Mutual Opacity between Multiple Adversaries
C K Sharpe, S L Ricker, Hervé Marchand

To cite this version:
C K Sharpe, S L Ricker, Hervé Marchand. Mutual Opacity between Multiple Adversaries. Wodes 2022 - 16th IFAC Workshop on Discrete Event Systems, Sep 2022, Prague, Czech Republic. pp.1-7. hal-03773490

HAL Id: hal-03773490
https://inria.hal.science/hal-03773490
Submitted on 9 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract: We investigate opacity, an information-flow privacy property, in a setting where there are two competing agents or adversaries whose objective is to hide their secrets and expose the secrets of the other agent. Each agent has only partial information about the state of the system. The agents can achieve their objective by enabling or disabling events from their set of controllable events. We examine two different scenarios. In the first problem, the agents are passive with no control capabilities, and we seek a global controller to enforce their mutual opacity. In the second problem, the formerly passive agents are autonomous and have control capabilities. We seek the plausibility of two controllers, one for each agent, to see if we can synthesize a winning control strategy so that one adversary can always discover the secrets of the other without revealing its own.

1. INTRODUCTION

Security is one of the most critical and challenging aspects in designing services deployed on large open networks such as the Internet, mobile phones, and e-voting systems. Such services are regularly subject to malicious attacks. Thus, developing methods to ensure their security is crucial. Over the past ten years, there has been significant research devoted to developing formal methods for designing secure systems, leading to a growing interest in the formal verification of security properties.

There are well-established notions of security, but we are interested in those that focus on characterizing the information flow from the system to an intruder (Riccardo Focardi (1994); Goguen and Meseguer (1982); McCullough (1988); Wittbold and Johnson (1990)). The notion of opacity, introduced in Mazare (2004), formalizes the absence of information flow in a system. More precisely, opacity refers to the impossibility that an intruder can infer the truth about some confidential activity of the system. For example, such activity could be the occurrence of some particular sequence of actions in the system or the fact that the system is in a specific configuration.

Opacity was initially modeled using trace-based systems (Bryans et al., 2008), but we are interested in examining opacity through the lens of supervisory control theory (Ramadge and Wonham, 1987). We model system behavior and, subsequently, its security properties using finite state automata. The work of Dubreil et al. (2008, 2010) introduced the use of supervisory control for enforcing opacity with a single malicious intruder pursuing secrets. Since then, this scenario with a single attacker has been extensively examined; refer to Jacob et al. (2016) for a general survey of the use of supervisory control theory and security properties like opacity.

We are interested in examining strategies where multiple agents want to acquire secret information without compromising their secrets. The first work to examine opacity with multiple agents appeared in Badouel et al. (2007) and, more recently, in Helouet et al. (2018). In particular, we examine a variation on the concurrent secrets problem (Badouel et al. (2007)) where the goal for each agent is to uncover the secret of the other agent.

The paper is organized as follows. Section 2 establishes the notation and background on using supervisory control to enforce opacity. In Section 3 we present the two multi-agent problems that we are interested in examining. We first consider a setting where a global control strategy ensures secrets are not directly or inadvertently leaked among members of a group of agents. We address this problem in Section 4. Secondly, we propose a setting where local control is the focus since the goal of each agent is to force the disclosure of the secrets of other agents while preventing its secrets from leaking. This problem is addressed in Section 5. We discuss extensions of this work in Section 6.

2. PRELIMINARIES

In this section, we provide a brief summary of the notation used in the paper, along with a relevant background for the opacity problem.

2.1 Languages and Automata

Let $Σ$ be a finite alphabet (set of events) while $Σ^*$ denotes the set of all finite-length sequences or words of events of $Σ$, including the empty word $ε$. The length of a word $w ∈ Σ^*$ is denoted by $|w|$, where $|ε| = 0$. A language $L$ is a subset of words in $Σ^*$, i.e., $L ⊆ Σ^*$. Given words $t, u ∈ Σ^*$, the word $t.u$ denotes the concatenation of $t$ and $u$, i.e., the sequence of events of $t$ followed by the sequence of events of $u$. Given two words $w, w′ ∈ Σ^*$, we say that $w′$ is a prefix of $w$ whenever there exists $w′′ ∈ Σ^*$ such that
Given $\Sigma^*$ said to be prefix-closed is given by $\text{pref}(\Sigma^*) = \{w \in \Sigma^* : \exists w' \in \Sigma^*, w.w' \in \mathcal{L}\}$. $\mathcal{L}$ is said to be prefix-closed if $\text{pref}(\mathcal{L}) = \mathcal{L}$.

Given $\Sigma' \subseteq \Sigma$, we define the projection operator, $\pi_{\Sigma'} : \Sigma^* \rightarrow \Sigma'^*$, on finite words that removes, from a word of $\Sigma^*$, all the events that do not belong to $\Sigma'$. Formally, $\pi_{\Sigma'}$ is recursively defined as follows: $\pi_{\Sigma'}(e) = e$; for $w \in \Sigma^*$, $\pi_{\Sigma'}(w.\sigma) = \pi_{\Sigma'}(w).\sigma$, if $\sigma \in \Sigma'$; $\pi_{\Sigma'}(w)$, otherwise. The definition of projection for words extends to languages. Let $\mathcal{L} \subseteq \Sigma^*$. Then $\pi_{\Sigma'}(\mathcal{L}) = \{\pi_{\Sigma'}(w) \mid w \in \mathcal{L}\}$.

Alternatively, let $\mathcal{L}' \subseteq \Sigma'^*$. The inverse projection of $\mathcal{L}'$ is $\pi_{\Sigma'}^{-1}(\mathcal{L}') = \{w \in \Sigma^* \mid \pi_{\Sigma'}(w) \in \mathcal{L}'\}$. We will use the notation $[w]_{\pi_{\Sigma'}}$ whenever we refer to $\pi_{\Sigma'}^{-1}(\mathcal{L}')(w)$.

A deterministic finite automaton (DFA) is commonly used to represent the behavior of systems at a very abstract level. It is composed of a finite number of states and transitions between those states, labeled by events from its associated alphabet. Formally, a DFA is defined as $G = (Q, \Sigma, \delta, q_0)$, where $Q$ is a finite set of states, $\Sigma$ is the alphabet, $\delta : Q \times \Sigma \rightarrow Q$ is the (possibly partially-defined) state transition function, and $q_0 \in Q$ is the initial state. Note that we can extend $\delta$ over $\Sigma^*$ in the usual way. We say that a transition $\delta(q, w)$ is defined, denoted $\delta(q, w)$!, if $\delta(q, w) \neq \epsilon$. For a given state $q \in Q$, we define the active event set at $q$ as follows: $\Gamma(q) = \{\sigma \in \Sigma \mid \delta(q, \sigma)\}$. We can extend this to be defined over a set of states $Q' \subseteq Q$ such that $\Gamma(Q') = \bigcup_{q \in Q'} \Gamma(q)$.

A run of $G$ is a sequence of transitions of the form $r = q_0 \sigma_{q_1} \cdots \sigma_{|r|-1} q_{|r|}$, where $\delta(q_j, \sigma_j)$!, for all $j \in \{0, \ldots, |r| - 1\}$. Thus, the word $w$ associated with the run $r$ is the sequence of transition labels: $w = \sigma_0 \sigma_1 \cdots \sigma_{|r|-1} \sigma_{|r|}$. We say that $w$ labels the run $r$ of $G$. Let $\text{Run}(G)$ denote the set of runs of $G$.

A DFA $G$ generates a regular language, which is defined as $\mathcal{L}(G) := \{w \in \Sigma^* \mid \delta(q_0, w)!\}$. Equivalently, $\mathcal{L}(G)$ is the set of words that label the runs of $G$.

Given two DFAs $G_1 = (Q_1, \Sigma_1, \delta_1, q_{10})$ and $G_2 = (Q_2, \Sigma_2, \delta_2, q_{20})$, their parallel composition is defined as follows: $G_1 || G_2 = (Q_1 \times Q_2, \Sigma_1 \cup \Sigma_2, \delta, (q_{10}, q_{20}))$, where $\delta((q_1, q_2), \sigma) = (\delta_1(q_1, \sigma), \delta_2(q_2, \sigma))$, if $\delta_1(q_1, \sigma)$! and $\delta_2(q_2, \sigma)$!; $(\delta_1(q_1, \sigma), q_2)$, if $\delta_1(q_1, \sigma)$!( and $\sigma \notin \Sigma_2$; $(q_1, \delta_2(q_2, \sigma))$, if $\delta_2(q_2, \sigma)$!( and $\sigma \notin \Sigma_1$; undefined, otherwise.

Due to the limitation of the sensors attached to the system or the distributed/concurrent nature of the system, not all events can be seen. Therefore, we decompose the alphabet of the system into observable and unobservable events and we denote the set of observable events by $\Sigma_o \subseteq \Sigma$. Based on this partition, one can define the observable reach, namely, the set of states that can be reached from one state by a sequence of unobservable events:

$$\varepsilon\text{-reach}(q, \Sigma_o) = \{q\} \cup \{q' \in \mathcal{Q} \mid \exists t \in (\Sigma \setminus \Sigma_o)^*, \delta(q, t) = q'\}.$$

We can now construct an observer, using subset construction (Rabin and Scott, 1959), namely, a DFA whose language is the set of words defined over $\Sigma_o$.

**Definition 1.** Consider a DFA $G = (Q, \Sigma, \delta, q_0)$. The observer automaton for $G$, which is constructed with respect to $\Sigma_o \subseteq \Sigma$, is $\mathcal{O}_{\Sigma_o} = (X, \Sigma_o, \delta_{\Sigma_o}, x_0)$, where $X \subseteq 2^Q$, $\delta_{\Sigma_o} : X \times \Sigma_o \rightarrow X$ is the (partial) transition function such that $\delta_{\Sigma_o}(x, \sigma) = \bigcup_{q \in x} \varepsilon\text{-reach}(\delta(q, \sigma), \Sigma_o)$, and $x_0 = \varepsilon\text{-reach}(q_0, \Sigma_o)$ is the initial state.

### 2.2 Opacity

We consider a setting where no runs of a system $G$ should reveal any secret information. In this paper, we represent a secret as a subset of states $S \subseteq \mathcal{Q}$. Elements of $S$ can represent states in which a system is vulnerable to attacks or where confidential information is in use and is not yet declassified.

There exists an information flow whenever an intruder or attacker, denoted by $A$, can deduce confidential information on the execution of a system $G$ from the observation of a subset of events $S_A \subseteq \Sigma$. In our model, we want to prevent an attacker from knowing that the system is in a secret state.

Given $w \in \mathcal{L}(G)$, the observation of the attacker $A$ is given by the natural projection $\pi_{\Sigma_o}(w)$. We consider that attackers collect information online with executions of the system and try to infer with certainty whether the system’s current state belongs to $S$. Intuitively, opacity says that an attacker is never able to infer that a run is currently in a secret state based on its observation. An attacker observes the system and builds from its observation a belief, i.e., an estimation of a set of states in which the system could be.

**Definition 2.** Consider $G$ with a set of secret states $S$, an attacker $A$, who observes $S_A$ and its observer automaton $\mathcal{O}_{S_A}$. Then $S$ is said to be opaque w.r.t. $G$ and $S_A$ if $\forall w \in \mathcal{L}(\mathcal{O}_{S_A})$, $\delta_{S_A}(x_0, w) \notin S$.

Alternatively, we could have considered formulating secrets from a language-based point of view, where a set of secrets $\tilde{S} \subseteq \mathcal{L}(G)$ is provided as a non-prefixed-closed regular language (Dubreil et al., 2010). With a DFA $M_{\tilde{S}}$, such that $\mathcal{L}(M_{\tilde{S}}) = \tilde{S}$, we can reduce this problem to a state-based setting by considering the product of $G$ and $M_{\tilde{S}}$ to identify the secret states.

### 2.3 Supervisory control

In this paper, we will consider opacity with supervisory control, which consists of restricting the system behavior by pairing it with an access control so that control can be used to protect a secret(s).

Given a prefix-closed behavior $K \subseteq \mathcal{L}(G) \subseteq \Sigma^*$ expected from the system $G$, the goal of supervisory control is to enforce this behavior $K$ on $G$ by pairing this system with a monitor $S$ (also called controller). In the context of opacity, $K$ is a language that contains words that do not pass through secret states. To facilitate opacity enforcement on $G$, the alphabet $\Sigma$ is partitioned into a set
of controllable events $\Sigma_c \subseteq \Sigma_o$ and a set of uncontrollable events $\Sigma \setminus \Sigma_c$. After each of its observations, the controller $S$ issues a command of enable or disable for elements in $\Sigma_c$, whereas the uncontrollable events in $\Sigma \setminus \Sigma_c$ are permanently enabled. Formally, the controller is given by a function $S : L(\Omega_{\Sigma_c}) \rightarrow 2^{\Sigma_c}$. Alternatively, $S$ can be seen as a (possibly infinite) DFA $S_G = (Y, \Sigma_o, \delta_{S_G}, y_0)$ such that for each $\mu \in L(G_{\Sigma_o})$, $\delta_{S_G}(x_o, \mu) = S(\mu)$.

For implementational reasons, we put conditions on the admissible restrictions of $G$, namely controllability and normality.

**Definition 3.** A prefix-closed language $K \subseteq L(G)$ is

- **controllable** w.r.t. $L(G)$ and $\Sigma_c$ if $K.(\Sigma \setminus \Sigma_c) \cap L(G) \subseteq K$;
- **normal** w.r.t. $L(G), \Sigma_o$ if $K = [K]_{\Sigma_o} \cap L(G)$.

Under the assumption that $\Sigma_c \subseteq \Sigma_o$, there exists a supremal controllable and normal prefix-closed sub-language $K^+\Sigma$ of $K$ corresponding to the largest language included in $K$ that can be enforced by control and this language is regular (Wonham and Ramadge (1987); Kumar et al. (1991)). If $K^+$ is not empty, there exists a maximal controller $S$ such that $L(G/S) = K^+$, where $G/S = G||S_G$. To ensure controllability under partial observation, $G_{\Sigma_o}$ and $\Sigma_c$, the valid control policies for $S_G$ at state $y \in Y$ are given by $C(y) = \{ \gamma \in 2^{\Sigma_c} \mid \gamma \cap \Sigma_c = \Gamma(y) \cap \Sigma_c \}$, where $\Gamma(y)$ is the active event set at $y$.

### 3. OPACITY IN A MULTI-AGENT SETTING

We are interested in the problem of two adversarial agents $A_1$ and $A_2$, each of which independently observes the system $G$. Each agent has its own set of observable events $\Sigma_{o,i} \subseteq \Sigma_o$ and its own set of secret states $S_i \subseteq Q$, for $i \in \{1, 2\}$.

---

**Problem 1.** Given a system $G$ and two agents $A_1$, $A_2$ with respective secrets $S_1, S_2 \subseteq Q$, does there exist a supervisor $S$ such that when $G$ is controlled by $S$, $A_1$ is not able to discover the secret of $A_2$ and vice-versa.

---

**Fig. 1.** A multi-agent secret-keeping scenario.

The problem here is that each agent has the aim of finding the secret of the other agent. In that sense, we shall call them *adversaries* rather than attackers in the sequel. We will consider the following problems under this scenario:

1. **Static concurrent secrets:** Does there exist a maximal controller that can avoid an information leak from one adversary to the other, neither of which controls the system?
2. **Dynamic concurrent secrets:** Does there exist a local controller for $A_1$ to guess the secret(s) of $A_2$, who is simultaneously trying to control the system to guess the secret(s) of $A_1$?

Let $G = (Q, \Sigma, \delta, q_0)$ be a DFA representing the system. We assume the existence of two adversaries, each of which observes a (sub)set of observable and (possibly) controllable events, denoted by $\Sigma_{o,1}$ and $\Sigma_{c,i}$ respectively, for $i \in \{1, 2\}$. A possible controller acting globally upon $G$ is observing $\Sigma_{o,1} \subseteq \Sigma$ and controlling $\Sigma_c \subseteq \Sigma_o$. Let $O_i = (X_i, \Sigma_{o,i}, \delta_i, x_{o,i})$ be the observer automaton defined over $\Sigma_{o,i}$ (as per Definition 1, where for readability we use $\delta_i$ in lieu of $\delta_{o,i}$) for $A_i$ (for $i \in \{1, 2\}$), where each observer state is a set of estimates of which states $G$ is currently in, based on $A_i$’s partial observation of $G$.

We will use a running example, Ex. 1, to demonstrate the concepts of the paper, beginning with the observers of $G$.

**Example 1.** Consider the system $G$ shown on the left side of Fig. 2. Let $\Sigma_{o,1} = \{a, c\}$ and $\Sigma_{o,2} = \{b, c\}$ be the observable alphabets for $A_1$ and $A_2$. The unique observer automata $O_1$ and $O_2$ are shown in the middle and right of Fig. 2.

---

**Fig. 2.** $G$ (left) and its corresponding observers $O_1$ (center) and $O_2$ (right).

---

The approach for solving each problem will depend on the power and observation of the controllers as well as the powers of control and observation of the adversaries. For the first problem, we must synthesize a global controller that avoids leaking secrets from one adversary to the other, as the adversaries are passive. For the second problem, there is a concurrent game between the adversaries as they try to discover the secret(s) of the other. In this setting, the system is passive.

### 4. STATIC CONCURRENT SECRETS

In this section, we suppose that the adversaries cannot control the system (i.e., $\Sigma_{c,i} = \emptyset$), but must rely on an external controller to keep their secrets. The aim of the system, then, is to exert control so that no secrets are leaked between adversaries.

**Problem 1.** Given a system $G$ and two agents $A_1$, $A_2$ with respective secrets $S_1, S_2 \subseteq Q$, does there exist a supervisor $S$ such that when $G$ is controlled by $S$, $A_1$ is not able to discover the secret of $A_2$ and vice-versa.

---

**Fig. 3.** Static concurrent secrets with a global controller.
Based on results from (Dubreil et al. (2008, 2010)),
\[ F_i = L_{\gamma_i}(O_{\Sigma_{n,i}}) \]
represents the set of observations leading to an information flow from \( A_j \) to \( A_i \), i.e., \( w \in \Sigma_{n,i} \) such that \( \delta_i(w, x_{0,i}) \subseteq 2^S_i \).

From \( F_i \), we can build the automata \( G_i' \) such that \( F_i = L_{\gamma_i}(G_i') \) with \( S'_i = 2^{\Sigma_i} \) as the secret states. Further, it is sufficient to build \( G||G_i'||2 \) and consider the set of states
\[ Bad = (Q \times S'_2 \times X_{G_i'}) \cup (Q \times S'_1 \times S_i') \]
Assuming \( S \) observes \( \Sigma_o \) and controls \( \Sigma_c \subseteq \Sigma_o \) one can compute the supremal controllable and normal controller \( S \) that avoids the set of states \( Bad \) to solve our problem (Wonham and Ramadge, 1987; Kumar et al., 1991).

Example 2. We continue to work with \( G \) as in Fig. 2, but now we identify the secret states for each adversary: \( S_1 = \{1\} \) and \( S_2 = \{2\} \). The observers remain the same and we set the controllable event sets as follows \( \Sigma_{n,1} = \{a, c\} \) and \( \Sigma_{n,2} = \{b, c\} \). Note that there are now states in each \( O_i \) where the secrets are revealed. In particular, at state \( \{2\} \) in \( O_1 \), \( A_1 \) can discover \( S_2 \). Similarly, at state \( \{1\} \) in \( O_2 \), \( A_2 \) can discover \( S_1 \).

Fig. 4. \( G_1' \) (left) and \( G_2' \) (right)

Fig. 4 generates the automata that generate the minimal languages that end in an adversary’s secret state. A portion of \( G||G_1'||G_2' \) is shown in Fig. 5, where states belonging to \( Bad \) are shaded. To ensure that \( A_2 \) does not discover \( A_1 \)’s secret state, the controller disables the first occurrence of \( c \), while it disables the second occurrence of \( a \) to prevent \( A_1 \) from discovering \( A_2 \)’s secret state.

Fig. 5. A portion of \( G||G_1'||G_2' \). States in \( Bad \) are shaded grey.

5. DYNAMIC CONCURRENT SECRETS

In this section, we again assume that each adversary has a partial observation of the system. Thus, \( \Sigma_{o,i} \subseteq \Sigma_o \) (resp., \( \Sigma_{n,i} \)) is the set of observable (resp., unobservable) events of \( A_i \). Each adversary \( A_i \) (for \( i \in \{1, 2\} \)) has an associated observer automaton \( O_i \). To influence the discovery of their opponent’s secret state(s), both adversaries are able to issue local (i.e., their own) control decisions at each state \( x_i \in X_i \). Each agent \( A_i \) can control a subset of observable events \( \Sigma_{c,i} \subseteq \Sigma_o \). A control decision consists of a set of events to be disabled: \( \gamma_i \subseteq \Sigma_{c,i} \). A set of valid control policies \( C \) for \( A_i \) is thus, a mapping between observer states and possible control decisions, i.e., \( C_i : X_i \rightarrow 2^{\Sigma_{c,i}} \), with the assumption that \( \forall x_i \in X_i \), and \( \gamma_i \subseteq \Gamma(x_i) \cap \Sigma_{c,i} \), i.e., \( A_i \) can only disable controllable events that are admissible at \( x_i \). In the sequel, we shall denote the set of valid control policies by \( PC_i : X_i \rightarrow 2^{\Sigma_{c,i}} \), for \( i \in \{1, 2\} \).

In terms of control in the system, an event \( \sigma \) may only occur in \( G \) if neither adversary has disabled \( \sigma \). Thus, after each adversary has issued a control decision, \( \sigma \) may occur in \( G \) as long as \( \sigma \in \Gamma(q) \setminus (\gamma_1 \cup \gamma_2) \). If all events in \( \Gamma(q) \) have been disabled by at least one adversary, the system is stalled and no events can occur. Note that fully uncontrollable events, i.e., \( \sigma \in \Sigma \setminus (\Sigma_{o,1} \cup \Sigma_{o,2}) \), cannot be disabled, so \( G \) can only become stalled in states that have no transitions labelled with uncontrollable events.

Definition 4. Given \( O_i \), the observer of \( A_i \) for \( i \in \{1, 2\} \), a sequence of valid control policies \( C_i \) is denoted by \( (\gamma^j \in C_i')_{j \in \mathbb{N}} \) such that there exists a run in \( O_i \) of the form
\[ r = x_0 \xrightarrow{\sigma_0} x_1 \xrightarrow{\sigma_1} x_2 \cdots \xrightarrow{\sigma_{j-1}} x_j \xrightarrow{\sigma_j} x_{j+1} \cdots \]
with \( \gamma^j \subseteq \Gamma(x_j) \cap \Sigma_{c,j} \) and \( \sigma_j \in \Gamma(x_j) \setminus \gamma_j \).

Based on this definition, we can now identify the Dynamic Concurrent Secrets problem:

Problem 2. Given a system \( G \) and two agents \( A_1, A_2 \) with respective secrets \( S_1, S_2 \subseteq Q \), does there exist a sequence of valid control policies \( (\gamma^j \in C_i')_{j \in \mathbb{N}} \) for agent \( A_i \), such that for all possible sequences of valid control policies of the form \( (\gamma^j \in C_i')_{j \in \mathbb{N}} \) for agent \( A_2, A_1 \) can discover \( S_2 ^j \)?

To solve this problem, we define an arena
\[ H = (V_0 \cup V_5, E_0 \cup \Sigma, \delta_0, \delta_5, v_0, Bad_1, Bad_2) \]
which contains all the possible control policies for each agent, as well as the possible outcomes in the system \( G \). Here we describe the arena in further detail.

- The arena features two types of states:
  - \( V_0 \subseteq Q \times X_1 \times X_2 \times PC_1 \times PC_2 \) is the set of decision states where
    - \( q \in Q \) is the current state of \( G \);  
    - \( x_i \in X_i \) is the current belief state in \( O_i \) for \( A_i \);

1 Equivalently, this problem can be solved with respect to \( A_2 \) always being able to discover \( S_1 \).
adversary can choose any valid control policy \( \gamma \).

\( H \) are free to choose any new control policy according to its belief state both evolve to a new belief state. Similarly, for a state \( x \), keep track of their previous decisions, i.e., \( v \in V \) for \( A \).

\( A \) does not observe anything and, thus, its control decision does not result in a state change for \( A \), and we assume that \( C_i(x, \sigma) = \{ \gamma \} \) from the preceding state of the arena.

We denote by \( \text{Run}(H) \) the set of possible runs in \( H \). Then \( \text{Run}_{H}(A) \subseteq \text{Run}(H) \) is the set of runs that ends in a state of \( \text{Bad} \). Intuitively, for any runs in \( \text{Run}_{B_2} \), \( A_1 \) will reach the secret of its adversary \( A_2 \).

The language of \( H \) is given by

\[
\mathcal{L}(H) = \{ (\gamma_1, \gamma_2), \sigma \}^* \]

such that there exists a run \( r \in \text{Run}(H) \) of the form

\[
(v_0, v_1, \ldots, v_n) \in V_0 \times V_1 \times \cdots \times V_n.
\]

\( \delta_{DS} : V_0 \times \Sigma \rightarrow V_0 \) maps system states to decision states and corresponds to an occurring event in \( G \). For a given \( v = (q, x_1, x_2, \{ \gamma_1 \}, \{ \gamma_2 \}) \in V_0 \) and \( \sigma = (\gamma, \gamma) \in \Gamma(q) \setminus (\{ \gamma \} \cup \{ \gamma \}) \), we define the transition function as follows:

\[
\delta_{DS}(q, x_1, x_2, \{ \gamma_1 \}, \{ \gamma_2 \}), \sigma = \begin{cases} 
(q', x_1', x_2', \{ \gamma_1' \}, \{ \gamma_2' \}) & \text{if } \sigma = (\gamma, \gamma) \\
(q', x_1', x_2', \{ \gamma_1 \}, \{ \gamma_2' \}) & \text{if } \sigma = (\gamma, \gamma) \\
(q', x_1', x_2', \{ \gamma_1' \}, \{ \gamma_2 \}) & \text{if } \sigma = (\gamma, \gamma) \end{cases}
\]

and we define the transition function as follows:

\[
\delta_{DS}(q, x_1, x_2, \{ \gamma_1 \}, \{ \gamma_2 \}), \sigma = \begin{cases} 
(q', x_1', x_2', \{ \gamma_1' \}, \{ \gamma_2' \}) & \text{if } \sigma = (\gamma, \gamma) \\
(q', x_1', x_2', \{ \gamma_1 \}, \{ \gamma_2' \}) & \text{if } \sigma = (\gamma, \gamma) \\
(q', x_1', x_2', \{ \gamma_1' \}, \{ \gamma_2 \}) & \text{if } \sigma = (\gamma, \gamma) \end{cases}
\]

Intuitively, \( H \) represents all the possible valid control decisions of \( A_i \) in each state of \( G \), as well as the consequences of their decisions.

We now explain, in more detail, the different components of the arena \( H \).

Given a state \( v_0 = (q, x_1, x_2, C_1(x_1), C_2(x_2)) \in V_0 \), each adversary can choose any valid control policy \( \gamma \in C_i(x) \) according to its belief state \( i \). By triggering the tuple \( (\gamma_1, \gamma_2) \), we enter into state \( v_1 = (q, x_1, x_2, C_1(x_1), C_2(x_2), (\gamma_1, \gamma_2)) \) that "memorizes" the control decisions.

Similarly, for a state \( v_0 = (q, x_1, x_2, C_1(x_1), C_2(x_2), (\gamma_1, \gamma_2)) \in V_5 \), the system \( G \) can choose any event that is allowed by \( \gamma_1 \cup \gamma_2 \), i.e., \( \sigma = \Gamma(q) \setminus (\{ \gamma_1 \} \cup \{ \gamma_2 \}) \). If \( \sigma \) is unobservable to both the adversaries, i.e., \( \sigma = \Gamma(q) \setminus (\{ \gamma_1 \} \cup \{ \gamma_2 \}) \) then the belief state of each adversary \( A_i \) does not change and we keep track of their previous decisions, i.e., \( C_i(x) \) is reduced to the singleton \( \{ \gamma_i \} \). Thus, \( A_i \) is free to choose any valid control policy in \( C_i(x) \), with \( x_i' = \delta_i(x, \sigma) \), while \( A_2 \) does not observe anything and, thus, its control decision remains \( \gamma_2 \) (and vice-versa if \( i \) is only observable by \( A_2 \)). Finally, when \( A_1 \) is observable by both adversaries, then both evolve to a new belief state \( x_i' = \delta_i(x, \sigma) \) and both are free to choose any new control policy according to \( x_i' \) (for \( i \in \{ 1, 2 \} \)).

A run of \( H \), which begins at \( v_0 \), is of the form:

\[
\begin{align*}
&v_0 = (q_0, x_{0,1}, x_{0,2}, C_1(x_{0,1}), C_2(x_{0,2})) \\
&\downarrow (\gamma_1, \gamma_2), \gamma_i \in C_i(x_{0,1}) \\
&\quad (q_0, x_{0,1}, x_{0,2}, (\gamma_1, \gamma_2)) \\
&\quad \downarrow \sigma \in \Gamma(q) \setminus (\{ \gamma_1 \} \cup \{ \gamma_2 \}) \\
&\quad \delta(q_0, \sigma), \delta(1(x, \sigma), \sigma), \delta(2(x, \sigma), \sigma), C_1(1(x_1, \sigma)), C_2(2(x_2, \sigma))) \quad \vdots
\end{align*}
\]

Note that if \( \sigma \in \Sigma_{\sigma_0} \), then a transition of \( \delta_i(x, \sigma) \) does not result in a state change for \( A_i \), and we assume that \( C_i(\delta_i(x, \sigma)) = \{ \gamma_i \} \) from the preceding state of the arena.

Based on the previous definitions, let us now define the notion of a state-based control strategy.

**Definition 6.** A **state-based control strategy** for \( A_1 \) is given by the function

\[
\Xi_{A_1} : V_D \rightarrow \Sigma_{\sigma_0}.
\]

That is, \( \Xi_{A_1}(q, x_1, x_2, C_1(x_1), C_2(x_2)) = \gamma_1 \in C_1(x_1) \) on the same. Similarly, we define a state-based control strategy for \( A_2 \), where

\[
\Xi_{A_2} : V_D \rightarrow \Sigma_{\sigma_0}.
\]

Given \( (q, x_1, x_2, C_1(x_1), C_2(x_2)) \), the goal of a state-based control strategy for \( A_1 \) is to choose or fix, in all states \( x_1 \in X_1 \), one possible valid control among the ones of \( C_1(x_1) \) and then free the choice of \( A_2 \) amongst all its possible valid control policies at \( x_2 \in X_2 \).
Given $\Xi_{A_1}$, the runs of $\mathcal{H}$ are reduced to runs of the form
\[
\begin{cases}
  v_0 = (q_0, x_{0.1}, x_{0.2}, \Xi_{A_1}(v_0), C_2(x_{0.2})) \\
  \downarrow (\Xi_{A_1}(v_0), \gamma_2), \gamma_2 \in C_2(x_{0.2}) \\
  (q_0, x_{0.1}, x_{0.2}, \Xi_{A_1}(v_0), \{\gamma_2\}) \\
  \downarrow \sigma \in \Gamma(q) \setminus (\Xi_{A_1}(v_0) \cup \gamma_2) \\
  (\delta(q_0, \sigma), \delta_1(x_{0.1}, \sigma), \delta_2(x_{0.2}, \sigma), \\
  \Xi_{A_1}(v_1), C_2(\delta_2(x_{0.2}, \sigma)))
\end{cases}
\]
with
\[
v_1 = (\delta(q_0, \sigma), \delta_1(x_{0.1}, \sigma), \delta_2(x_{0.2}, \sigma), \\
C_1(\delta_1(x_{0.1}, \sigma)), C_2(\delta_2(x_{0.2}, \sigma))).
\]

Similarly, if $\sigma \in \Sigma_{run, 1}$ for $A_1$, then a transition of $\delta(x, \sigma)$ does not result in a state change for $A_1$, and we assume that $\Xi_{A_1}(v_0) = \Xi_{A_1}(v_1)$ from the preceding state of the arena and we denote this set of runs by $\text{Run}^{A_1}(\mathcal{H}) \subseteq \text{Run}(\mathcal{H})$.

Only part of the runs of $\text{Run}_{B_2}(\mathcal{H})$ solve Problem 2, as it should be true for any choices of $A_2$ along the trajectories that lead to $B_{ad}$ and, equivalently, for the language $L_{B_2}(\mathcal{H}) \subseteq \Sigma^* \Sigma_{o, 1}$.

Moreover, given $\Xi_{A_1}$, there might exist a run $r$ such that, once the choice of $A_1$ is fixed for every state in $V_0$, there exists a choice of $A_2$ that forms a loop in $\mathcal{H}$. If such a loop exists, then $A_2$ can go against the decisions of $A_1$ w.r.t. $\Xi_{A_1}$ so that any runs cannot lead to $B_{ad}$, whatever the decisions of $\Xi_{A_1}$. If $\Xi_{A_1}$ does not meet such a problem then it is called a loop-free strategy. We denote this set of loop-free strategies, w.r.t. $A_2$, by $L^{\text{fl}}_{B_2}(\mathcal{H})$.

Next, we define a winning control strategy of $A_1$ that leads to a state in $B_{ad}$ and is loop-free w.r.t. $A_2$.

**Definition 7.** A winning state-based control strategy is given by the function $\Xi_{A_1}^{B_2}$ of Def 6 such that, for $v_i = (q, x_{1.1}, x_{2.1}, C_1(x_{1.1}), C_2(x_{2.1}))$,
\[
(\Xi_{A_1}^{B_2}(v_i))^{*} \subseteq \{(\gamma_1^{*})^{*} \in \Sigma_{o, 1}^* | \\
\forall \gamma_2^{*} \in C_2(x_{2.1}), \exists \sigma_i \in \Gamma_H(v_i) \setminus (\gamma_1^{*} \cup \gamma_2^{*}), \text{ s.t. } \\
((\gamma_1^{*}, \gamma_2^{*}), \sigma_i)^{*} \in L^{\text{fl}}_{B_2}(\mathcal{H})\}.
\]

We denote the language that fulfills the definition of $L(\mathcal{H})$, according to the Definition 7, by $L^{\Xi_{A_1}}(\mathcal{H})$.

**Theorem 1.** Given $\Xi_{A_1}^{B_2}$ as in Definition 7, if $L^{\Xi_{A_1}}(\mathcal{H}) \subseteq \emptyset$, then there exists a winning control strategy that solves Problem 2.

Let us now explain how one can prune the arena $\mathcal{H}$ in order to see whether adversary $A_1$ can win against $A_2$. Let $B \subseteq V_0$ be an arbitrary set of states in $\mathcal{H}$.

\[
\text{Pre}_{A_1}(B) = \{ v_0 = (q, x_{1.1}, x_{2.1}, C_1(x_{1.1}), C_2(x_{2.1})) \in V_0 \setminus B | \\
\exists \gamma_1 \in C_1(x_{1.1}), \forall \gamma_2 \in C_2(x_{2.1}), \\
\exists \sigma \in (\Gamma(q) \setminus (\gamma_1 \cup \gamma_2)) \cap \Sigma_{o, 1} \\
s. t. \delta_1(x_{1.1}, \sigma), \sigma) \in B\}
\]

Then $\text{Pre}_{A_1}(B)$ is the set of “previous” states of $V_0$ from which $A_1$ has a possible control decision $\gamma_1$ such that regardless of $A_2$’s control decision $\gamma_2$, there is an enabled event $\sigma$ that leads to a state in $B$ and is observable to $A_1$.

In other words, if we consider $B = B_{ad}$, we compute the set of states of $V_0$ for which there exists a valid control policy for $A_1$ such that for any possible valid control policy of $A_2$, the system can trigger an event enabled by all the adversaries that leads to $B_{ad}$, while avoiding loops (as we only consider new states in $\text{Pre}_{A_1}(B)$). Further, the set of states of $V_0$ that lead to $B_{ad}$ is given by the following fixed-point computation:
\[
E_0 = B_{ad}; \quad E_1 = \text{Pre}_{A_1}(B_{ad}); \quad E_{k+1} = E_k \cup \text{Pre}_{A_1}(E_k), k > 1.
\]

Let $E$ be the result of this fixed-point computation.

As noted previously, $E_1$ is the set of states that allow us to reach $B_{ad}$ in one step by choosing the set of possible valid control policies $\gamma_1 \in C_1(x_1)$ w.r.t. $x_1$ so that whatever the choice of $A_2$, the system $G$ can reach $B_{ad}$ by an observable event. Meanwhile, starting from any state of $E$, it is easy to see that there exists a sequence that reaches $B_{ad}$.

**Theorem 2.** Given the following arena
\[
\mathcal{H} = ((V_0 \cup V_5), E_0 \cup \Sigma, \delta_D, \delta_{SD}, v_0, B_{ad}, B_{ad}).
\]

Then consider
\[
E_0 = (\text{pre} E \cup V_5), E_0 \cup \Sigma, \delta_D, \delta_{SD}, v_0, B_{ad}, B_{ad}).
\]
If $E_0$ is not reduced to the empty automaton, then there exists a winning control strategy that solves Problem 2.

**Example 3.** We reuse $G$ and its observers $O_1$ and $O_2$ from Ex. 2, but examine it within the setting of Problem 2. As a reminder, $\Sigma_{o, 1} = \Sigma_{o, 1} = \{a, c\}$ and $\Sigma_{o, 2} = \Sigma_{o, 2} = \{b, c\}$. A portion of the arena $\mathcal{H}$ is shown in Fig 7. As noted before, the system is not opaque for either adversary. Even without displaying the entire arena, there are, in fact, winning strategies for $A_1$. For example, at the initial state, if $A_1$ chooses $\gamma_1 = \{c\}$, then $G$ is forced to execute either the sequence $aa$ or $baa$, leading to a winning configuration for $A_1$.

6. CONCLUSIONS

We presented two problems involving the use of control to prevent information leakage between two adversaries. In the first problem, we used an existing result from Dubreil et al. (2010) to determine whether or not there is a global control policy, where the agents are passive, to preserve the secret of each adversary. For the second problem, where the system is passive and each agent exerts local control to discover the secret of its adversary, we presented a sufficient condition for the existence of a winning strategy for $A_1$ to always discover $A_2$’s secret. Future extensions of this work include extending the results to multiple secrets and incorporating a quantitative component to the secret to present different incentives for the interactions of the adversaries.
Fig. 7. A portion of the arena $\mathcal{H}$ for the dynamic concurrent secrets version of Ex. 1. States in $V_D$ are denoted as ellipses, while states in $V_S$ are denoted as rectangles. The states shown in grey are in $Bad_1 \cup Bad_2$.

REFERENCES


