Deceptive Decision-Making Against Adversaries: *Theory, Algorithms, and User Studies*

Yagiz Savas

yagiz.savas@utexas.edu

In collaboration with Christos Verginis, Michael Hibbard, Emilie Thome, and Ufuk Topcu

autonomous SYSTEMS GROUP

TEXAS
The University of Texas at Austin

Deception is a critical capability that helps ...

animals to survive.

teams to win games.

armies to win battles.

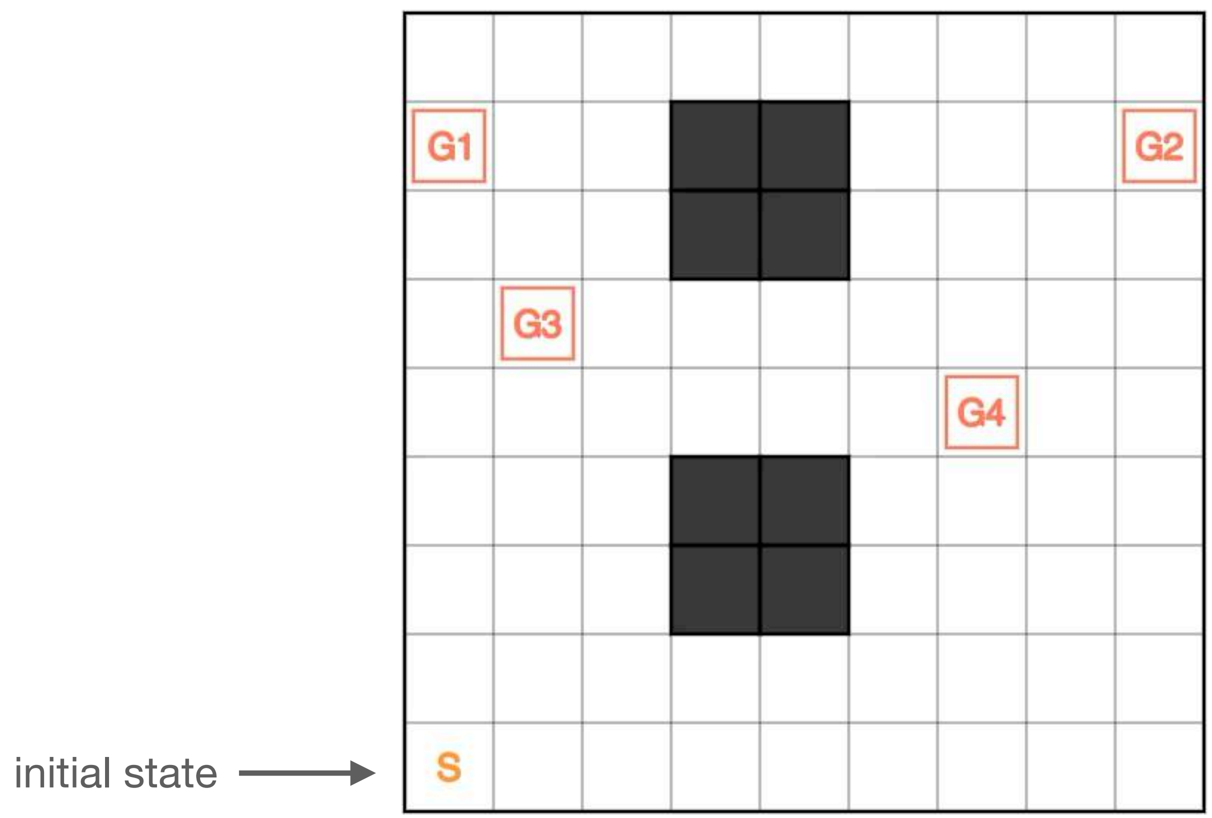
Deceptive capabilities in autonomy will lead to enhanced security …

in surveillance missions.

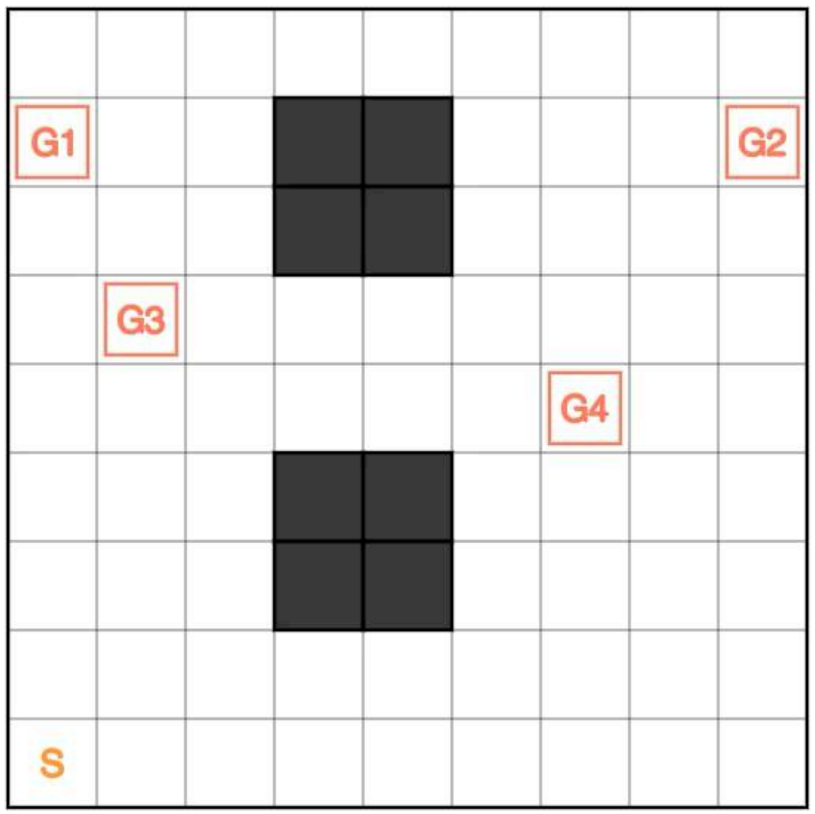
in aerial battles.

in cyber space.

Is autonomous deception really possible?

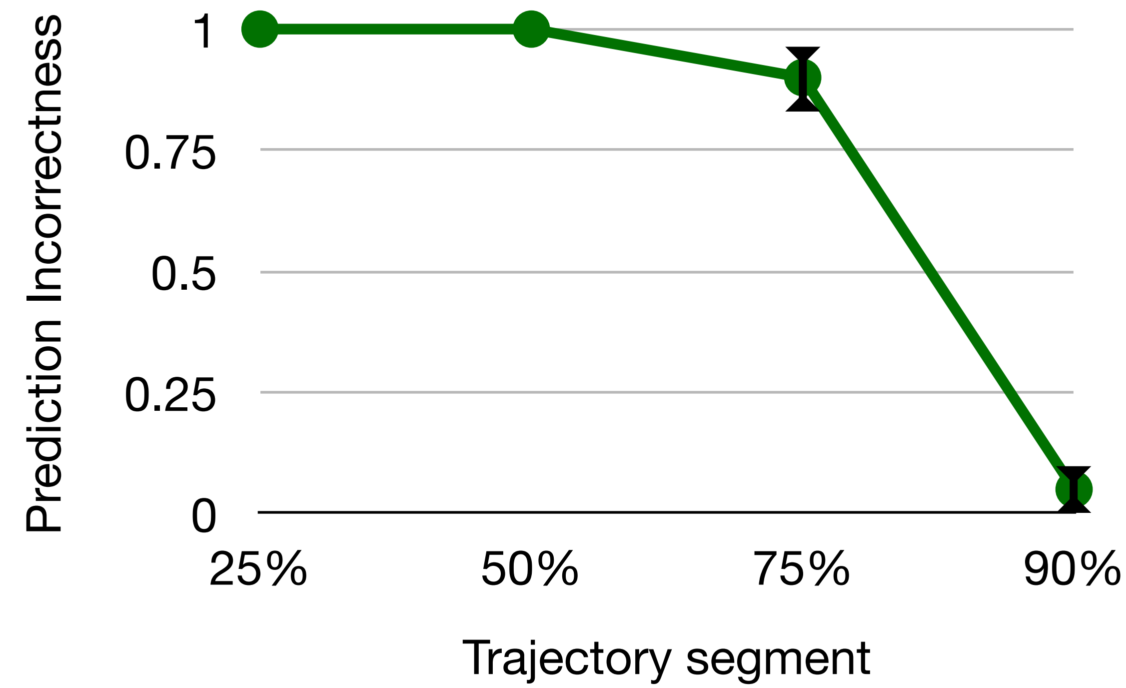


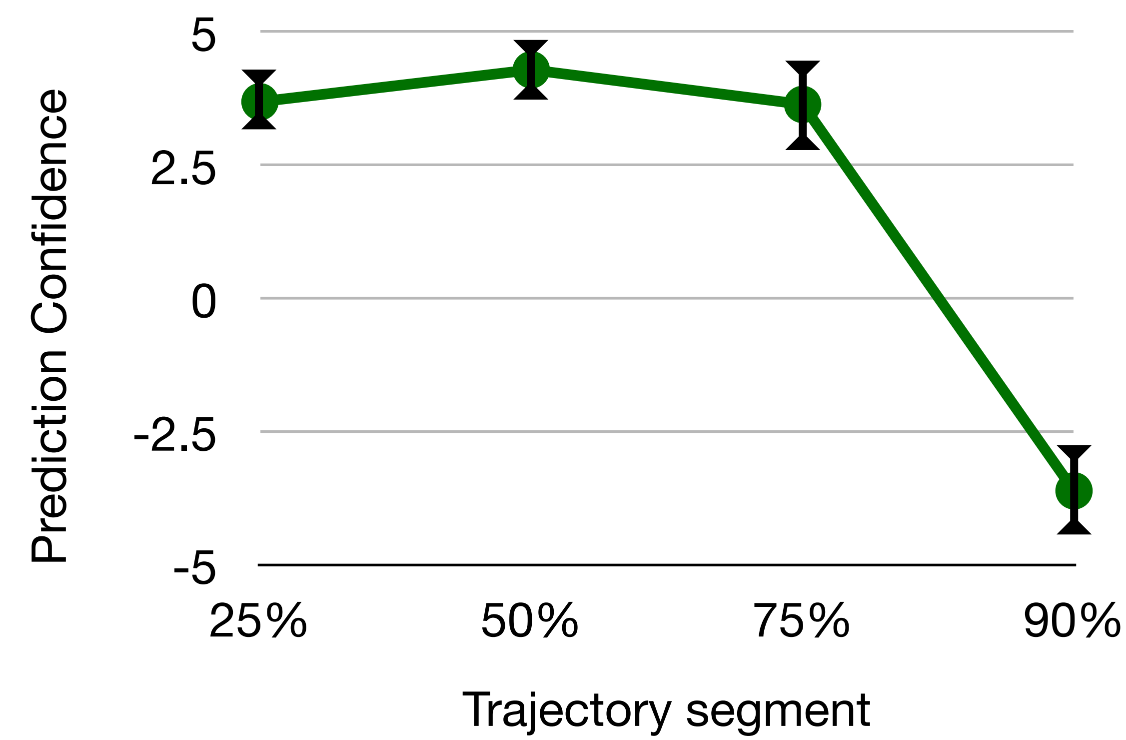
Is autonomous deception really possible?



Yes, it is possible!
but still in the early stages of development…

* Based on the user study in [1]

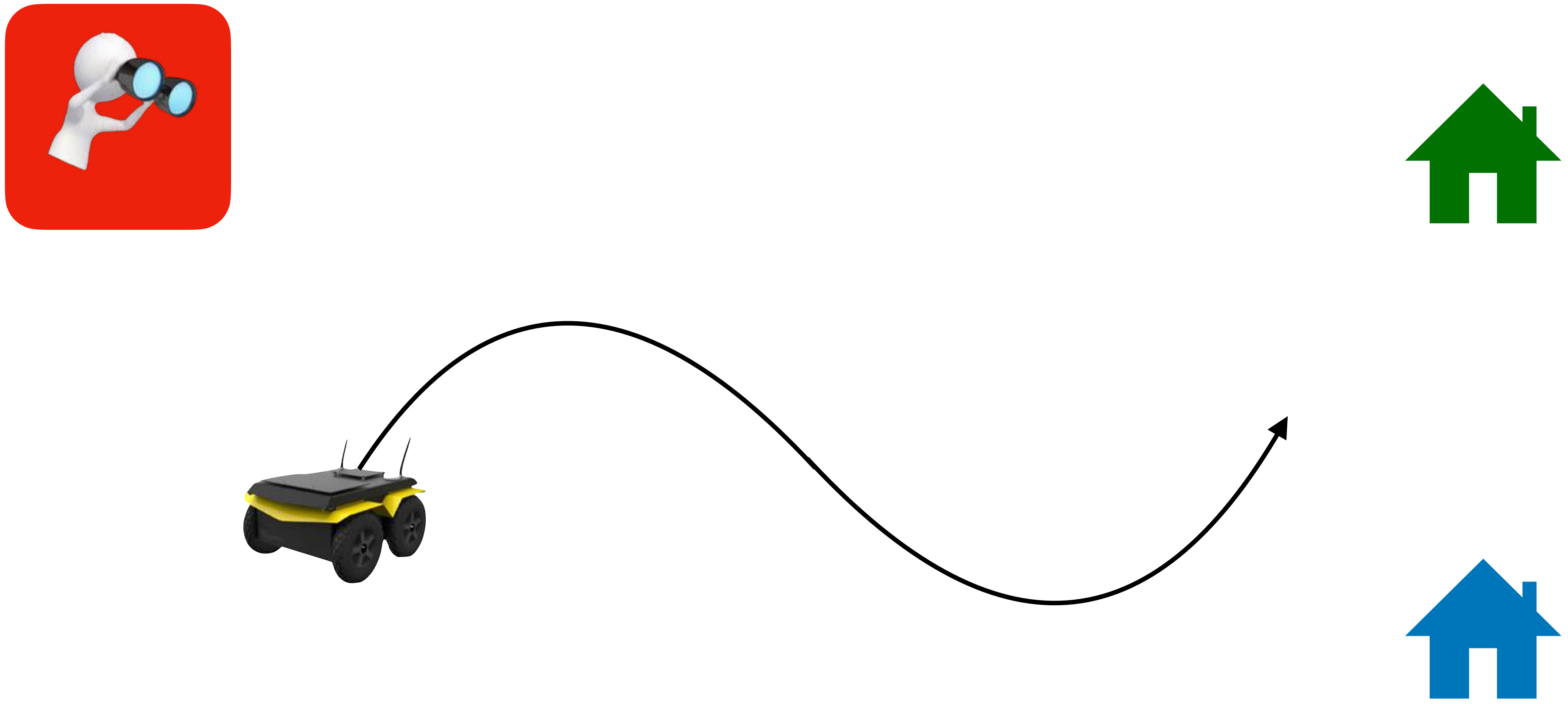




[1] Y. Savas, C. Verginis, U. Topcu, ``Deceptive decision-making under uncertainty”, AAAI Conference on Artificial Intelligence, 2021 (under review)

Overview

- Observer's prediction model
- Deception as a constrained optimization problem
- Technical considerations
- User studies and a case study in Manhattan, New York



Related work and contributions

- Game-theoretic approaches with demanding computational requirements [1,2]
- Heuristic approaches tailored to specific scenarios [3,4]
- Gradient descent-based approaches that have only local optimality guarantees [5]

Contribution: An efficient deception algorithm that works in stochastic environments, adjusts behavior according to predictions, and has global performance guarantees [6].

[1] R. Wagner, and R. Arkin, ``Acting deceptively: Providing robots with the capacity for deception”, International Journal of Social Robotics, 3(1):5–26, 2011.

[2] A. Anwar, and C. Kamhoua, ``Game theory on attack graph for cyber deception”, International Conference on Decision and Game Theory for Security, 445–456, 2020.

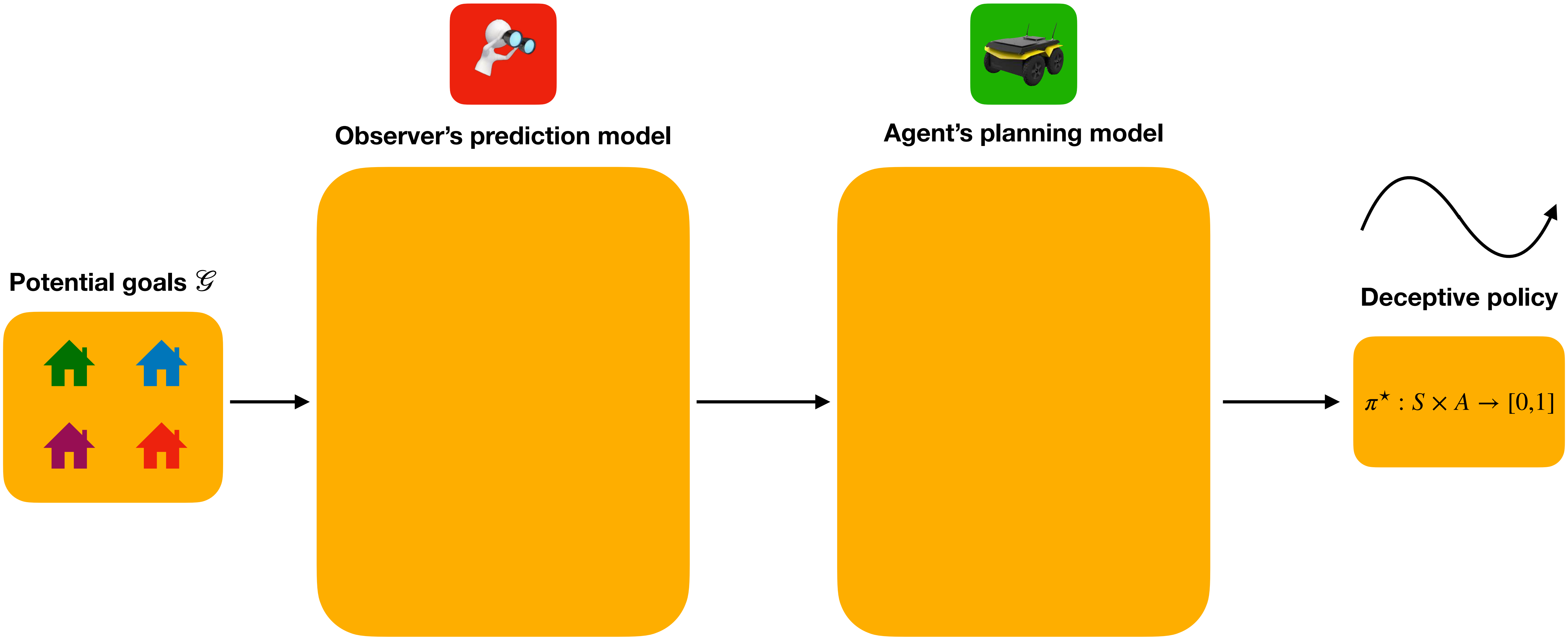
[3] P. Masters, and S. Sardina, ``Deceptive path-planning”, International Joint Conference on Artificial Intelligence, 2017.

[4] M. Pettinati, and R. Arkin, ``Push and pull: Shepherding multi-agent robot teams in adversarial situations”, International Conference on Advanced Robotics and its Social Impacts, 2019.

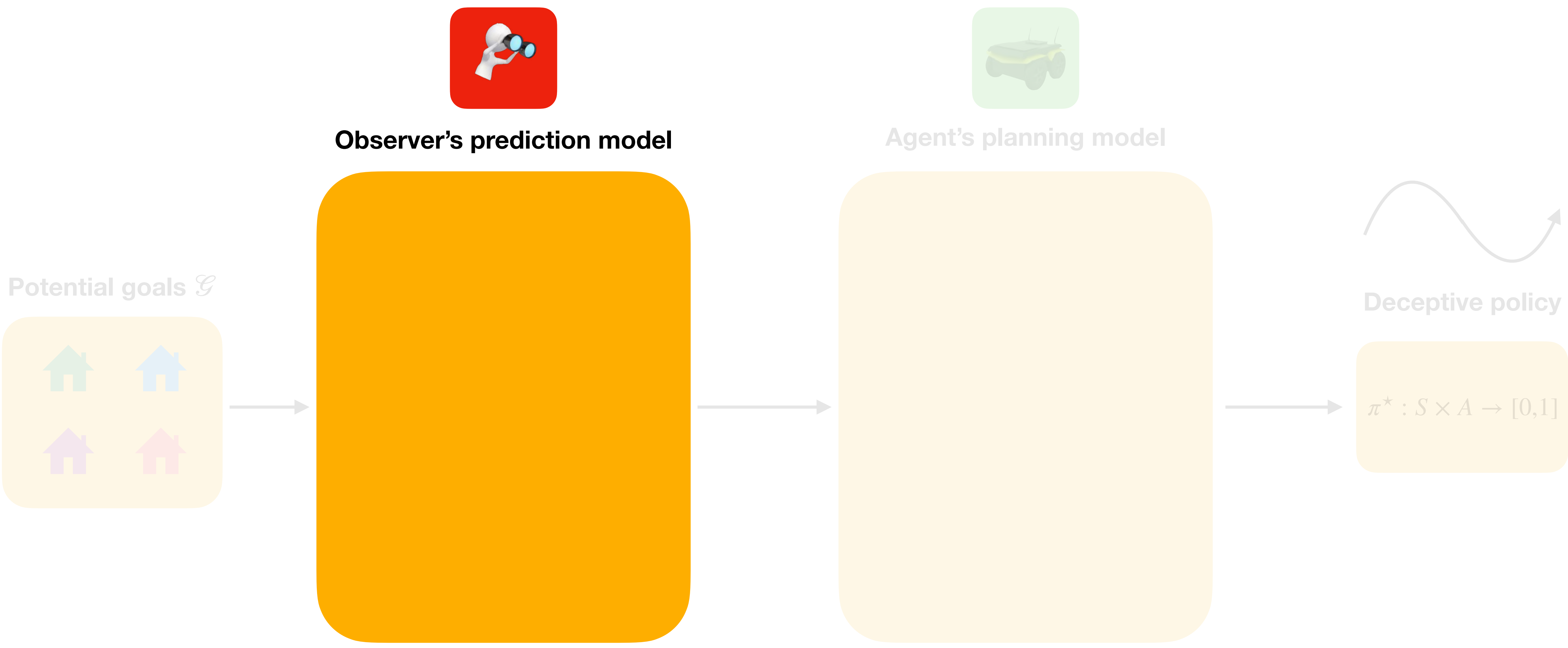
[5] A. Dragan, A, R. Holladay, and S. Srinivasa, ``Deceptive robot motion: synthesis, analysis and experiments”, Autonomous Robots, 39(3):331–345, 2015.

[6] Y. Savas, C. Verginis, U. Topcu, ``Deceptive decision-making under uncertainty”, AAAI Conference on Artificial Intelligence, 2021 (under review)

Overall system model

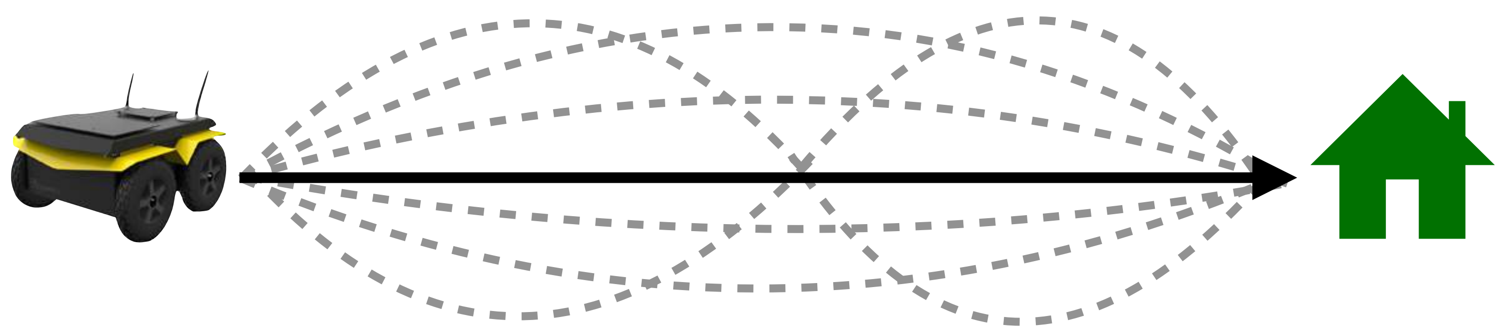


Overall system model



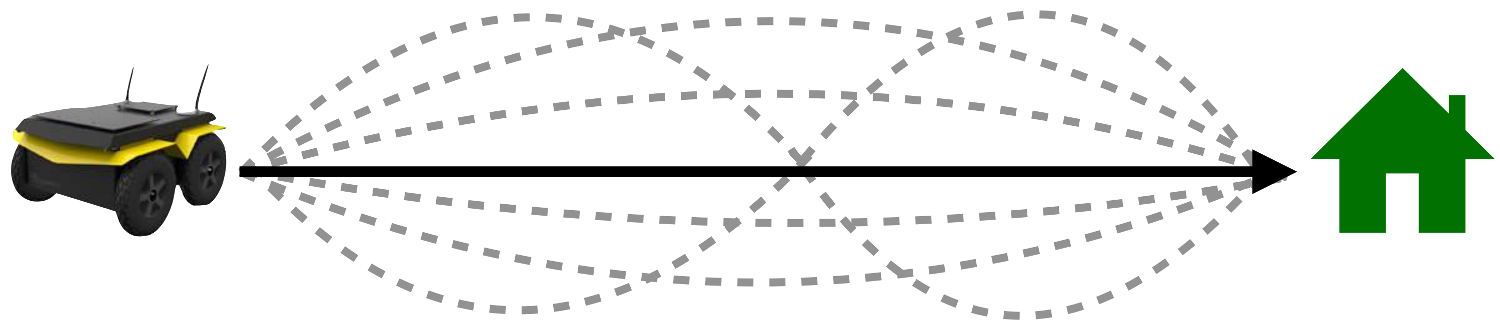
Observer's prediction model: the principle of maximum entropy

Observers expect the agent behavior to be goal-directed with a certain degree of efficiency.



Observer's prediction model: the principle of maximum entropy

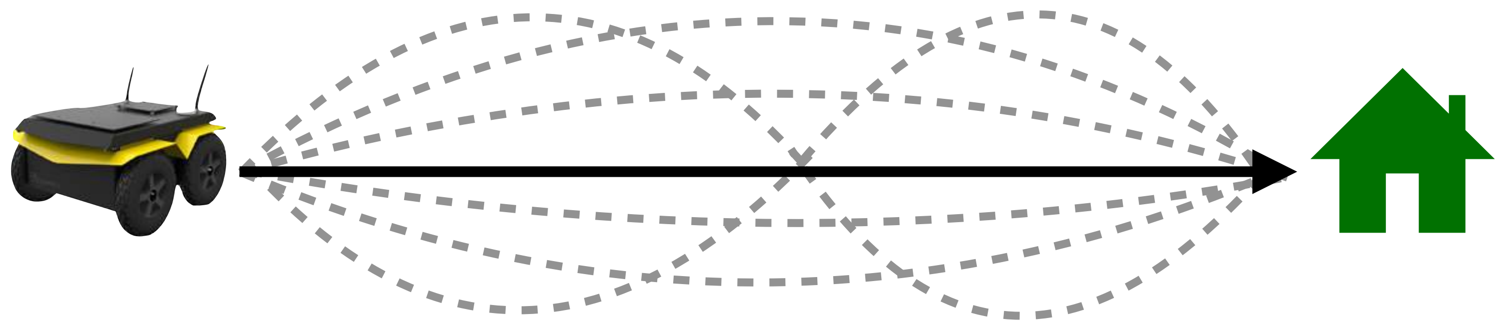
Observers expect the agent behavior to be goal-directed with a certain degree of efficiency.



The principle of maximum entropy: the distribution that best represent the current state of knowledge is maximally noncommittal with regard to missing information

Observer's prediction model: the principle of maximum entropy

Observers expect the agent behavior to be goal-directed with a certain degree of efficiency.



The principle of maximum entropy: the distribution that best represent the current state of knowledge is maximally noncommittal with regard to missing information

We can formally express the agent's expected goal-directed behavior $\overline{\pi}_G$ as

cost function
$c : S \times A \rightarrow [0,1]$

efficiency parameter α

entropy regularization

$$\overline{\pi}_G \in \arg\min_{\pi \in \Pi} \ \mathbb{E}^{\pi}\Big[\sum_{t=0}^{\infty} \gamma_o^t \Big(c(s_t, a_t) - \alpha H(\pi(s_t, \cdot\,)) \Big)\Big]$$

subject to: $\Pr^{\pi}(Reach[G]) = R_{\max}(G)$

reach the goal G
with maximum probability

Observer's prediction model: computing predictions

The observer knows that the agent is moving towards one of N potential goals $\mathcal{G} = \{G_1, G_2, \ldots, G_N\}$.

Given a partial agent trajectory $\zeta_{1:T}$, the observer aims to predict the agent's true goal $G^\star \in \mathcal{G}$.

Observer's prediction model: computing predictions

The observer knows that the agent is moving towards one of N potential goals $\mathcal{G} = \{G_1, G_2, \ldots, G_N\}$.

Given a partial agent trajectory $\zeta_{1:T}$, the observer aims to predict the agent's true goal $G^\star \in \mathcal{G}$.

Bayes' rule

$$\Pr(G \mid \zeta_{1:T}) = \frac{\Pr(\zeta_{1:T} \mid G)\Pr(G)}{\sum_{G' \in \mathcal{G}} \Pr(\zeta_{1:T} \mid G')\Pr(G')}$$

Prior beliefs on potential goals

How is the agent expected to reach the goal G'?

Observer's prediction model: computing predictions

The observer knows that the agent is moving towards one of N potential goals $\mathcal{G} = \{G_1, G_2, \ldots, G_N\}$.

Given a partial agent trajectory $\zeta_{1:T}$, the observer aims to predict the agent's true goal $G^\star \in \mathcal{G}$.

Bayes' rule

$$\Pr(G \mid \zeta_{1:T}) = \frac{\Pr(\zeta_{1:T} \mid G)\Pr(G)}{\sum_{G' \in \mathcal{G}} \Pr(\zeta_{1:T} \mid G')\Pr(G')}$$

Prior beliefs on potential goals

How is the agent expected to reach the goal G'?

Compute the conditional probabilities $\Pr(\zeta_{1:T} \mid G)$ using the expected goal-directed behavior $\overline{\pi}_G$.

$$\overline{\pi}_G(s,a) = e^{\alpha\left(Q_G(s,a) - V_G(s)\right)}$$

$$\left.\begin{aligned} Q_G(s,a) &= -c(s,a) + \gamma_o \sum_{s' \in \mathcal{S}} P(s,a,s')V_G(s') \\ V_G(s) &= \operatorname*{softmax}_a Q_G(s,a)\,. \end{aligned}\right\}$$

softmax value iteration efficiently computable

Observer's prediction model: computing predictions

The observer knows that the agent is moving towards one of N potential goals $\mathcal{G} = \{G_1, G_2, \ldots, G_N\}$.

Given a partial agent trajectory $\zeta_{1:T}$, the observer aims to predict the agent's true goal $G^\star \in \mathcal{G}$.

Bayes' rule

$$\Pr(G \mid \zeta_{1:T}) = \frac{\Pr(\zeta_{1:T} \mid G)\Pr(G)}{\sum_{G' \in \mathcal{G}} \Pr(\zeta_{1:T} \mid G')\Pr(G')}$$

Prior beliefs on potential goals

How is the agent expected to reach the goal G'?

Compute the conditional probabilities $\Pr(\zeta_{1:T} \mid G)$ using the expected goal-directed behavior $\overline{\pi}_G$.

$$\overline{\pi}_G(s,a) = e^{\alpha\left(Q_G(s,a) - V_G(s)\right)}$$

$$\left.\begin{aligned} Q_G(s,a) &= -c(s,a) + \gamma_o \sum_{s' \in \mathcal{S}} P(s,a,s')V_G(s') \\ V_G(s) &= \operatorname*{softmax}_a Q_G(s,a)\,. \end{aligned}\right\}$$

softmax value iteration
efficiently computable

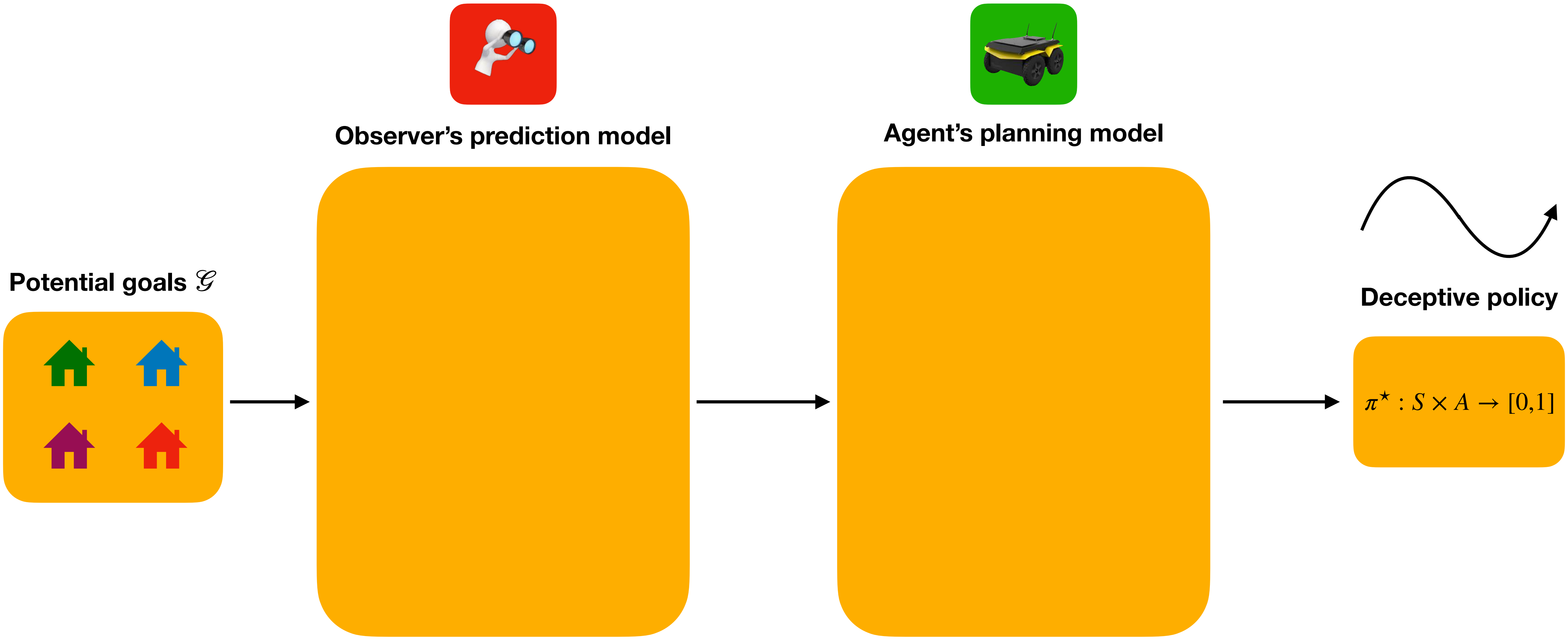
As a result, we have

$$\Pr(G \mid \zeta_{1:T}) \approx \frac{e^{V_G(s_T) - V_G(s_1)}\Pr(G)}{\sum_{G' \in \mathcal{G}} e^{V_{G'}(s_T) - V_{G'}(s_1)}\Pr(G')}\,.$$

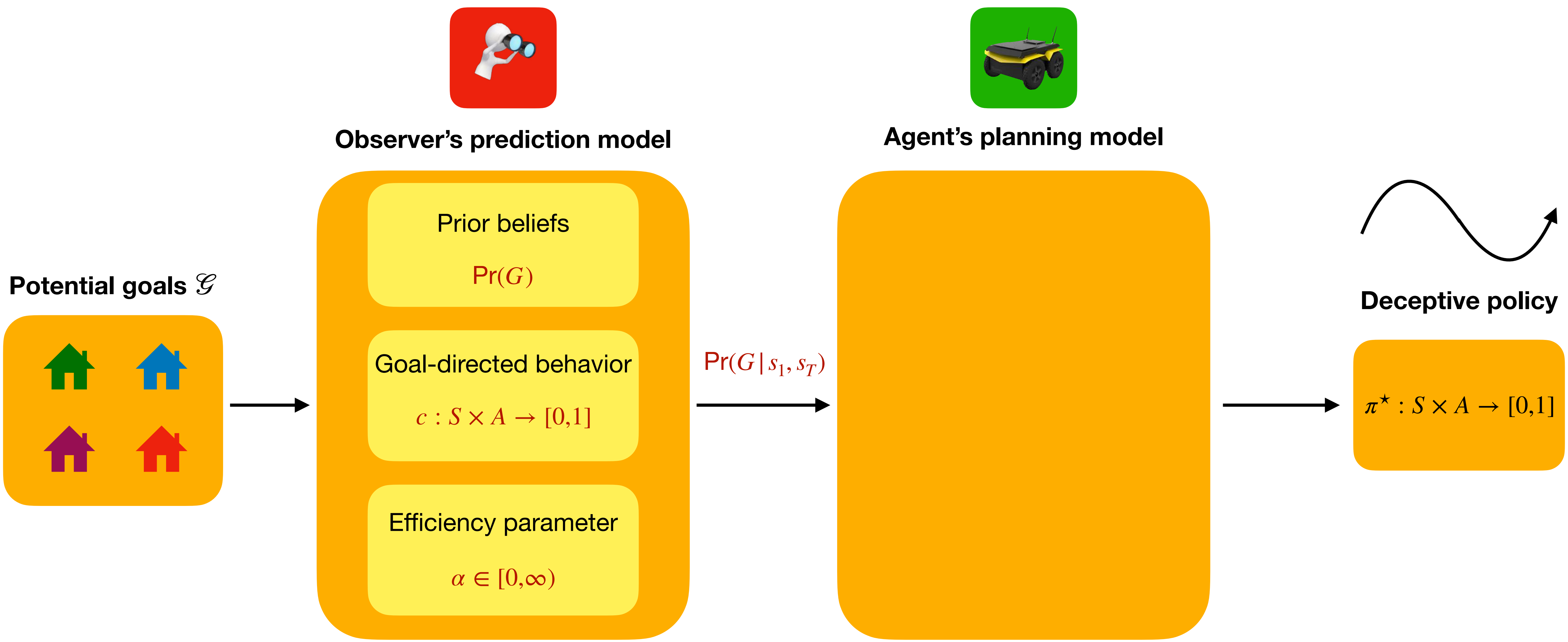
$\Pr(G \mid \zeta_{1:T}) = \Pr(G \mid s_1, s_T)$

only a function of the initial state
and the current state

Overall system model



Overall system model



Agent’s planning model: expressing deception as a cost function

We express deception objective as a generic cost function

$$f : S \times A \rightarrow [0,1]$$

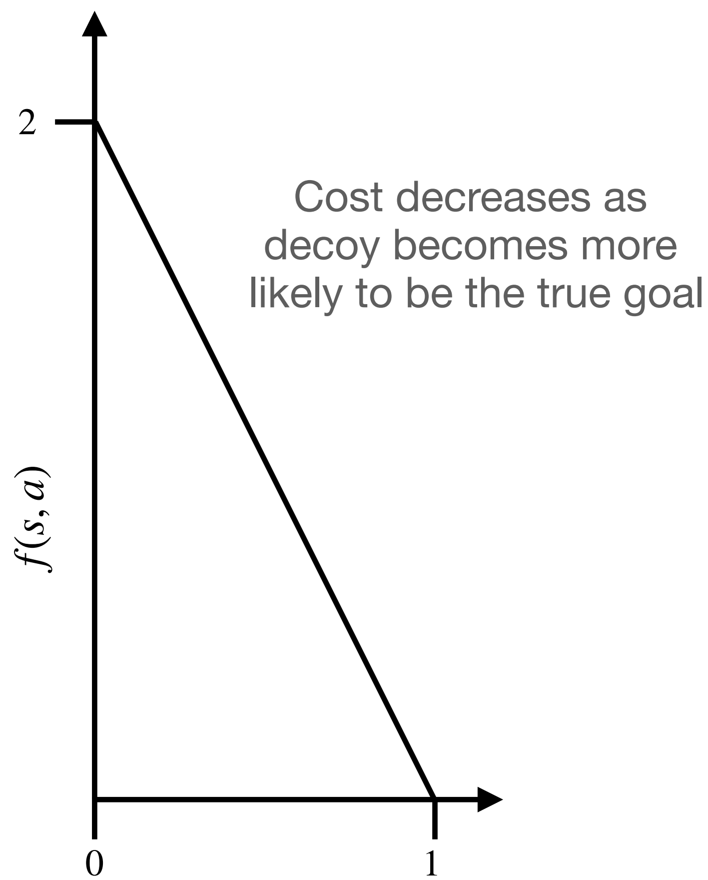
Agent's planning model: expressing deception as a cost function

We express deception objective as a generic cost function

$$f : S \times A \to [0,1]$$

Exaggeration:

$$f(s,a) = 1 + \Pr(G^\star \mid s_1, s) - \max_{G \in \mathcal{G} \setminus \{G^\star\}} \Pr(G \mid s_1, s)$$



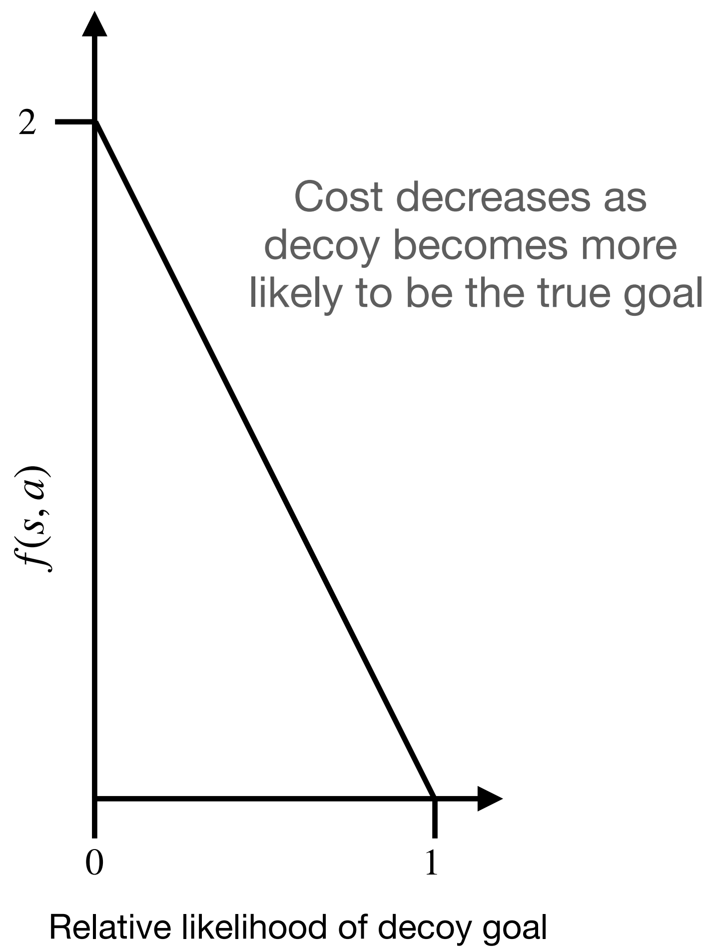
Agent's planning model: expressing deception as a cost function

We express deception objective as a generic cost function

$$f : S \times A \rightarrow [0,1]$$

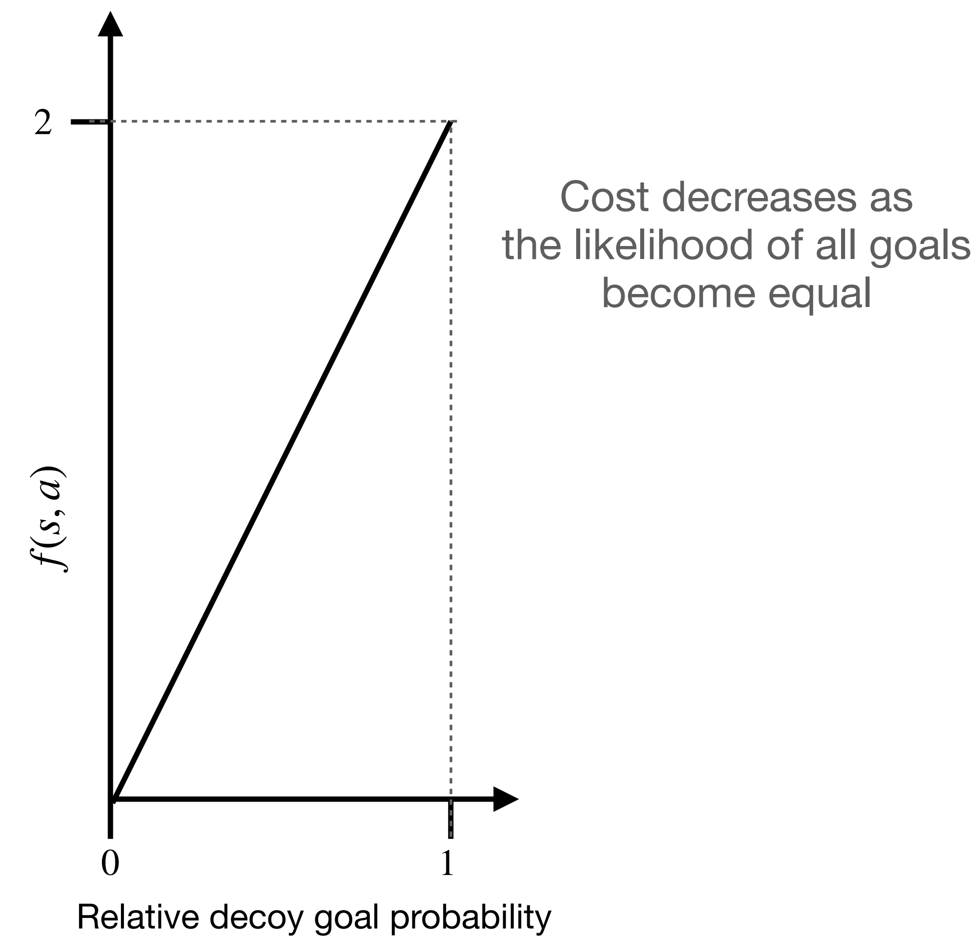
Exaggeration:

$$f(s,a) = 1 + \Pr(G^\star \mid s_1, s) - \max_{G \in \mathcal{G} \setminus \{G^\star\}} \Pr(G \mid s_1, s)$$



Ambiguity:

$$f(s,a) = \sum_{G \in \mathcal{G}} \sum_{G' \in \mathcal{G}} \left| \Pr(G \mid s_1, s) - \Pr(G' \mid s_1, s) \right|$$



Agent's planning model: a constrained optimization problem

The agent's objective is to reach its goal while deceiving the observer about its goal *for as long as possible*

$$\pi^\star \in \arg\min_{\pi \in \Pi} \; \mathbb{E}^\pi \Big[\sum_{t=0}^{\infty} \gamma_a^t f(s_t, a_t) \Big]$$

minimize total **discounted** cost

subject to: $\Pr^\pi(Reach[G^\star]) = R_{\max}(G^\star)$

reach the true goal $G^\star$
with maximum probability

Agent's planning model: a constrained optimization problem

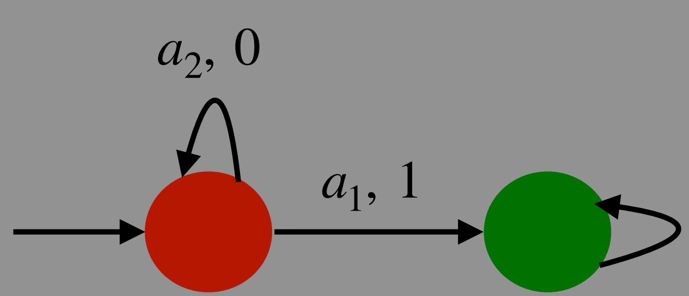
The agent's objective is to reach its goal while deceiving the observer about its goal *for as long as possible*

$$\pi^\star \in \arg\min_{\pi \in \Pi} \; \mathbb{E}^\pi \Big[\sum_{t=0}^{\infty} \gamma_a^t f(s_t, a_t) \Big]$$

⟵ minimize total **discounted** cost

subject to: $\Pr^\pi(Reach[G^\star]) = R_{\max}(G^\star)$

⟵ reach the true goal $G^\star$ with maximum probability



⟵ Infimum is zero and not attainable

An optimal policy $\pi^\star$ **may not exist**

Minimizing total discounted cost subject to reachability constraints

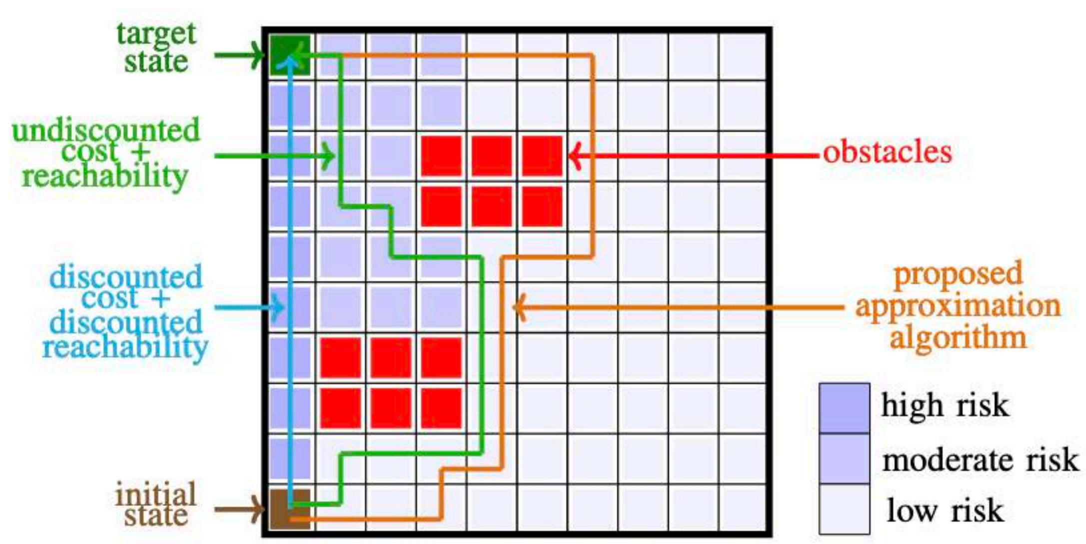
We provide a comprehensive analysis of this problem in [1].

We provide **necessary and sufficient** conditions for the existence of optimal policies.

We show that an $\epsilon-$optimal **stationary** policy exists and can be synthesized **efficiently.**

We show that it is **NP-hard** to synthesize an optimal **stationary deterministic** policy.

We show that a stationary deterministic policy **with suboptimality guarantees** can be synthesized **efficiently.**



[1] Y. Savas, C. Verginis, M. Hibbard, U. Topcu, ``On minimizing total discounted cost in MDPs subject to reachability constraints", IEEE Transactions on Automatic Control, 2022. (accepted)

Synthesizing policies via linear programming

Variables: $x(s, a)$ for all $s \in S, a \in A$

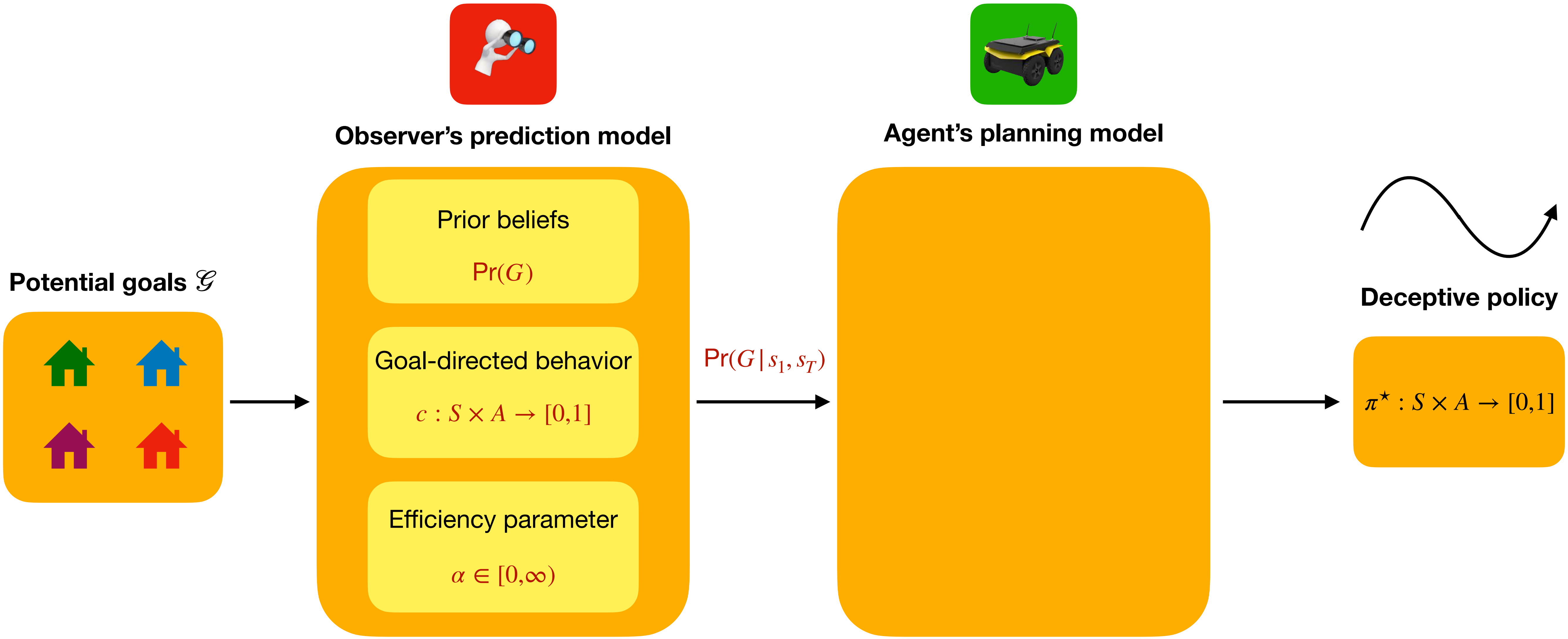
Constraints: $x(s, a) \geq 0$

$$\sum_{a \in A} x(s, a) - \sum_{t \in S} \sum_{a \in A} \mathbb{P}_{t,a,s} x(t, a) = \mathbb{I}\{s = s_1\}$$

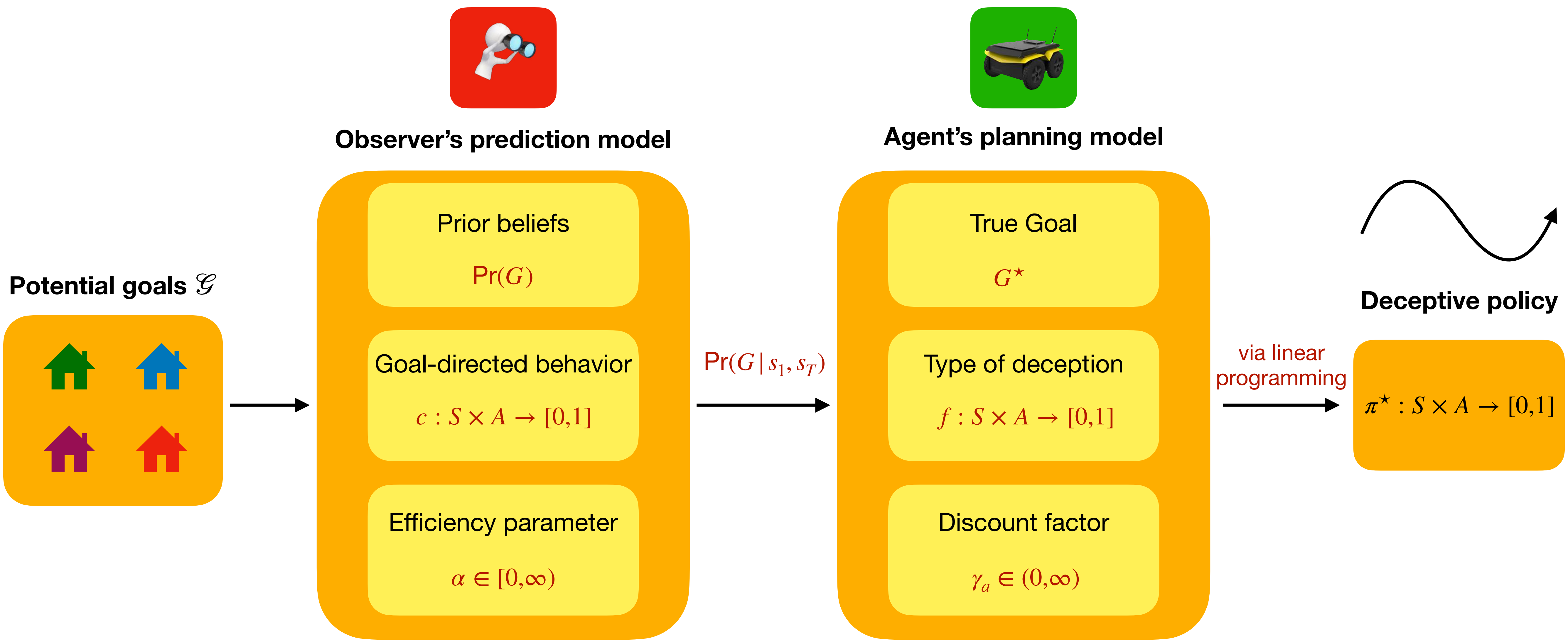
$$\sum_{t \in S} \sum_{a \in A} x(t, a) P_{t,a,G^\star} = R_{\max}(G^\star)$$

Objective: $\min_{x(s,a)} \quad \sum_{s \in S} \sum_{a \in A} x(s, a) g(s, a)$

Overall system model

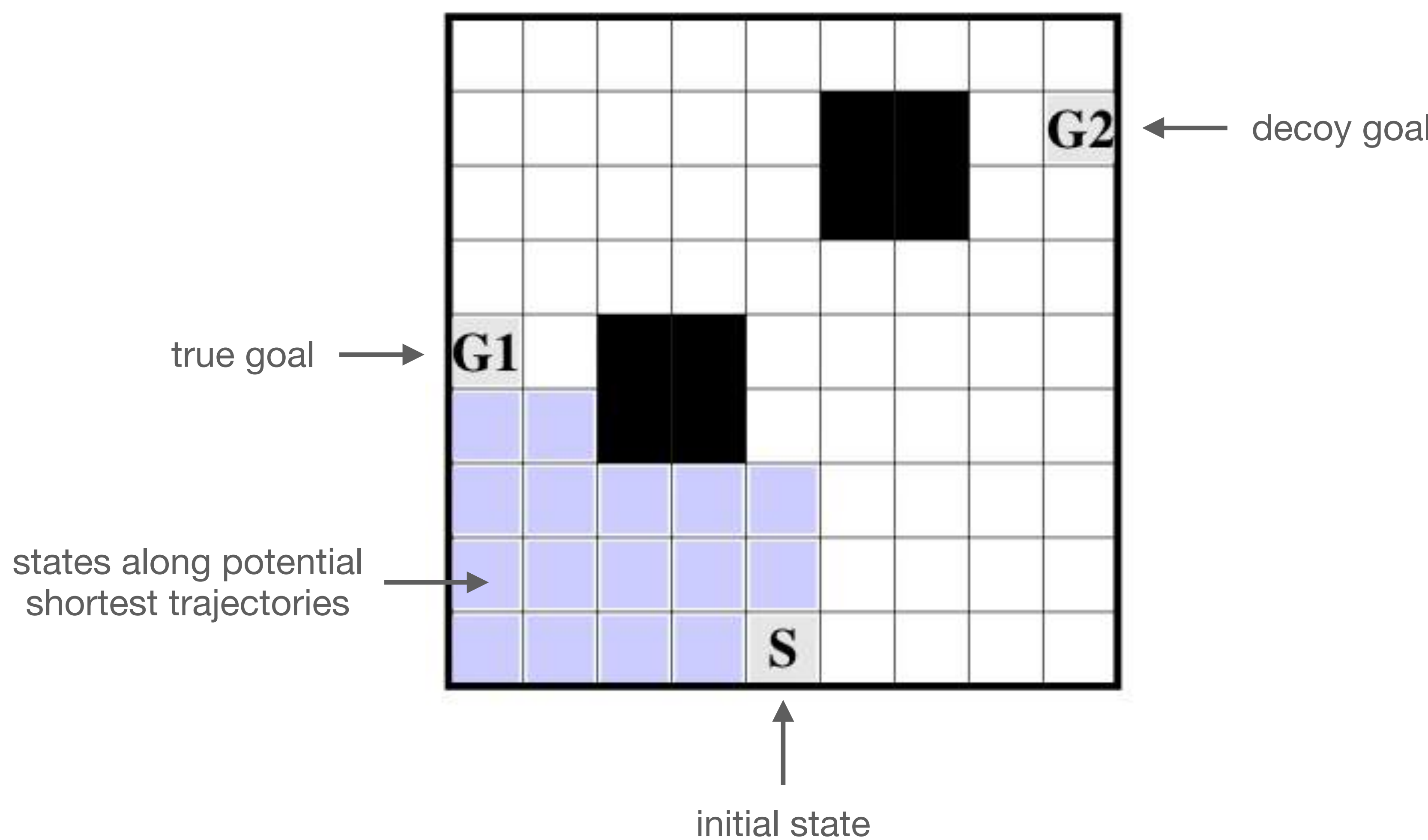


Overall system model



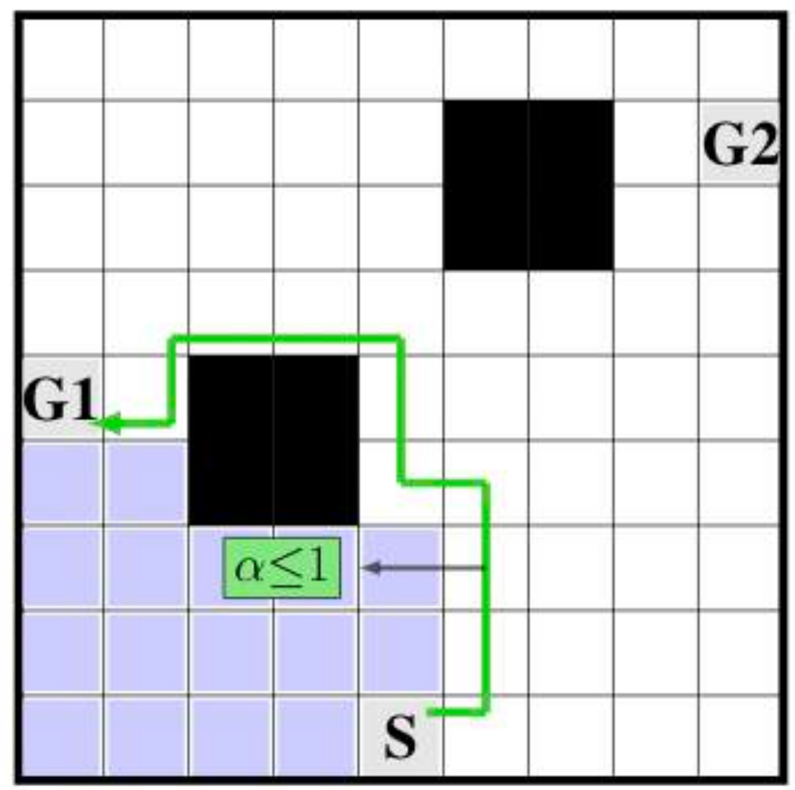
Generating tunable deceptive behavior

The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2 .



Generating tunable deceptive behavior

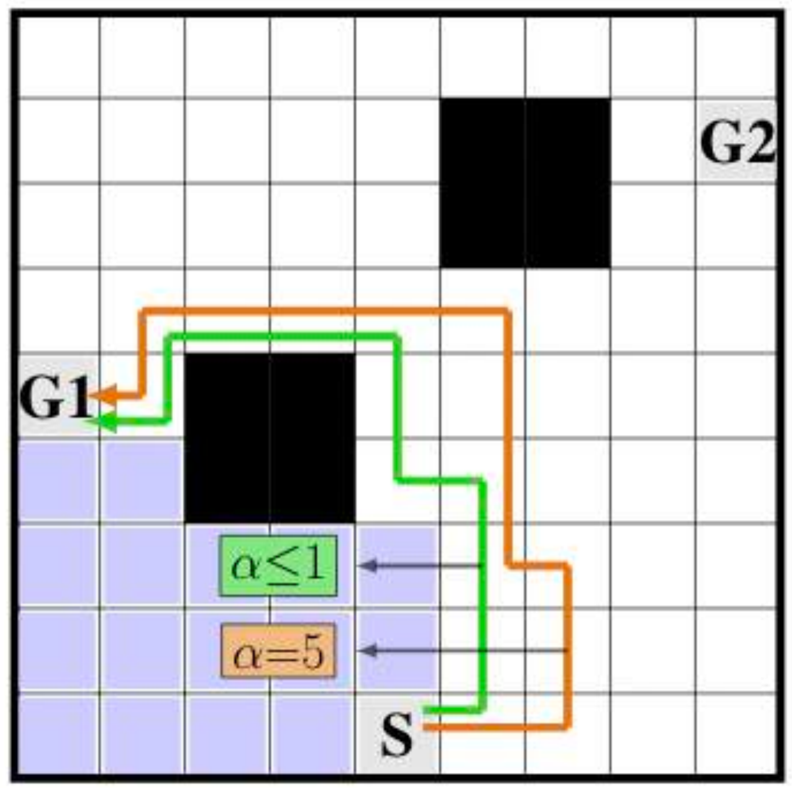
The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2.



Effect of α: is the agent expected to be efficient?

Generating tunable deceptive behavior

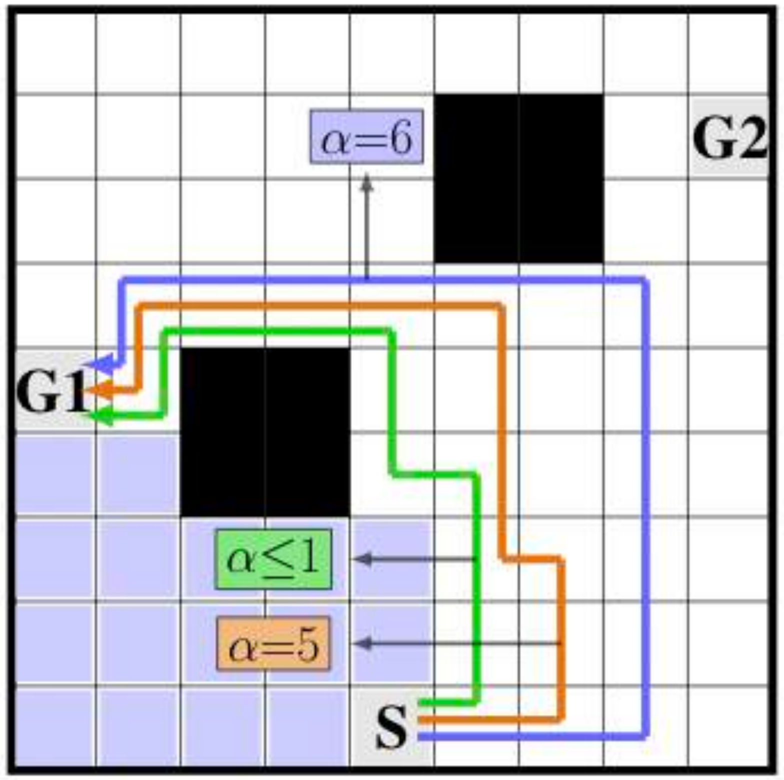
The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2 .



Effect of α: is the agent expected to be efficient?

Generating tunable deceptive behavior

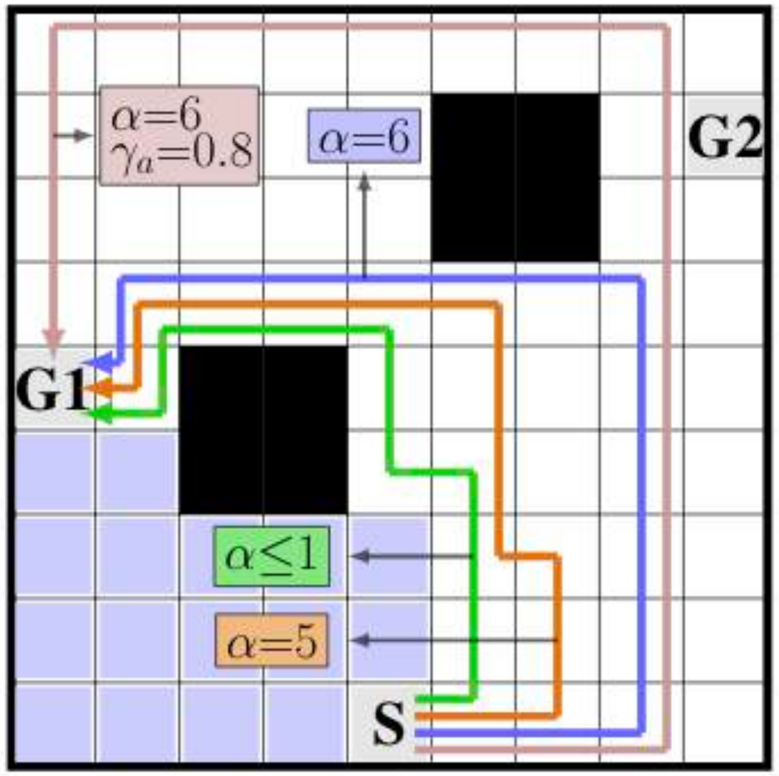
The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2.



Effect of α: is the agent expected to be efficient?

Generating tunable deceptive behavior

The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2 .

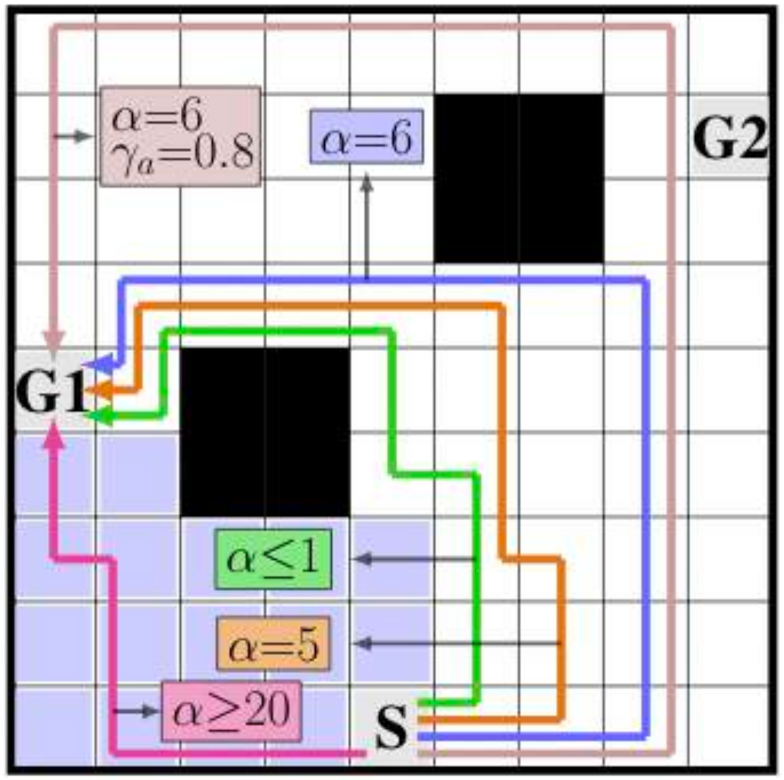


Effect of α: is the agent expected to be efficient?

Effect of γ_a: let's exploit further regions for deception.

Generating tunable deceptive behavior

The agent starts from S and aims to reach its goal G_1 while **exaggerating** its behavior towards G_2 .

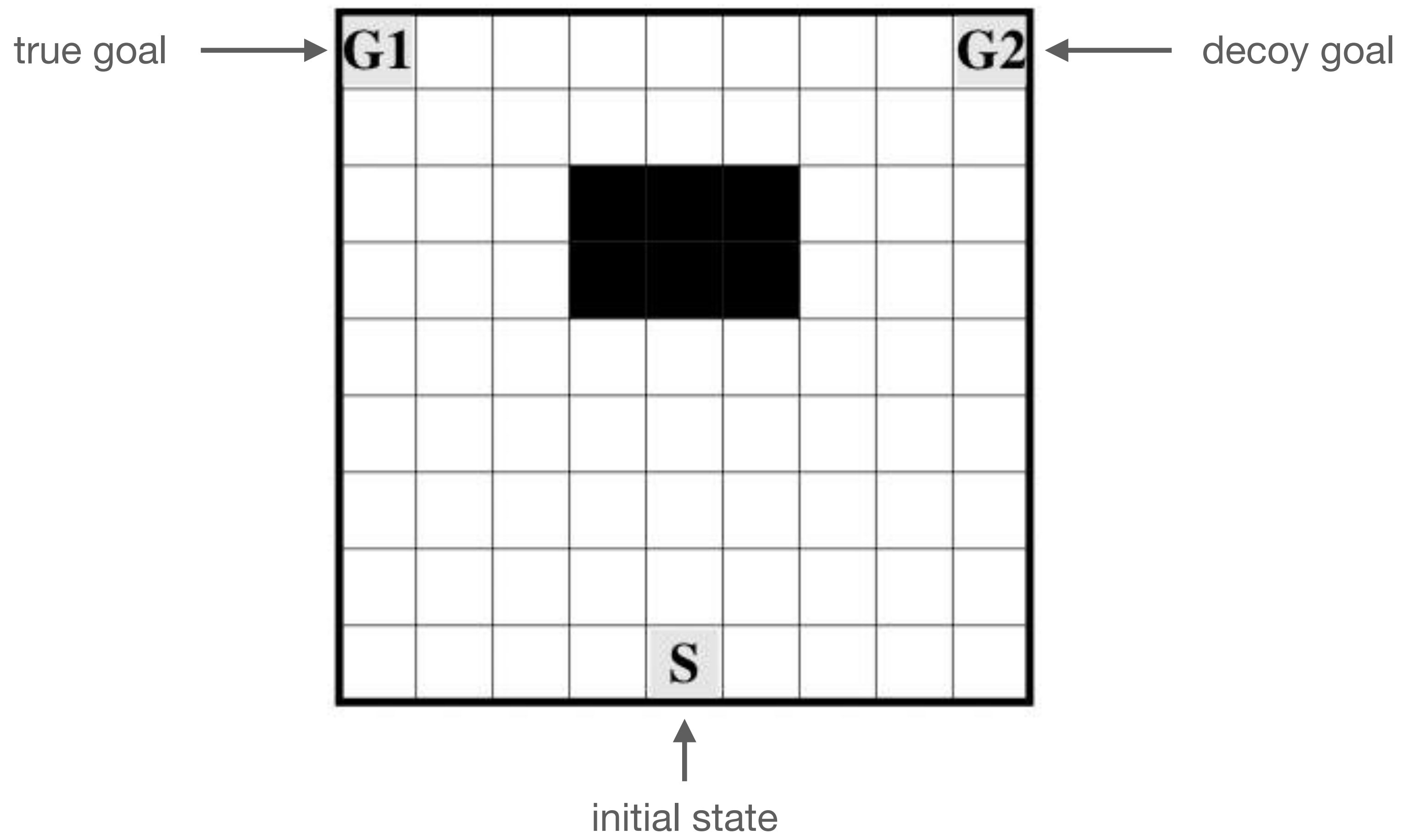


Effect of α: is the agent expected to be efficient?

Effect of γ_a: let's exploit further regions for deception.

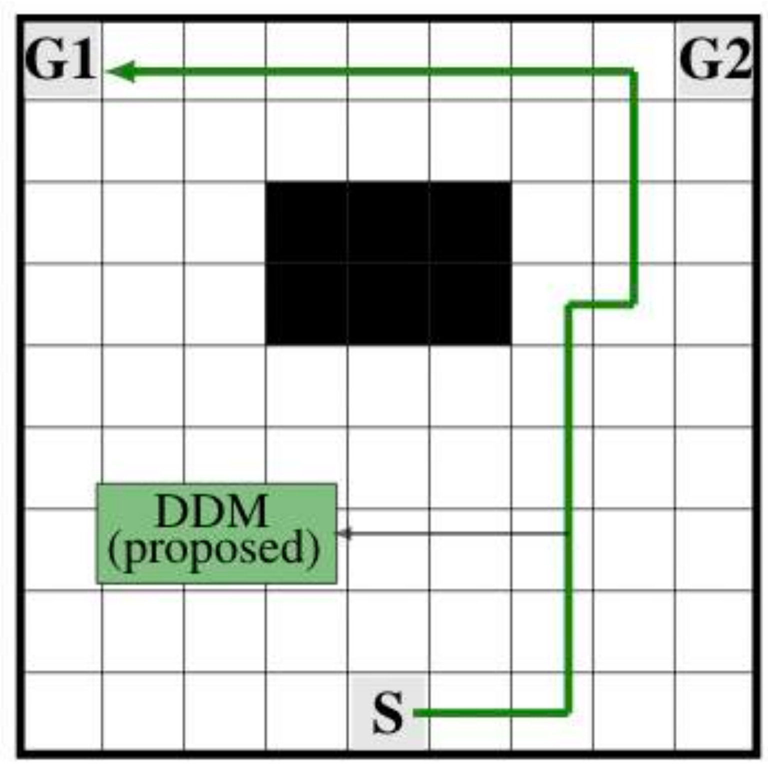
User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.



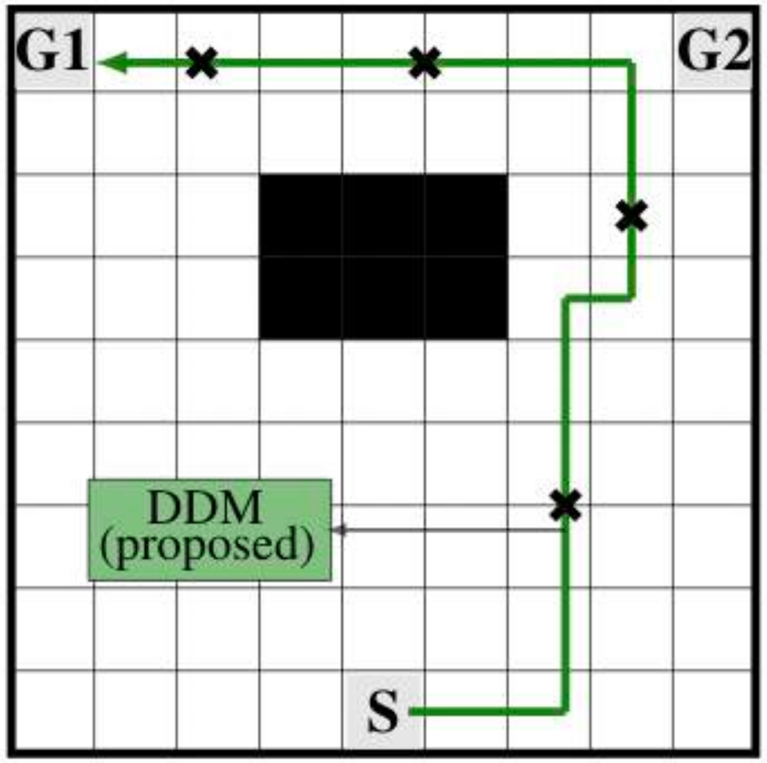
User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.



User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.

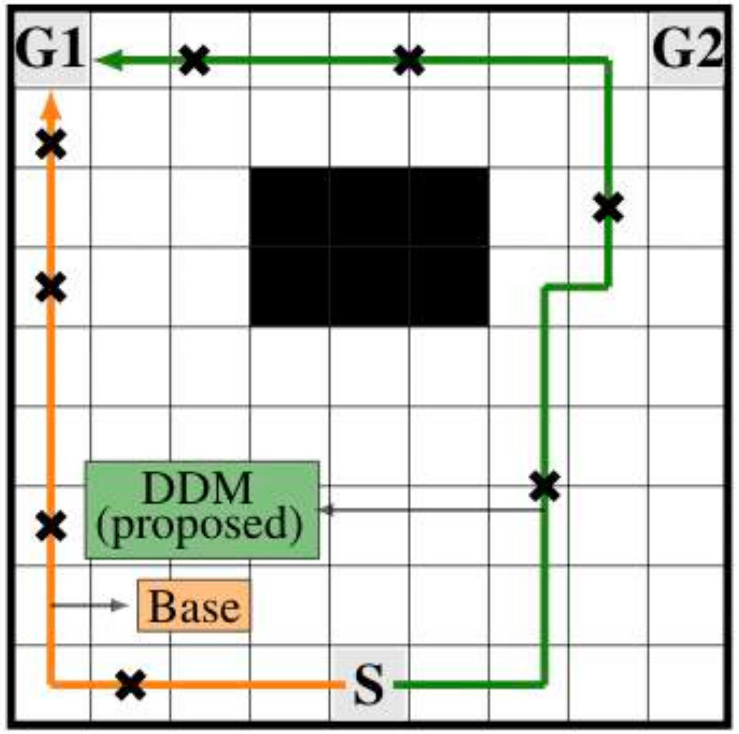


Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

Question 2: How **confident** are you in the robot's goal?

User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.



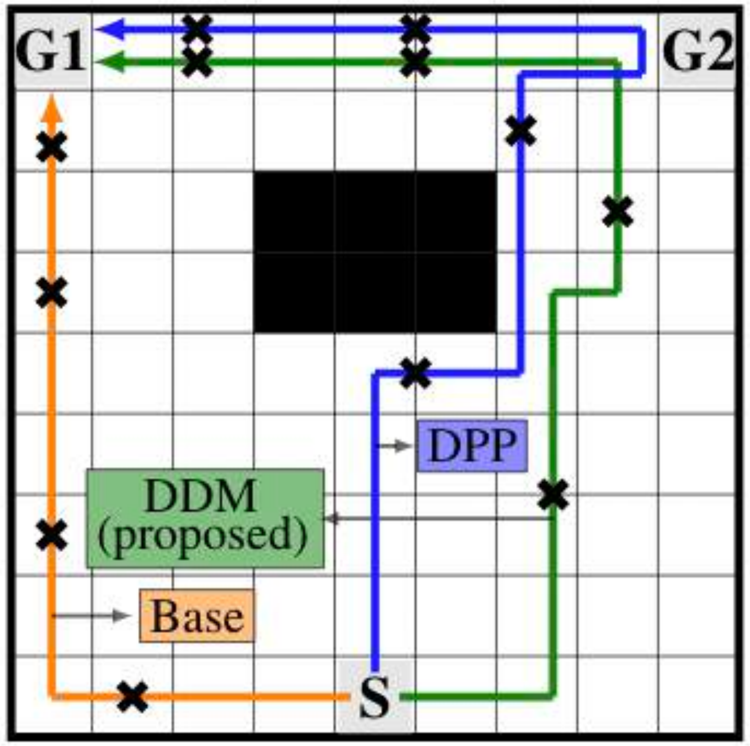
Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

Question 2: How **confident** are you in the robot's goal?

Base: Baseline trajectory (shortest path to the goal)

User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.



Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

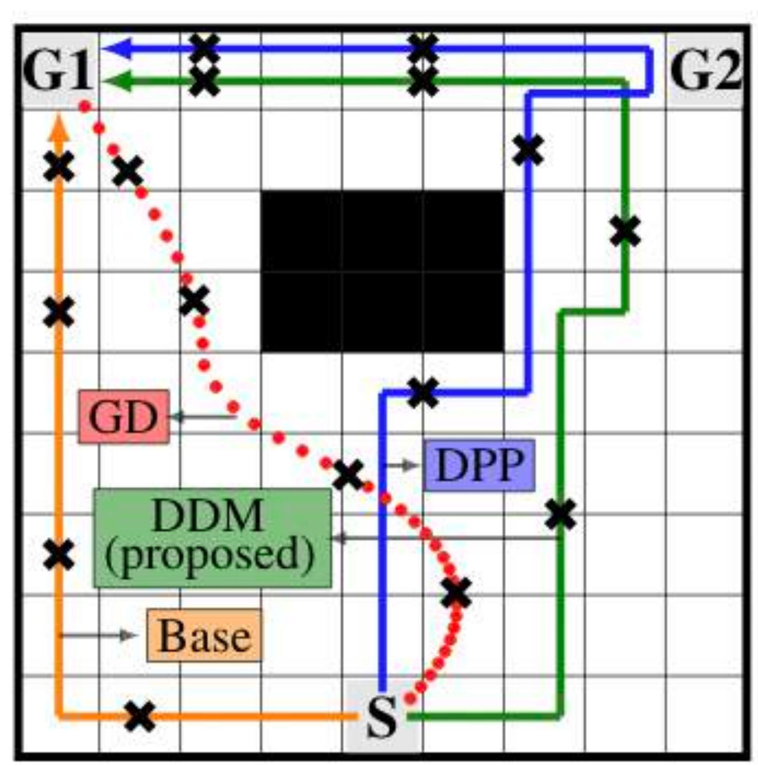
Question 2: How **confident** are you in the robot's goal?

Base: Baseline trajectory (shortest path to the goal)

DPP: A heuristic approach utilizing least deceptive point

User Study 1 - the importance of global optimality

A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.



Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

Question 2: How **confident** are you in the robot's goal?

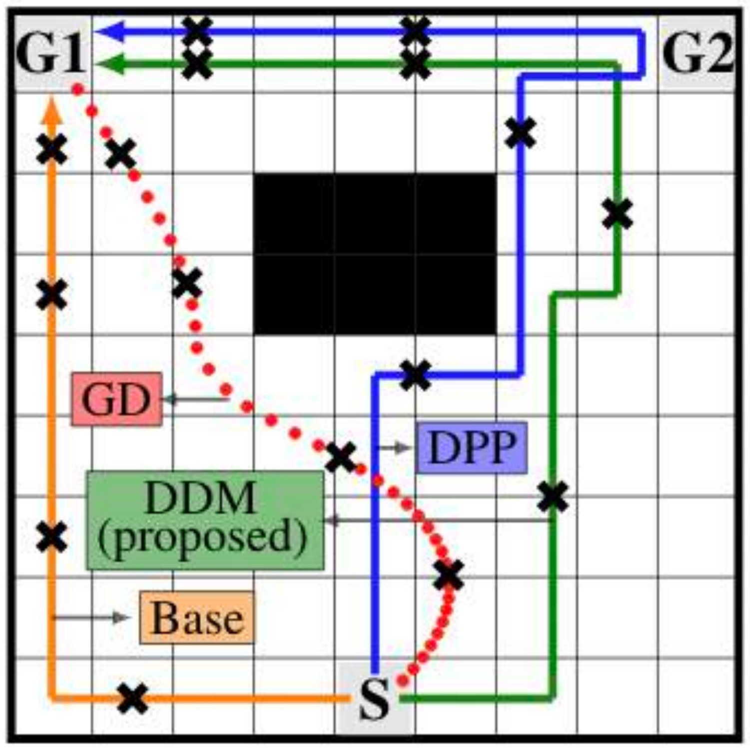
Base: Baseline trajectory (shortest path to the goal)

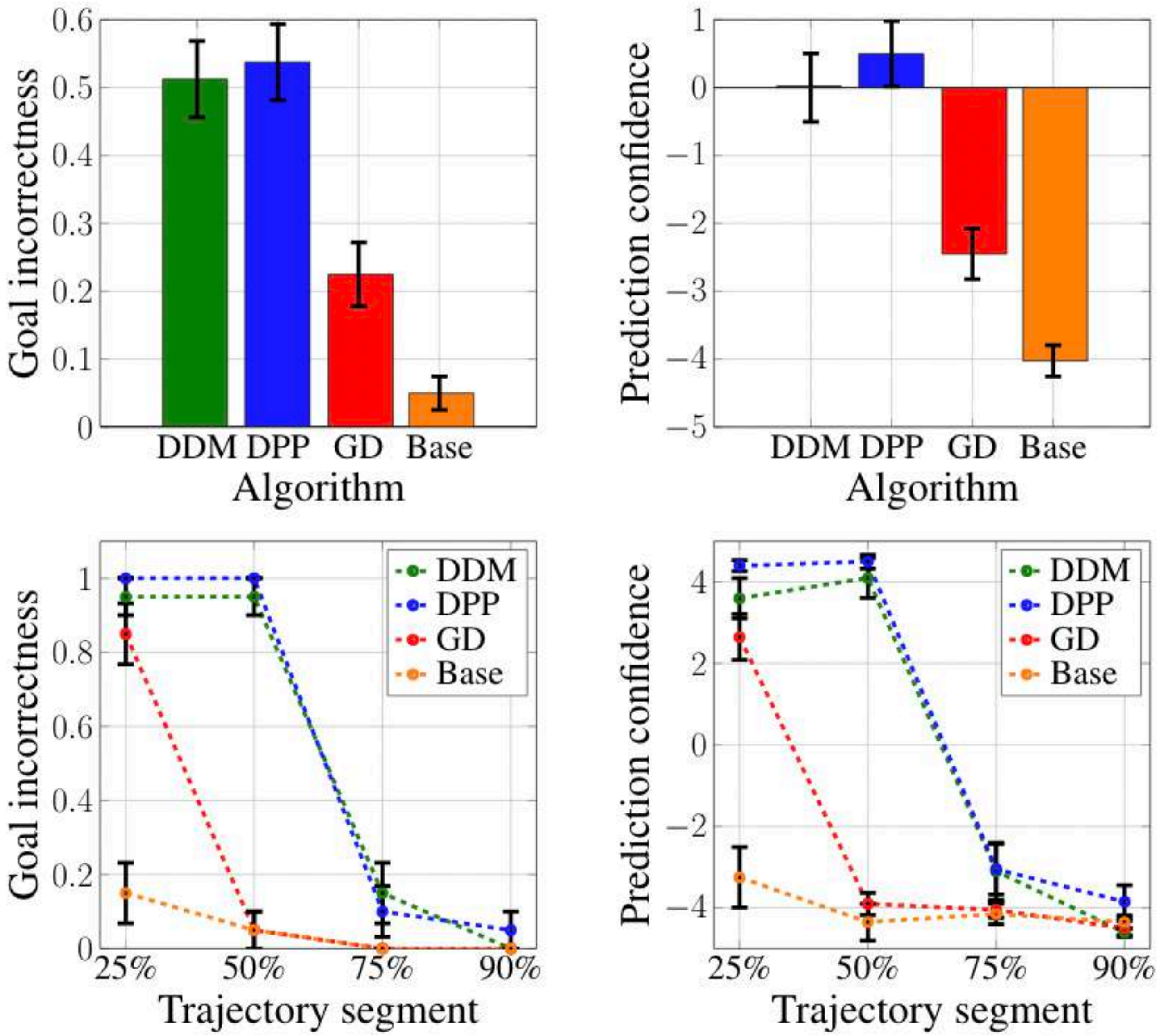
DPP: A heuristic approach utilizing least deceptive point

GD: A functional gradient descent-based approach

User Study 1 - the importance of global optimality

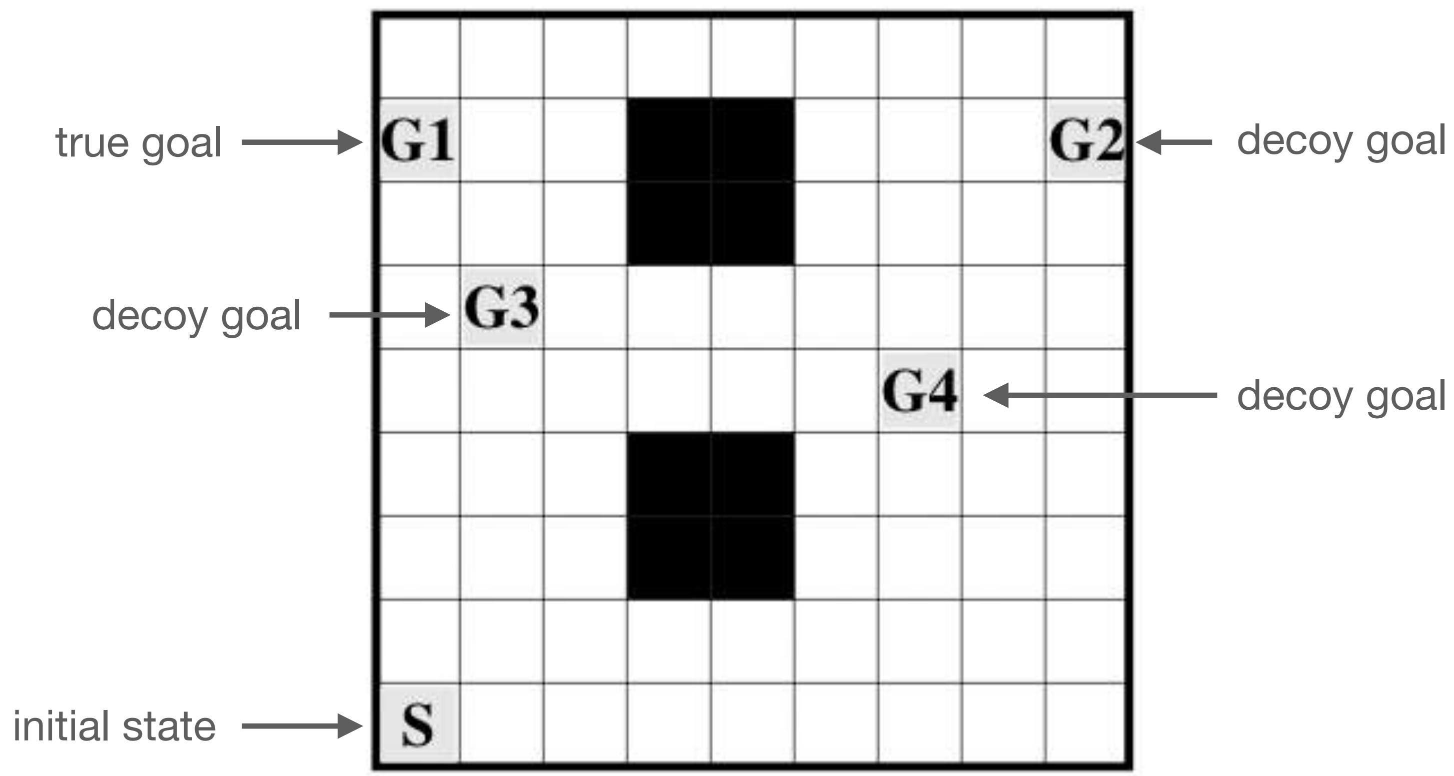
A user study via Amazon MTurk (320 participants) to illustrate the benefits of global guarantees.





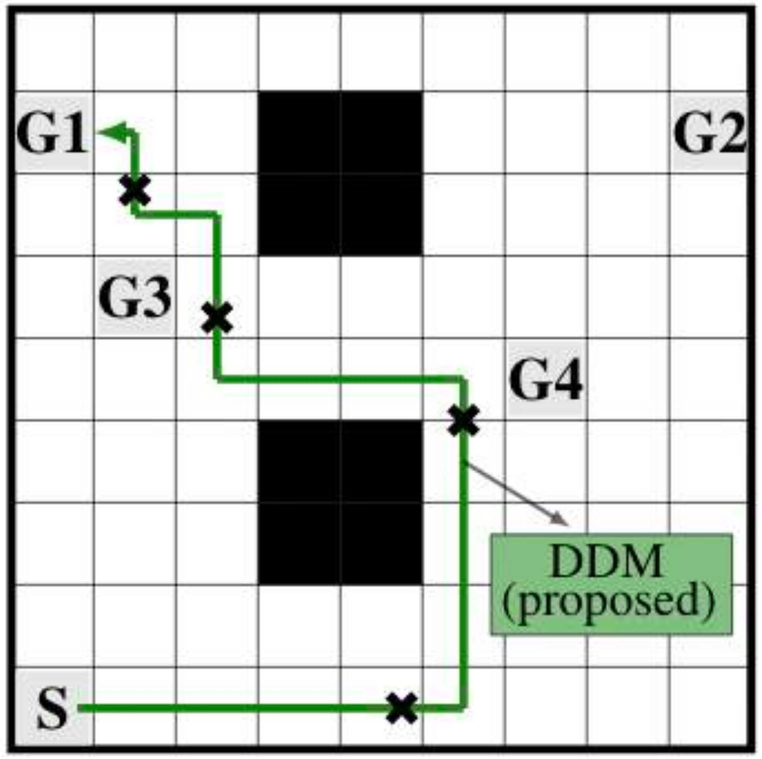
User Study 2 - the importance of prediction-awareness

A user study via Amazon MTurk (240 participants) to illustrate the benefits of prediction-awareness.



User Study 2 - the importance of prediction-awareness

A user study via Amazon MTurk (240 participants) to illustrate the benefits of prediction-awareness.



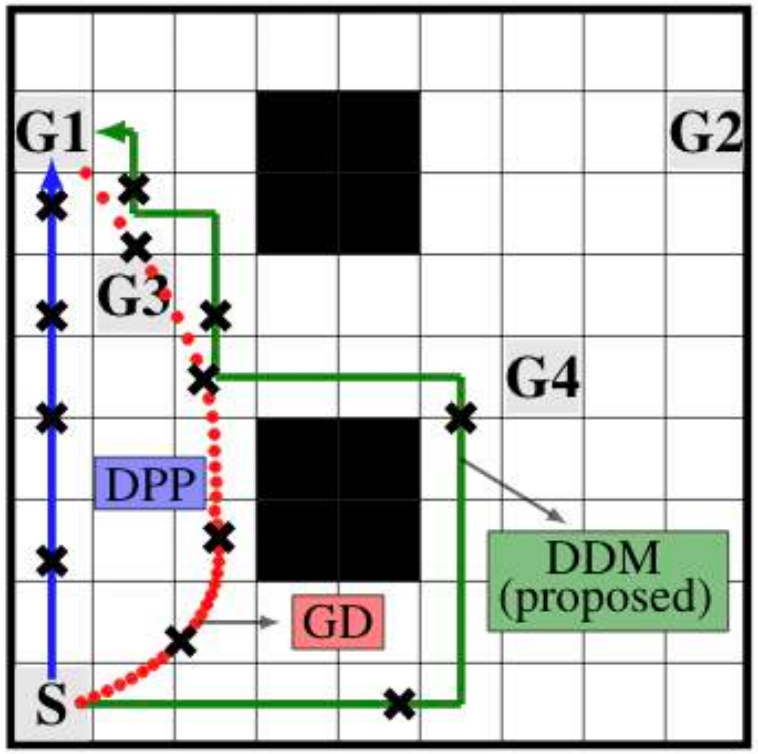
Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

Question 2: How **confident** are you in the robot's goal?

Question 3: Based on the robot's partial trajectory, which one do you think is the **robot's second most likely goal**?

User Study 2 - the importance of prediction-awareness

A user study via Amazon MTurk (240 participants) to illustrate the benefits of prediction-awareness.



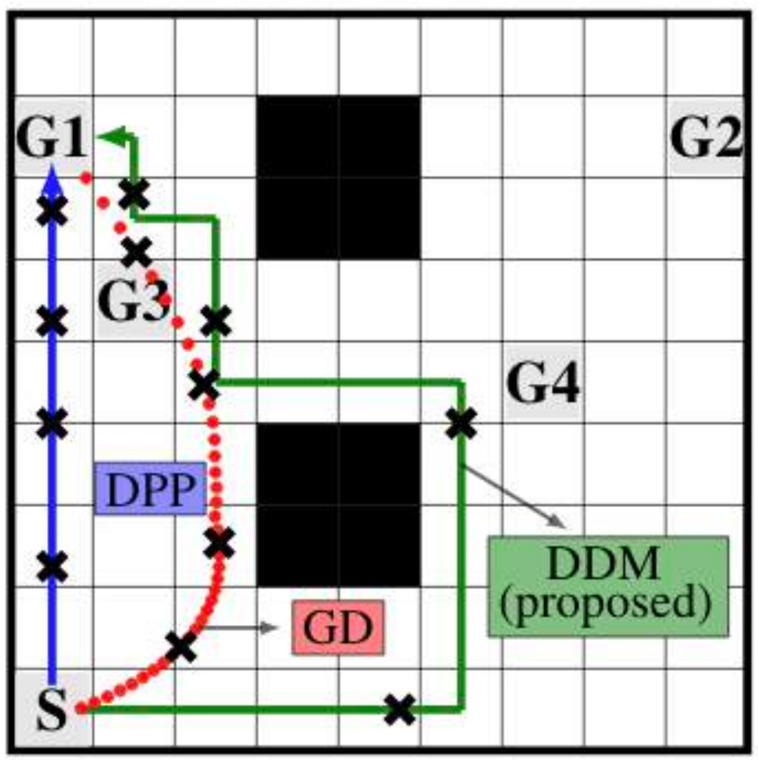
Question 1: Based on the robot's partial trajectory, which one do you think is the **robot's goal**?

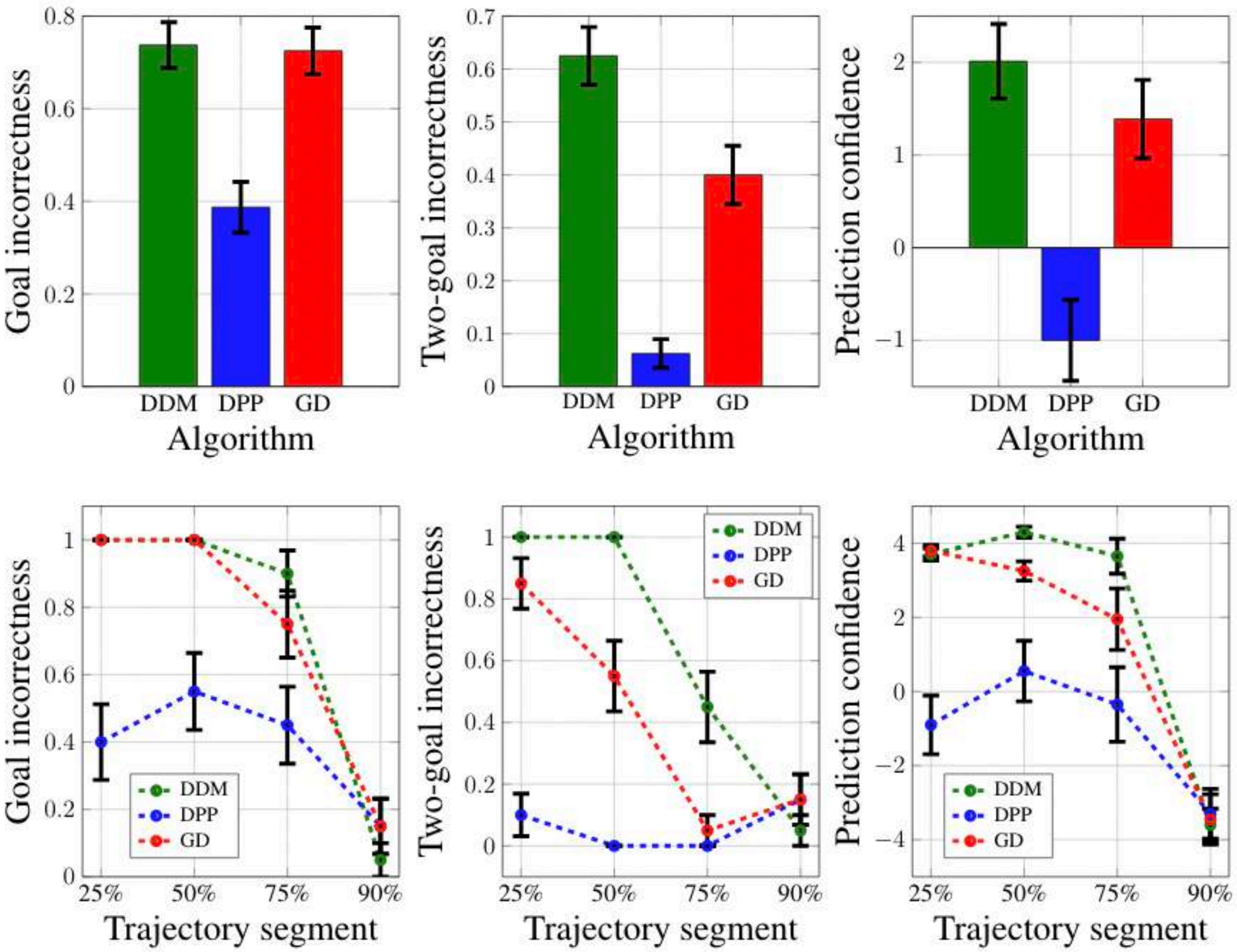
Question 2: How **confident** are you in the robot's goal?

Question 3: Based on the robot's partial trajectory, which one do you think is the **robot's second most likely goal**?

User Study 2 - the importance of prediction-awareness

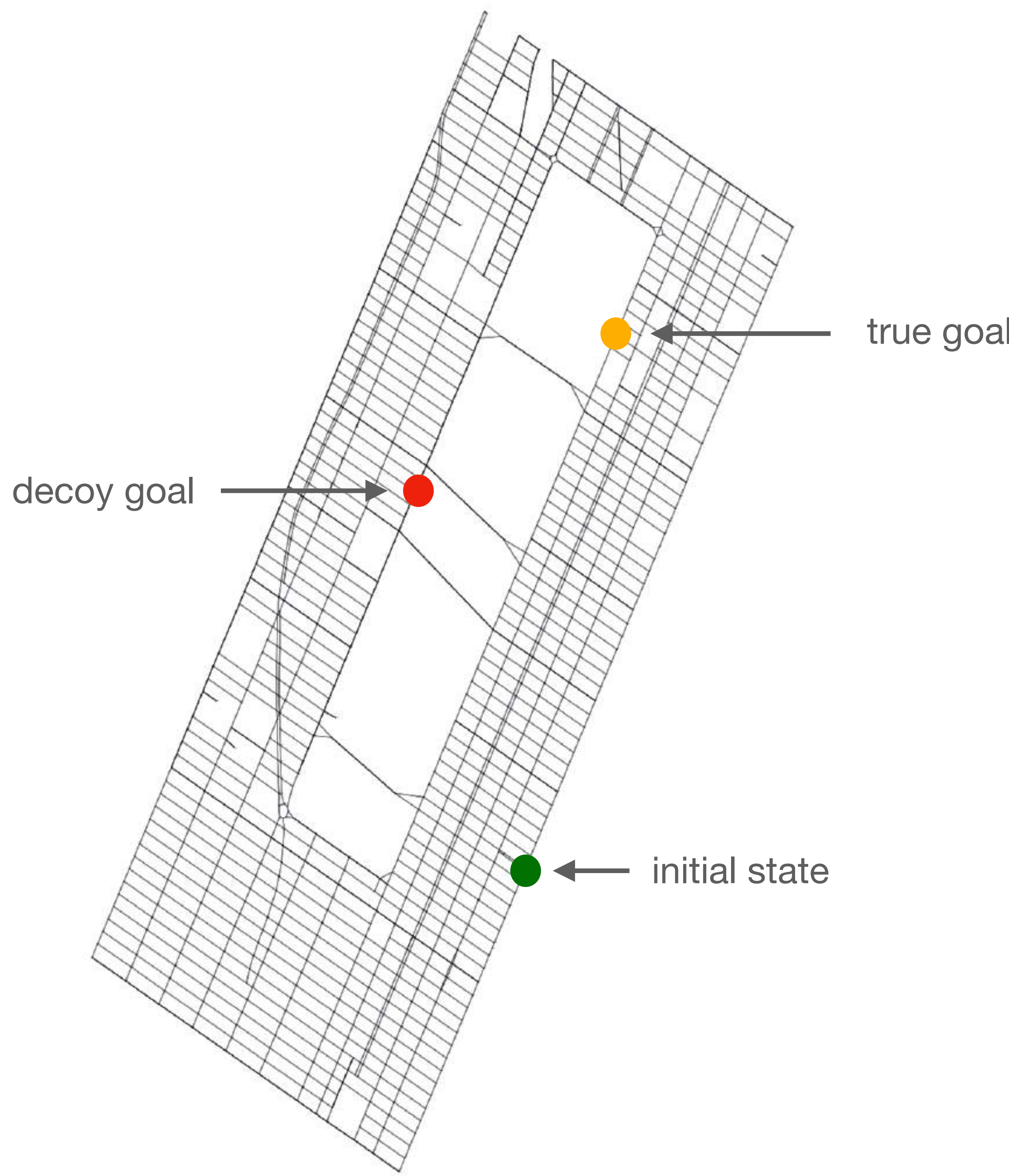
A user study via Amazon MTurk (240 participants) to illustrate the benefits of prediction-awareness.

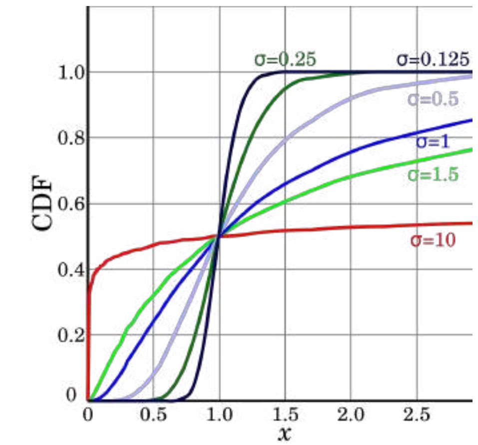




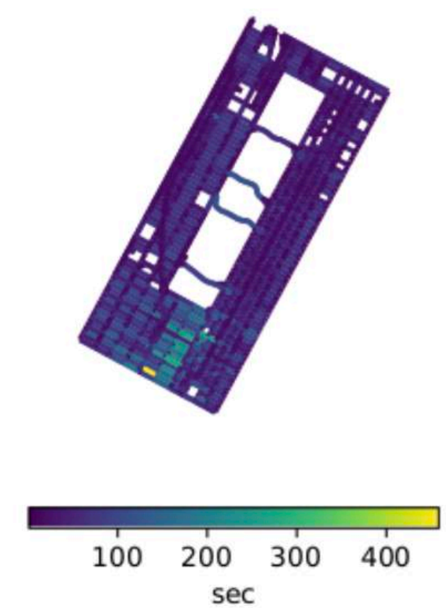
A case study in the streets of Manhattan, New York

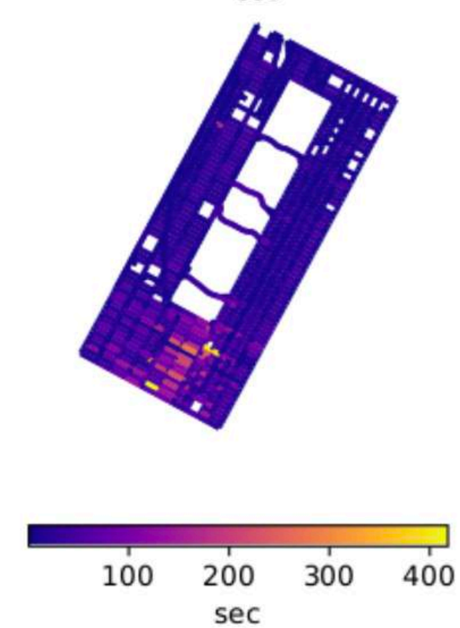
Real travel speed data from an open-source database in [1].





Lognormal travel time distribution

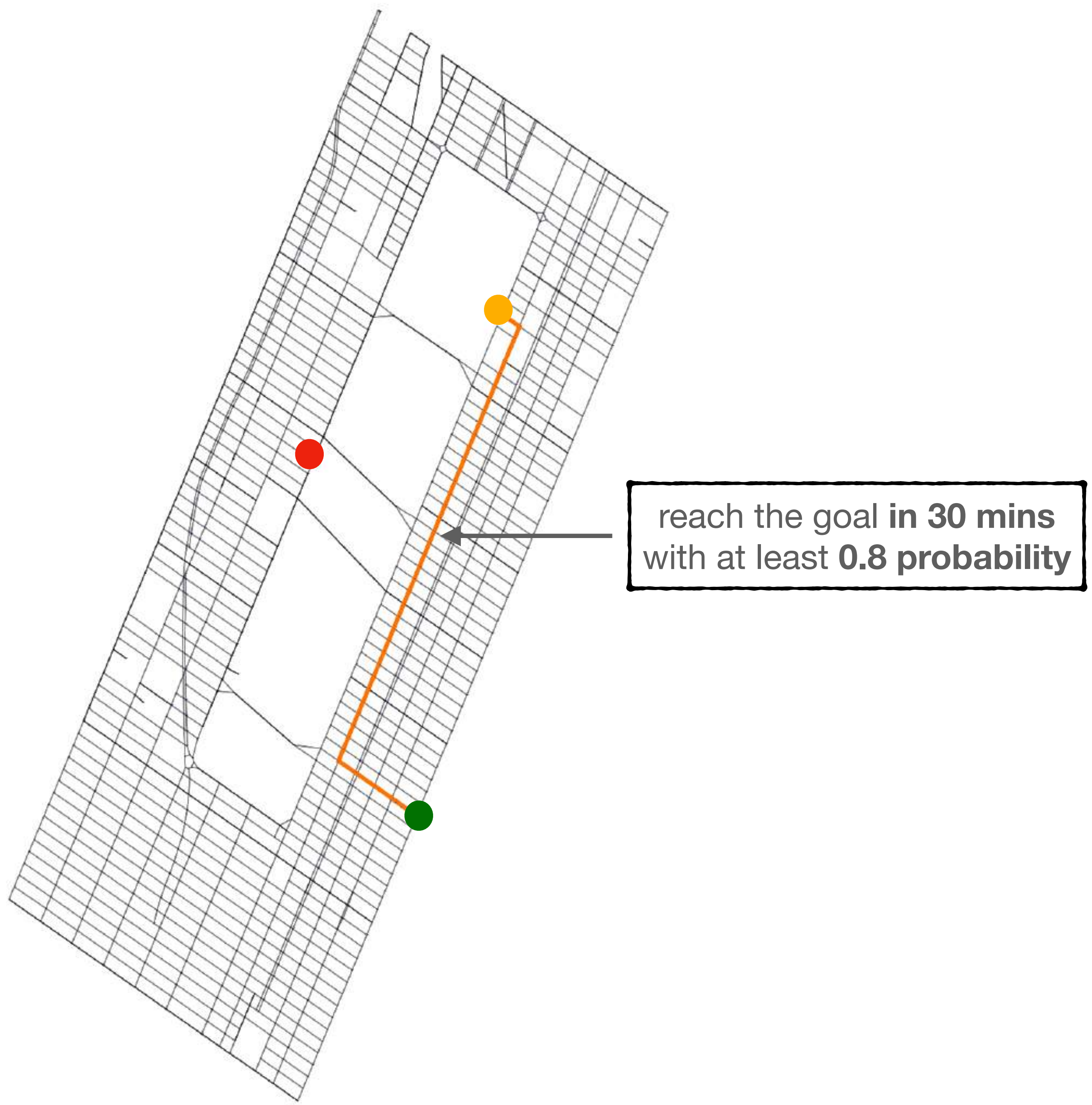


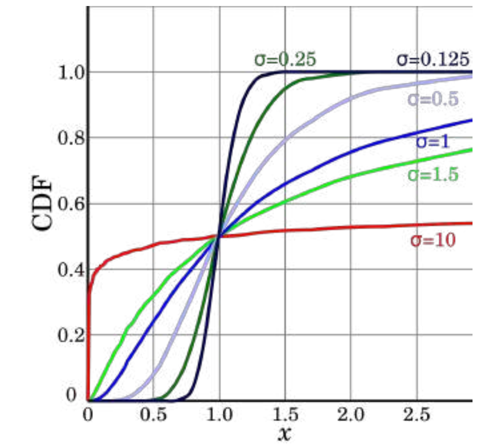


[1] Uber Technologies, I. 2021. Uber movement.

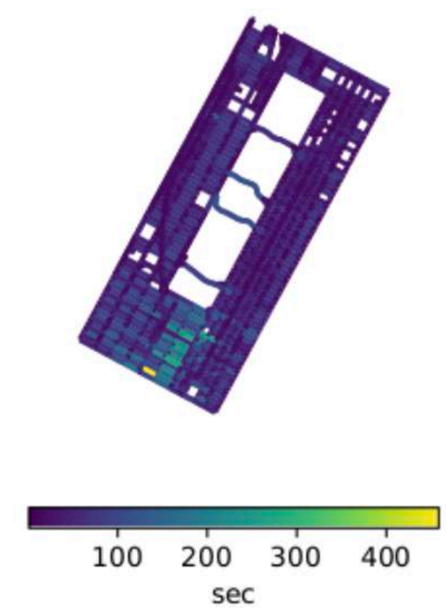
A case study in the streets of Manhattan, New York

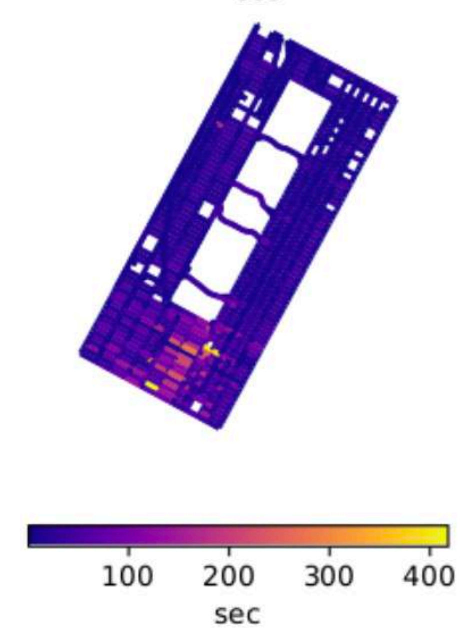
Real travel speed data from an open-source database in [1].





Lognormal travel time distribution

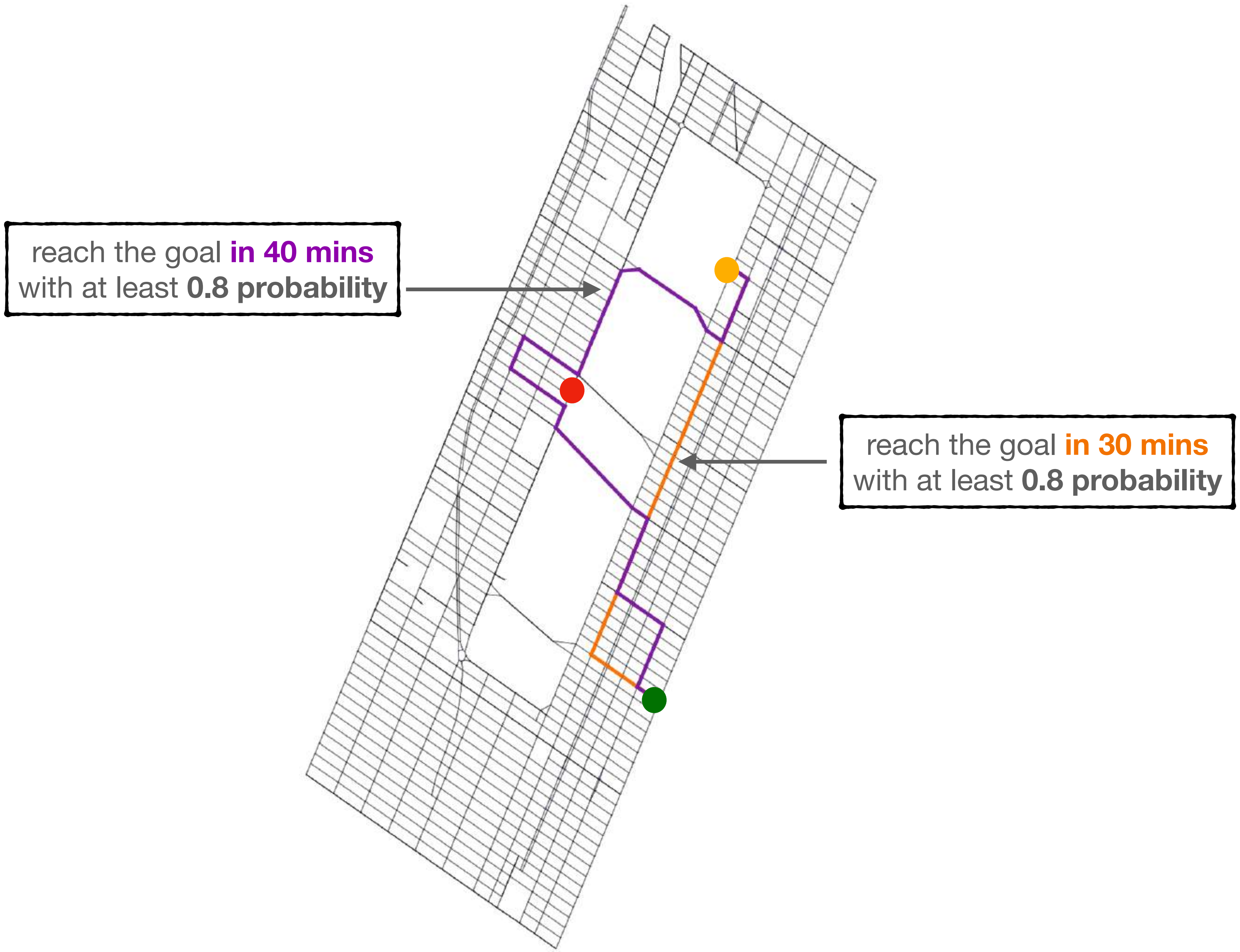


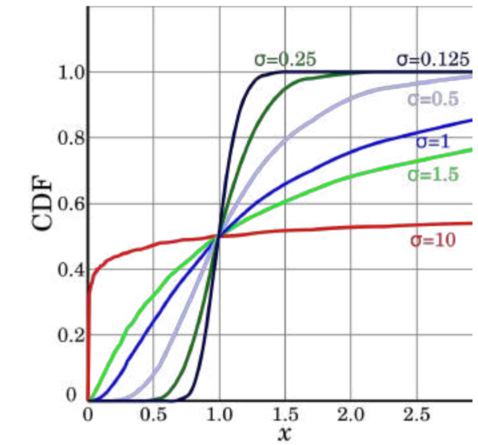


[1] Uber Technologies, I. 2021. Uber movement.

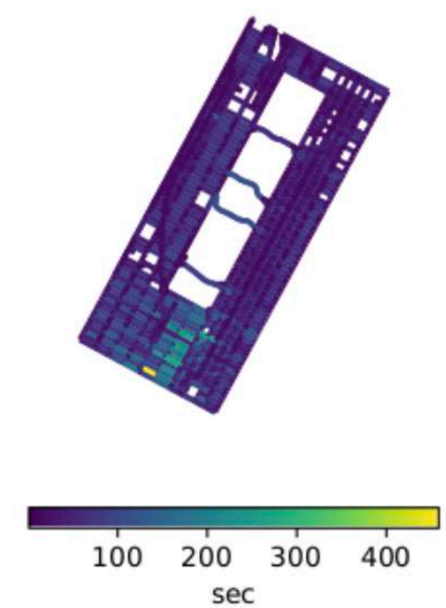
A case study in the streets of Manhattan, New York

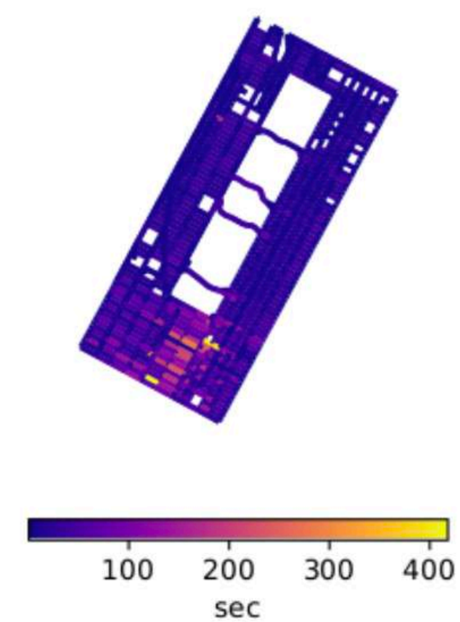
Real travel speed data from an open-source database in [1].





Lognormal travel time distribution





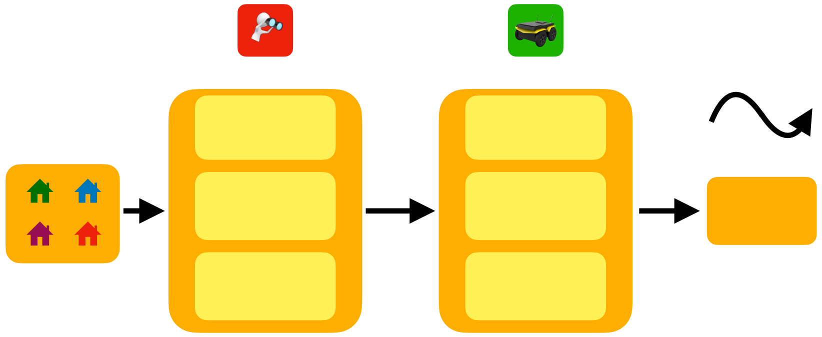
[1] Uber Technologies, I. 2021. Uber movement.

Conclusions

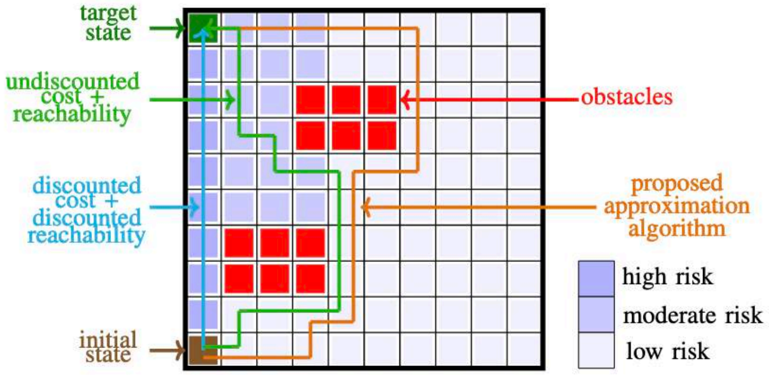
Deceptive capabilities have the potential to improve security in autonomy.

We propose an efficient deception algorithm that

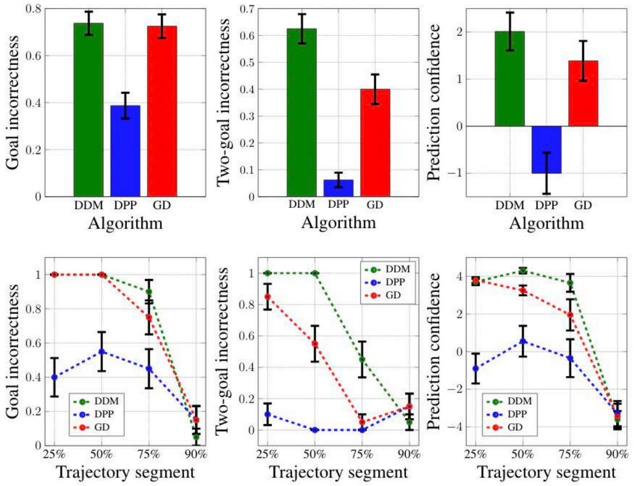
- works in stochastic environments,
- adjusts behavior according to predictions,
- and has global performance guarantees.



We present a comprehensive analysis for minimizing total discounted cost in MDPs subject to reachability constraints.



We show the effectiveness of the proposed method through user studies.



Thanks for listening!

Yagiz Savas

yagiz.savas@utexas.edu

autonomous SYSTEMS GROUP

TEXAS The University of Texas at Austin