In Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)
Auckland, New Zealand (Virtual)

Signaling Friends and Head-Faking Enemies Simultaneously: Balancing Goal Obfuscation and Goal Legibility*

Extended Abstract

Anagha Kulkarni
Arizona State University
anaghak@asu.edu

Siddharth Srivastava
Arizona State University
siddharths@asu.edu

Subbarao Kambhampati
Arizona State University
rao@asu.edu

ABSTRACT

In order to be useful in the real world, an AI agent needs to plan and act in the presence of other agents, who may be helpful or disruptive. In this paper, we consider the problem where an autonomous agent needs to act in a manner that clarifies its objectives to cooperative agents while simultaneously preventing adversarial agents from inferring those objectives. We call it *Mixed-Observer Controlled Observability Planning Problem* (MO-COPP). We develop two new solution approaches: one provides an optimal solution to the problem given a fixed time horizon by using an integer programming solver, the other provides a satisficing solution using heuristic-guided forward search to achieve prespecified amount of obfuscation and legibility for adversarial and cooperative agents respectively.

KEYWORDS

obfuscation, legibility, integer program, planning

ACM Reference Format:
Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. 2020. Signaling Friends and Head-Faking Enemies Simultaneously: Balancing Goal Obfuscation and Goal Legibility. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020*, IFAAMAS, 3 pages.

1 MO-COPP

In a multi-agent environment, the activities performed by an agent may be observed by other agents. In such an environment, an agent should perform its tasks while taking into account the observers' sensing capabilities and its relationship with the observers. Several prior works have explored the generation of legible behavior to convey necessary information to a cooperative observer [1, 5, 8] and obfuscating behavior to hide sensitive information from an adversarial observer [3, 4, 6, 7]. However, in real-world settings of strategic importance, an agent might encounter both types of observers *simultaneously*. This would necessitate synthesizing a behavior that is simultaneously legible to friendly entities and obfuscatory to adversarial ones. For instance, in soccer, a player may perform a feinting trick to confuse an opponent while signaling a teammate. This problem gives rise to a novel optimization space that involves trading-off the amount of obfuscation desired for adversaries with the amount of legibility desired for friends.

A MO-COPP setting involves an actor (**A**) and two observers, where one is adversarial observer (**X**) while the other is cooperative (**C**). The actor has full observability of its own activities and knows the sensor models used by the observers. The observers have different sensor models. When the actor takes an action and reaches a new state, an observation is emitted. After obtaining the observations, the observers update their belief. The actor leverages the known limits in the observers' sensors to control the observability of multiple observers in the environment simultaneously. Given a set of candidate goals, the objective of the actor is to convey information about its goal to **C** and to hide it from **X**.

Formally, a **mixed-observer controlled observability planning problem** is a tuple, MO-COPP $= \langle \Lambda, \mathcal{P}, \mathcal{G}, \{\Omega_i\}_{i\in\Lambda}, \{O_i\}_{i\in\Lambda}, \{\mathcal{B}_0^i\}_{i\in\{\mathbf{X},\mathbf{C}\}}\rangle$. $\Lambda = \{\mathbf{A}, \mathbf{C}, \mathbf{X}\}$ is the set of agents. $\mathcal{P} = \langle \mathcal{F}, Op, \mathcal{I}, G_A\rangle$ is **A**'s task captured as a planning problem [2], where $\mathcal{F}$ is the set of fluents, Op is the set of actions, $\mathcal{I}$ is the initial state and goal G_A is a subset of fluents. Also, for $a \in Op$, $pre(a)$, $add(a)$, $delete(a)$ are each a subset of fluents representing preconditions, add effects and delete effects of a. $\mathcal{G} = \{G_1, G_2, \ldots, G_{n-1}, G_A\}$ is the set of candidate goals, where G_A is the true goal of **A**, which is not known to either **C** or **X**. Ω_i is the set of observation symbols for agent i, which are emitted when **A** takes an action and reaches a new state. Further, $\Omega_A = \{o^{\mathbf{A}}_{a,s} | a \in Op, s \in \mathcal{S}\}$. $O_i : Op \times \mathcal{S} \rightarrow \Omega_i$ is agent i's deterministic sensor model. $\mathcal{S}$ is the set of states, where each state is an instantiation of all fluents. Further, $O_{\mathbf{A}}$ maps each action-state pair to a unique observation, giving **A** full observability. While, $O_{\mathbf{X}}$ and $O_{\mathbf{C}}$ are noisy sensor models that map multiple action-state pairs to the same observation symbol, giving observers partial observability. $\mathcal{B}_0^i$ is the initial belief of an observer, $i \in \{\mathbf{X}, \mathbf{C}\}$. The initial belief is a set of states inclusive of $\mathcal{I}$.

Although the observers are aware of the planning domain of the actor and of the candidate goals, they do not know which candidate goal is the actor's true goal, $G_{\mathbf{A}}$. The observers' partial observability is due to mapping of multiple $\langle a, s\rangle$ pairs to an observation, i.e., $\forall i \in \{\mathbf{X}, \mathbf{C}\}, \exists a, a' \in Op, s, s' \in \mathcal{S}, a \neq a' \wedge s \neq s' : O_i(a, s) = O_i(a', s')$. Each observer $i \in \{\mathbf{X}, \mathbf{C}\}$ maintains its belief, which is a set of states. $\Gamma(\cdot)$ is a transition function, such that, $\Gamma(s, a) = \perp$ if $s \not\models pre(a)$; else $\Gamma(s, a) = s \cup add(a) \setminus delete(a)$. Now we can define the belief update: (1) at time step $t = 0$, the initial belief of i is given by $\mathcal{B}_0^i$, (2) at time step $t \in \{1, \ldots, \mathcal{T}\}$, $\mathcal{B}_t^i = \{\hat{s} \mid \exists \hat{a}, \bar{s}\ \Gamma(\bar{s}, \hat{a}) = \hat{s}; \bar{s} \in \mathcal{B}_{t-1}^i; O_i(\hat{a}, \hat{s}) = o_t^i\}$, where $\mathcal{T}$ is the time horizon. That is, the belief is updated using the previous belief and the observation received. A plan π (or sequence of actions) is associated with a sequence of observations, $ObsSeq_i(\pi) = \{o_1^i, \ldots, o_{\mathcal{T}}^i\}$ for each observer i.

A plan π is a valid solution to a MO-COPP $= \langle \Lambda, \mathcal{P}, \mathcal{G}, \{\Omega_i\}_{i\in\Lambda}, \{O_i\}_{i\in\Lambda}, \{\mathcal{B}_0^i\}_{i\in\{\mathbf{X},\mathbf{C}\}}\rangle$, iff: $\Gamma(\mathcal{I}, \pi) \models G_A$. In other words, any

*The full version of this paper is available at https://arxiv.org/abs/1905.10672.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

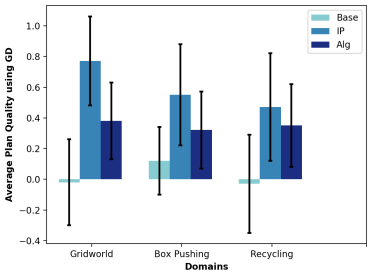

Figure 1: Average plan quality (GD) using a baseline planner, IP planner and search algorithm over three domains.

solution to $\mathcal{P}$ is a solution to MO-COPP. We measure the quality of a valid MO-COPP solution in terms of the extent to which the plan is legible to **C** and obfuscating to **X**. Given a solution plan, π, $\mathcal{G}^{\pi}_{\mathbf{X}} = \{G \in \mathcal{G} : \exists s \in \mathcal{B}^{\mathbf{X}}_{\mathcal{T}},\ s \models G\}$, denotes the goals present in **X**'s final belief and $\mathcal{G}^{\pi}_{\mathbf{C}} = \{G \in \mathcal{G} : \exists s \in \mathcal{B}^{\mathbf{C}}_{\mathcal{T}},\ s \models G\}$, denotes the goals present in **C**'s final belief. $|\mathcal{G}^{\pi}_{\mathbf{X}}|$ and $|\mathcal{G}^{\pi}_{\mathbf{C}}|$ represent the amount of **goal obfuscation for X** and the amount of **goal legibility for C**. By increasing (or decreasing) the number of goals in $\mathcal{G}^{\pi}_{\mathbf{X}}$ (or $\mathcal{G}^{\pi}_{\mathbf{C}}$), we can improve the goal obfuscation (or goal legibility).

Given a solution plan π that solves MO-COPP, where $|\mathcal{G}| = n$, the **goal difference**, of π is given by: $GD(\pi) = \frac{|\mathcal{G}^{\pi}_{\mathbf{X}}| - |\mathcal{G}^{\pi}_{\mathbf{C}}|}{n-1}$ where the denominator represents the difference between ideal values of $\mathcal{G}^{\pi}_{\mathbf{X}}$ and $\mathcal{G}^{\pi}_{\mathbf{C}}$. An optimal solution to MO-COPP maximizes the trade-off between amount of goal obfuscation and goal legibility. That is, it maximizes the difference between the number of goals in $\mathcal{G}^{\pi}_{\mathbf{X}}$ and $\mathcal{G}^{\pi}_{\mathbf{C}}$. Equivalently, closer the $GD(\pi)$ value to 1, better is the plan quality. A solution plan with $GD(\pi) = 1$ is an optimal plan.

2 PLAN SYNTHESIS

We now present two solution synthesis approaches: **(1)** we formulate MO-COPP as a constraint optimization problem and provide an IP encoding to solve it in $\mathcal{T}$ steps, **(2)** we use heuristic-guided search to achieve preset levels of goal obfuscation and legibility simultaneously.

MO-COPP as Integer Program The IP encoding provides an optimal solution for the given horizon by maximizing the trade-off between the amount of obfuscation and legibility. Let $x_{a,t}, y_{s,t}, w_{o,t}$ be indicator variables for action a, state s and observation o at time t respectively, $b^i_{s,t}, h^i_{s,a,t}$ for state s and action a being applicable in state s in observer i's belief at time t respectively, $g^i_{G,\mathcal{T}}$ for a goal G present in observer i's final belief. The objective function is essentially the numerator of $GD(\cdot)$ metric, i.e.,

$max\ \sum_{G\in\mathcal{G}} g^{\mathbf{X}}_{G,\mathcal{T}} - \sum_{G\in\mathcal{G}} g^{\mathbf{C}}_{G,\mathcal{T}}$.

The IP constraints are as follows:

(1) $\forall s \in \mathcal{S}, s = \mathcal{I} :\ y_{s,0} = 1;\ s \neq \mathcal{I} :\ y_{s,0} = 0;\ \sum_{G_A \in s} y_{s,\mathcal{T}} = 1$

(2) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s \in \mathcal{S}, s \in \mathcal{B}^i_0 :\ b^i_{s,0} = 1;\ s \notin \mathcal{B}^i_0 :\ b^i_{s,0} = 0$

(3) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, G \in \mathcal{G}, m > |\{s|\ G \in s\}| : m * g^i_{G,\mathcal{T}} - \sum_{G\in s} b^i_{s,\mathcal{T}} \geqslant 0$

(4) $\forall a \in Op, t \in \{1, \ldots, \mathcal{T}\}, pre_a = \{s|pre(a) \in s\} : x_{a,t} \leqslant \sum_{s\in pre_a} y_{s,t-1}$

(5) $\forall s, s' \in \mathcal{S}, t \in \{1, \ldots, \mathcal{T}\}, add_{s'} = \{a|pre(a) \in s, add(a) \setminus delete(a) \in s'\}, pre_{s'} = \{s|\ pre(a) \in s \wedge add(a) \setminus delete(a) \in s'\} :$
$\sum_{a\in add_{s'}} x_{a,t} + \sum_{s\in pre_{s'}} y_{s,t-1} - 2\ y_{s',t} \geqslant 0$

(6) $\forall a \in Op, t \in \{1, \ldots, \mathcal{T}\}, post_a = \{s'\ |\ add(a) \setminus delete(a) \in s'\} :$
$\sum_{s\in pre_a, s'\in post_a} y_{s,t-1}\ y_{s',t} = x_{a,t}$

(7) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, o \in \Omega_i, t \in \{1, \ldots, \mathcal{T}\} :\ w^i_{o,t} = \sum_{a,s'\in O^i_o} x_{a,t}\ y_{s',t}$

(8) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s \in \mathcal{S}, t \in \{1, \ldots, \mathcal{T}\},\ a \in add_s,$
$add_s = \{a|\ pre(a) \in s\} :\ b^i_{s,t-1} + w^i_{o,t} - h^i_{s,a,t} \leqslant 1$

(9) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s \in \mathcal{S}, o \in \Omega_i, t \in \{1, \ldots, \mathcal{T}\}, a \in add_s,$
$add_s = \{a|\ pre(a) \in s\} : h^i_{s,a,t} - b^i_{s,t-1} \leqslant 0$

(10) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s \in \mathcal{S}, t \in \{1, \ldots, \mathcal{T}\}, a \in add_s, s' \in post_s$
$add_s = \{a|pre(a) \in s\}, post_s = \{s'|add(a) \setminus delete(a) \in s'\} : h^i_{s,a,t} - b^i_{s',t} \leqslant 0$

(11) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s \in \mathcal{S}, o \in \Omega_i, t \in \{1, \ldots, \mathcal{T}\}, a \in add_s,$
$add_s = \{a|\ pre(a) \in s\} :\ h^i_{s,a,t} - w^i_{o,t} \leqslant 0$

(12) $\forall i \in \{\mathbf{X}, \mathbf{C}\}, s, s' \in \mathcal{S}, t \in \{1, \ldots, \mathcal{T}\}, add_{s'} = \{a|\ pre(a) \in s \wedge add(a) \setminus delete(a) \in s'\}, pre_{s'} = \{s|\ pre(a) \in s \wedge add(a) \setminus delete(a) \in s'\} :$
$\sum_{s\in pre_{s'}, a\in add_{s'}} h^i_{s,a,t} - b^i_{s',t} \geqslant 0$

(13) $\forall t \in \{1, \ldots, \mathcal{T}\} : \sum_{a\in Op} x_{a,t} \leqslant 1$

Constraint (1) sets the initial state and ensures that a state that satisfies the true goal is achieved in the last time step for **A**. Constraint (2)-(3) set the initial belief and the goal satisfaction constraint for the final belief for both the observers. Constraint (4)-(6) enforce the transition function for the actor. Constraint (7) sets the observation symbols for each observer depending on the $\langle a, s'\rangle$ pair in the plan. Constraints (8)-(12) enforce a belief update for the observers. Constraint (13) ensures only one action is possible at each step.

Search Algorithm We adapt Kulkarni et al. [4] approach to address goal obfuscation and goal legibility simultaneously. We specify bounds on the amount of goal obfuscation and goal legibility desired: obfuscate true goal with at least $k-1$ goals, make true goal legible with at most j goals. These bounds, $\Phi = \langle \Phi_{\mathbf{X}}, \Phi_{\mathbf{C}} \rangle$, are given as input to the search algorithm. Each search node maintains the associated beliefs for both observers: b^i_{Δ}, of size Δ (i.e. cardinality of b^i_{Δ} is Δ). b^i_{Δ} is always inclusive of the true state of the actor, this is because the actor can only take actions that are consistent with its true state. For each Δ, all Δ-sized unique combinations of belief (that include the actual state of the actor) are explored. This allows systematic and complete exploration of multiple paths to a given search node. The algorithm consists of two loops. The outer loop incrementally increases the cardinality of the belief i.e. Δ from $1, \ldots, |\mathcal{S}|$. The inner iteration performs heuristic guided forward search (we use greedy best first search) to find a plan while tracking at most Δ states in each b^i_{Δ}. In the inner loop, the node expansion is guided by (1) customized heuristic function: $h_{node} = h_{G_A}(s) + h_{\mathcal{G}_{k-1}}(\mathcal{B}^{\mathbf{X}}) - h_{\mathcal{G}_{\mathcal{G}-j}}(\mathcal{B}^{\mathbf{C}})$, and (2) goal test, which checks for satisfaction of true goal and constraints given by Φ. The algorithm stops either when a solution is found or when all the Δ iterations have been explored.

Empirical Evaluation The average and standard deviation *GD* for 3 domains are reported in Figure 1. The IP has higher *GD* for all 3 domains. While, the search algorithm generates satisficing solutions with lower *GD* that meet the goal constraints. The baseline planner that achieves satisficing solution to a single goal produces worst quality solutions (lowest *GD*).

Acknowledgements. This work is supported in part by ONR grants N00014-16-1-2892, N00014-18-1-2442, N00014-18-1-2840, N00014-19-1-2119, AFOSR grant FA9550-18-1-0067, DARPA SAILON grant W911NF-19-2-0006, NSF grants 1936997 (C-ACCEL), 1844325, NASA grant NNX17AD06G, and a JP Morgan AI Faculty Research grant.

REFERENCES

[1] Anca Dragan and Siddhartha Srinivasa. 2013. Generating Legible Motion. In *Proceedings of Robotics: Science and Systems*. Berlin, Germany.

[2] Hector Geffner and Blai Bonet. 2013. A concise introduction to models and methods for automated planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 8, 1 (2013), 1–141.

[3] Sarah Keren, Avigdor Gal, and Erez Karpas. 2016. Privacy Preserving Plans in Partially Observable Environments.. In *IJCAI*. 3170–3176.

[4] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. 2019. A unified framework for planning in adversarial and cooperative environments. In *AAAI*.

[5] Aleck M MacNally, Nir Lipovetzky, Miquel Ramirez, and Adrian R Pearce. 2018. Action Selection for Transparent Planning. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1327–1335.

[6] Peta Masters and Sebastian Sardina. 2017. Deceptive Path-Planning. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 4368–4375. https://doi.org/10.24963/ijcai.2017/610

[7] Shashank Shekhar and Ronen I Brafman. 2018. Representing and Planning with Interacting Actions and Privacy. In *Twenty-Eighth International Conference on Automated Planning and Scheduling*.

[8] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz H Zhuo, and Subbarao Kambhampati. 2017. Plan Explicability and Predictability for Robot Task Planning. In *International Conference on Robotics and Automation (ICRA)*.