cost shortcuts, hence crossing each such baiting gadget costs L. Then traversing $\widetilde{g}$ results in an expected cost of at least $L_1 + L$, and as $L_1 + 2 < L_1 + L$, we again have that executing π' is cheaper than traversing any edges of $\widetilde{g}$.

case 1.c: Traverse the exam section path. Recall that o is identified with r_5. Suppose that observation edge (r_4, r_5) is blocked. At o, denote the following partial policy by π_1: *cross (r_5, r_1'); if (r_1', r_2') is unblocked, continue with any optimal policy, otherwise, return to o, and execute π', which can still be executed, for an expected cost of $C(\pi') < L_1 + 2$.* Then we have

$$C(\pi_1) \geq 1 + p_1(1 + C(\pi'))$$

and as $p_1 > 1 - 2/(3L + 1)$, we have that

$$C(\pi') < 1 + p_1(1 + C(\pi')) < C(\pi_1)$$

Therefore executing π' is cheaper than executing π_1.

23

Now suppose that (r_4, r_5) is unblocked. Then we can either execute π_1 (or the symmetric case in which (r_2, r_3) is being inspected) with the same analysis, or we can extend π_1 with the following policy denoted by π_2:

execute π_1; upon returning to o (after (r'_1, r'_2) is found blocked), cross (r_5, r_4) and (r_4, r_3); if (r_3, r_2) is unblocked, continue with any optimal policy; otherwise return to r_5, and execute π', which can still be executed for an expected cost of $C(\pi') < L_1 + 2$. Then we have that

$$C(\pi'_2) \geq 1 + 2p_1 + p_1^2(1 + C(\pi'))$$

However, $p_1 > 1 - 2/(3L + 1)$ entails $C(\pi') < C(\pi'_2)$. Therefore executing π' is cheaper than executing π_2 as well. The policy in which (r_2, r_3) is the first edge among (r_2, r_3) and (r'_1, r'_2) to be inspected is symmetric to π_2. Hence we see that traversing any edges of the exam section path is suboptimal.

case 2: (v_1, v_4) is blocked.

In this case the following partial policies can be executed at v_2:

case 2.a: Take the shortcut edge (v_2, t) for a cost of $3L/2$. Denote this policy by π'.

case 2.b: Traverse (v_2, v_3) and (v_3, o) for a cost of L_1 and at o traverse edges of another observation gadget. As in case 1.b we have that traversing edges of $\widetilde{g}$ results in an expected cost of at least $L_1 + 3L_1/4$. Then, as $3L/2 < L_1 + L_1 + 3L_1/4$, we have that π' is cheaper than reaching o and traversing any edges of $\widetilde{g}$.

case 2.c: Traverse (v_2, v_3) and (v_3, o), for cost of L_1, and at o traverse the exam path. We define π_1 and π_2 as in case 1.c. Recall that $C(\pi') = 3L/2$. However as $p_1 > 1 - 2/(3L + 1)$, we have that $C(\pi') < L_1 + C(\pi_1)$, and $C(\pi') < L_1 + C(\pi'_2)$. Hence π' is cheaper than traversing to o and traversing the any edges of the exam section path.

$\square$

Appendix B.

Appendix B.1. Behavior of reasonable policies

Claim 4. *At r_0, any reasonable policy acts as follows. If all the edges in the exam section were observed to be unblocked, cross $(r_0, r_1^1, \cdots, r_4^{m+1}, t)$ until reaching t for a cost of $2(m+1)$. Otherwise, cross the cost L shortcut edge (r_0, t).*

PROOF. Retracing $BG(z_0, z_1)$ clearly results in a cost of at least L. We first see that unless all the edges in the exam section were observed to be unblocked, any partial policy executed at r_0 results in a cost of at least L, therefore it is cheaper to take the shortcut (r_0, t) for a cost of L.

At every vertex r_i^l, $l \leq m+1$, $, i \leq 5$, any unblocked edge on the exam section path, incident on r_i^l, can be traversed. At r_2^l there is an additional option to cross either $BG(z_l, z_{l+1})$, or $BG(z_{l-1}, z_l)$ which hold no unblocked shortcut edges, hence crossing these results in a cost of at least L. If r_5^l is identified with an observation point of some observation gadget g', there is an additional option to traverse edges of g. However by an argument identical to case 1.b of Appendix A.2, traversing any edges of g results in a cost of at least L. Hence any deviation from the exam section path results in a cost of at least L.

Suppose that all the edges of the exam section are known to be unblocked. Then, as the exam section contains $2(m+1)$ always traversable cost 1 edges, and as $2(m+1) < L$, the optimal policy is to cross the exam section $(r_0, r_1^1, \cdots, t)$ for the cost of $2(m+1)$.

Otherwise, suppose there are edges in the exam section with unknown status. As all the guard edges were observed upon crossing the guard section, then such an edge is a clause edge. Hence let $e_l = (r_4^l, r_5^l)$ be the first unknown clause edge such that every edge in the path $(r_0, \cdots, r_4^l)$ is known to be unblocked. Finding (r_4^l, r_5^l) blocked results in either retracing the exam section to r_0 and taking the cost L shortcut to t, or in deviating from the exam section. Hence, as (r_3^l, r_4^l) costs 1, $L > 1$ and $p_1 > 1 - 2/(3L+1)$, traversing from r_0 to e_l results in an expected cost of at least $1 + p_1(1 + L) > L$. Hence traversing the shortcut edge (v_2, t) is cheaper. Obviously, the same argument holds for traversing e_l where e_l is previously known to be blocked. $\square$

Appendix B.2. Polynomial size representation

We show the computation of h, the cost of the default edge (s, t). Recall that $L = 8m + 16$, $N = 2^{\lceil \log_2(4L) \rceil} - 1$ and $p_1 = 1 - 2^{-\lceil \log_2(\frac{3L+1}{2}) \rceil}$.

From Section 3.4 we have that

$$h = B_0 + (\frac{1}{4})^{\frac{n}{2}} m P_{r_0}$$

and

$$B_0 = D_{st} + P_{r_0}(D_{pt} + P_{rt}2(m+1) + (1 - P_{rt})L)$$

where for a reasonable policy π, P_{r_0} is the probability of reaching r_0 by executing π, D_{pt} is the total cost from s to r_0 while executing π in a full-trip weather, D_{st} is the expected cost of executing π over shortcut weathers, and P_{rt} is the probability that all the edges in the exam section are unblocked.

From section 3.4, $P_{rt} = (1 - p_1)^{3m+2}$ and

$$D_{pt} = 1 + (2 + \frac{19mL + 4}{4})n + (n + m + 1)L$$

Hence it is left to compute P_{r_0} and D_{st}. To do that we define the following.

For $k > 0$, let $G(k)$ be a CTP instance composed of a series of gadgets g_i, $1 \le i \le k$, such that for every $i < k$, $Exit(g_i)$ is identified with $Entry(g_{i+1})$. The gadgets in $G(k)$ are either all baiting gadget with a parameter $L > 1$ (then $G(k)$ is denoted by $BG(L, k)$), or are all observation gadgets (then $G(k)$ is denoted by $OG(k)$). Set s to be $Entry(g_1)$. Then we denote the following policy for $G(k)$ as π_k: *For every $i \le k$, cross g_i.*

Let $q(G(k))$ be the probability that $Exit(g_k)$ is reached by executing π_k. Let $w_1(G(k))$ be the expected traversal cost when $Exit(g_k)$ is reached while executing π_k. Let $w_2(G(k))$ be the expected cost in case a shortcut to t is taken while executing π_k. Then we have for $k > 1$

$$w_2(G(k)) = w_2(G(1)) + q(G(1))(w_1(G(1)) + w_2(G(k-1))) \qquad \text{(B.1)}$$

Next, if $G(k)$ is a series of baiting gadgets we have that $q(BG(L, k)) = 2^{-kN}$ and $w_1(BG(L, k)) = kL$. From (A.2) we obtain:

$$w_2(BG(L, 1)) = \frac{2L}{N+1}(1 - 2^{-(N+1)}) - 2^{-N}L \qquad \text{(B.2)}$$

If $G(k)$ is a series of observation gadgets, set

26

$$N_1 = 2^{\left\lceil \log_2\left(\frac{3}{2}4L\right)\right\rceil} - 1$$

Then we have $q(OG(k)) = 2^{-k(2N+N_1+4)}$, and $w_1(OG(k)) = \frac{k(19L+4)}{4}$. We compute $w_2(OG(1))$ to be:

$$w_2(OG(1)) = w_2(BG(L,1)) + 2^{-N}(w_1(BG(L,1))$$
$$+2^{-N}(w_2(BG(\frac{3L}{2},1)) + 2^{-N-N_1}(\frac{3L}{2}) + 2^{-N-N_1-4}(2L_1 + 1 + w_2(BG(L,1))))))$$
$$\text{(B.3)}$$

We can now compute P_{r_0}. Assume w.l.o.g. n is even. Due to symmetry of the *true/false-paths*, and the subsections of the variable sections, we have that

$$P_{r_0} = \left(q(BG(L,1))q(OG(L,m))\frac{3}{4}q(BG(L,1))q(OG(L,m))\right)^{\frac{n}{2}} q(BG(L,m+2))$$
$$\text{(B.4)}$$

To find D_{st} we note the following. Assume w.l.o.g n is even. Due to symmetric considerations, there are parameters q_{st}, w_{st} and z_{st}, independent of i, such that by executing π at v_i', v_{i+2}' is reached with probability q_{st} and an expected cost of w_{st}, and v_{i+2}' is *not* reached (i.e. a shortcut it taken) with an expected cost of z_{st}. Assume w.l.o.g. the subsection X_i is universal. Then

$$q_{st} = q(BG(L,1))\frac{3}{4}q(OG(L,m))q(BG(L,1))q(OG(L,m)) \qquad \text{(B.5)}$$

and

$$w_{st} = 2L + 4 + 2mw_1(OG(L,1)) \qquad \text{(B.6)}$$

We now compute

$$z_{st} = w_2(BG(L,1)) + 2^{-N}(w_1(BG(L,1)) + \frac{1}{4}L + \frac{3}{4}(1 + w_2(OG(L,m)) +$$
$$q(OG(L,m))(w_1(OG(L,m)) + 1 + w_2(BG(L,1)) +$$
$$2^{-N}(w_1(BG(L,1)) + 1 + w_2(OG(L,m)) + q(OG(L,m)) + 1)))) \quad \text{(B.7)}$$

Now define $\bar{z}_{st} = z_{st} + q_{st}w_{st}$, set $D_{st}^1 = \bar{z}_{st}$, and for every $k > 1$ set

$$D_{st}^k = D_{st}^1 + q_{st}D_{st}^{k-1} \tag{B.8}$$

Then

$$D_{st} = D_{st}^{\frac{n}{2}} + (q_{st})^{\frac{n}{2}} w_2(BG(L, m+1)) \tag{B.9}$$

which concludes the computation of h. $\qquad\square$