

# Toward Safe and Efficient Human–Robot Interaction via Behavior-Driven Danger Signaling

Mehdi Hosseinzadeh<sup>ID</sup>, *Member, IEEE*, Bruno Sinopoli<sup>ID</sup>, *Fellow, IEEE*, and Aaron F. Bobick, *Fellow, IEEE*

**Abstract**—This article introduces the *notion of danger awareness* in the context of human–robot interaction, which decodes whether a human is aware of the existence of the robot, and illuminates whether the human is willing to engage in ensuring safety. This article also quantifies the notion as a single binary variable, called *danger awareness coefficient*, and provides a game-theoretic interpretation for that. Employing an online Bayesian learning method to update the robot’s belief about the human’s danger awareness by observing their actions, it is shown how the robot can build a predictive human model to anticipate the human’s future actions. To enrich robot’s observations, and thus to improve safety and efficiency of the robot, the robot is equipped with a *danger signaling* system to generate awareness in the human. Finally, a planning scheme is proposed to provide an efficient and probabilistically safe plan for the robot. The effectiveness of the proposed scheme is demonstrated through simulation studies on an interaction between a self-driven car and a pedestrian.

**Index Terms**—Behavior-driven planning, danger awareness in humans, danger signaling, human–robot interaction, robot action planning.

## I. INTRODUCTION

ENABLING efficient and safe interaction among all participating agents (i.e., humans and robots) is a challenging task in human–robot interaction. Safety enforcement techniques depend on some presumptions and assumptions that will not necessarily be true in practice. Thus, robots will inevitably encounter incomplete and possibly erroneous knowledge of the environment and other agents, humans in particular, which may degrade the efficiency of the robot. This implies that robots must safely and timely reason over the uncertainties of the environment they operate in to maintain a safe and efficient interaction.

### A. Notion of Danger Awareness

Reasoning in uncertain environments is an area where humans excel compared to current robots. With that motiva-

tion, this article uses models of human decision-making from cognitive science to develop a framework that enables robots to reason over the uncertainties inherent in predicting the actions of humans to improve safety and efficiency. More specifically, this article introduces the *notion of danger awareness* which can be used to decode whether a human is aware of the existence of other agents and possible dangers that they may present, and to explicate whether human is willing to engage in ensuring safety.

This article quantifies the notion of danger awareness as a binary variable which is called *danger awareness coefficient*. Supported with a game-theoretic interpretation, it is shown that a binary variable is appropriate and sufficient to model the impact of human’s danger awareness on their behavior. This article also proposes a method to continually learn the coefficient based upon robot’s real-time observations. A robot action planning scheme which incorporates human’s danger awareness is also proposed to provide a *probabilistically safe* and efficient plan for the robot.

### B. Danger Signaling

Most of the existing methods in robotics literature to learn humans’ behavior (e.g., [1], [2], [3]) are based upon *passive* observations of humans’ states and actions. We argue that it is implausible to accurately learn the humans’ behavior through passive observations, as their trajectories may not encode sufficient information about them. As a result, any robot action planner developed based upon passive observations might be tremendously inefficient, leading to conservative solutions. One possible way to address this issue is to enable the robot to influence humans to enrich its observations [4], [5], [6]; this would reveal humans intention [7], [8] about cooperating with robots to ensure safety. For this purpose, we assume that the robot is equipped with a *danger signaling system*. This system creates a communication channel between the robot and the human, where the robot can convey a message to the human and receives the human’s replies manifested in their actions. By means of the danger signal, the robot can *actively* aptly perturb the environment so that the bird’s-eye view of the human’s behavior observed by the robot is rich enough to reason about their opinion on the cooperative safety.

### C. Contribution

Two key contributions of this article are: 1) introduction of a quantifiable notion of danger awareness and 2) development of a learning-based robot action planning scheme to obtain a probabilistically safe and efficient plan by leveraging the aforementioned notion.

Manuscript received 21 June 2023; accepted 3 August 2023. This work was supported by the National Science Foundation under Award ECCS-2020289. Recommended by Associate Editor G. Notarstefano. (*Corresponding author: Mehdi Hosseinzadeh.*)

Mehdi Hosseinzadeh is with the School of Mechanical and Materials Engineering, Washington State University, Pullman, WA 99164 USA (e-mail: mehdi.hosseinzadeh@wsu.edu).

Bruno Sinopoli is with the Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: bsinopoli@wustl.edu).

Aaron F. Bobick is with the Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: afb@wustl.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCST.2023.3305100>.

Digital Object Identifier 10.1109/TCST.2023.3305100

The main features of the proposed scheme are: 1) it is general and can be applied to any human–robot interaction satisfying the posed setting; 2) it is modular, meaning that any other objective function or belief update rule can be incorporated into the scheme without changing its structure; 3) danger awareness coefficient is automatically and continuously updated by the robot as a result of the observed human’s actions, and therefore takes into account possible time-varying effects due to danger signaling; and 4) the developed robot action planning scheme takes safety into account, thus avoiding the design of often complex, rarely comprehensive, rule-based, ad hoc safety rules.

#### D. Organization

Section II discusses selected related work. Section III formulates the problem and introduces the notion of danger awareness. Section IV discusses how to build a predictive human model, learn from humans’ actions, and predict humans’ actions and states in the future. In Section V, a safe and efficient robot action planning is proposed. Section VI validates the proposed approach through intensive simulation studies on an interaction between a self-driven car and a pedestrian. Section VIII concludes the article and discusses future work.

#### E. Notation

We denote the set of real numbers, the set of positive real numbers, and the set of non-negative real numbers by  $\mathbb{R}$ ,  $\mathbb{R}_{>0}$ , and  $\mathbb{R}_{\geq 0}$ , respectively. We use  $\mathcal{N}(\mu, \Sigma)$  to indicate the Gaussian distribution with mean  $\mu$  and covariance  $\Sigma$ . We denote proportionality by  $\propto$ , and the transpose of matrix  $A$  by  $A^\top$ . For a given set  $X$ , we use  $|X|$  to denote its cardinality.

## II. RELATED WORK

### A. Predictive Human Model

In recent years, there have been several studies on predicting humans’ actions in the context of human–robot interaction. In some work (e.g., [9]), it is assumed that the robot has complete knowledge about the environment. However, this assumption may not be reasonable in real-world scenarios due to uncertainties in human’s behavior. As a result, many researchers have focused on developing a method to enable robots to use the history of humans’ actions to predict their future actions and states. In [10], propagation networks have been utilized to detect partially ordered sequential actions of the humans. Albanese et al. [11] introduced the concept of constrained probabilistic Petri nets and showed how this concept can be used to predict humans’ actions. In [12], Gaussian mixture distribution techniques have been used to model humans’ actions and predict their timing. Markov models have been used in a variety of studies [13], [14] to predict the timing of humans’ actions. In [15], an interaction primitive framework for predicting humans’ the most likely future movements is developed. The anticipatory temporal conditional random fields have been used in [16] to predict humans’ future actions. Some *ad hoc* methods (e.g., [17]) have also been proposed in the literature.

Extensive work in cognitive science has shown that human behavior can be well modeled by objective-driven optimization [18], [19], [20]. With that, a goal-based algorithm is proposed in [21] to predict pedestrians’ future trajectories. A Bayesian framework is provided in [22] and [23] to reason about humans’ rationality, and thus predict humans’ actions. Hawkins and Tsotras [24] assume that humans are rational and build a predictive model to anticipate the timing of their actions. In [25] and [26], online Bayesian method has been exploited to infer human’s latent states and generate a predictive model.

### B. Robot Action Planning

Once a predictive human model is developed, the robot can use this model to generate a safe and efficient plan. Several robot action planning schemes have been proposed in the literature. Wilcox et al. [27] introduce the adaptive preferences algorithm that computes a flexible optimal policy for robot scheduling and control in assembly manufacturing. In [28], a method has been proposed to optimize the task assignment such that the cycle time is shortened, and consequently the productivity is increased. Probabilistic wait-sensitive task planning have been proposed in [29] and [30] to optimize the robot tasks with respect to the posterior human action distributions, such that the human’s total wait time is reduced. Tanaka et al. [31] and Kanazawa et al. [32] propose a motion planning scheme based on human’s trajectory prediction to improve efficiency. Genetic algorithms have also been utilized in some robot action planners, e.g., [33]. The notion of the virtual plane is used in [34] for path planning and navigation in dynamic environments. Aoude et al. [35] have developed a path planning framework to safely navigate robots, while avoiding dynamic obstacles with uncertain motion patterns.

## III. PROBLEM FORMULATION

Consider a human–robot interaction in which one robot and one human are moving to different goal locations.

### A. Robot Model

The robot can be modeled as

$$x_R(t+1) = f_R(x_R(t), u_R(t)) \quad (1)$$

where  $x_R(t) \in \mathbb{R}^{n_R}$  and  $u_R(t) \in \mathcal{U}_R \subset \mathbb{R}^{m_R}$  are respectively the robot’s state and action at time  $t$ . Note that  $\mathcal{U}_R$  is the set of admissible robot actions, and  $n_R$  and  $m_R$  are the dimensions of robot’s state-space and control action, respectively.

Let  $g_R \in \mathbb{R}^{n_R}$  be the goal state of the robot. Analogous to [30] and [36], we assume that the robot uses a receding horizon control strategy [37] to reach the state  $g_R$ , while avoiding collisions with the human. Let  $Q_g^R(x_R(t), u_R(t), g_R) : \mathbb{R}^{n_R} \times \mathbb{R}^{m_R} \times \mathbb{R}^{n_R} \rightarrow \mathbb{R}$  be the robot’s objective function corresponding to the goal state  $g_R$ , defined as follows<sup>1</sup>:

$$Q_g^R(\cdot) = \theta_1 \|x_R(t) - g_R\|^2 + \theta_2 \|u_R(t)\|^2 \quad (2)$$

<sup>1</sup>For the sake of brevity, we denote  $Q_g^R(x_R(t), u_R(t), g_R)$  by  $Q_g^R(\cdot)$  whenever there is no confusion.

where  $\theta_1, \theta_2 \in \mathbb{R}_{>0}$  are weighting parameters. Given  $T_R \in \mathbb{R}_{>0}$  as the prediction horizon, the robot solves the following optimization problem to compute the optimal control actions over the prediction horizon:

$$u_R^*(t : t + T_R - 1) = \begin{cases} \arg \min \sum_{k=t}^{t+T_R-1} Q_g^R(\cdot) \\ \text{s.t. } u_R(k) \in \mathcal{U}_R \quad \forall k \\ \text{Model (1)} \\ \mathbb{P}_{\text{Coll}}(k) \leq P_{\text{th}} \quad \forall k \end{cases} \quad (3)$$

where  $k = t + 1, \dots, t + T_R$ ,  $u_R^*(t : t + T_R - 1) = [(u_R^*(t))^T \cdots (u_R^*(t + T_R - 1))^T]^T$ ,  $\mathbb{P}_{\text{Coll}}(k) \in [0, 1]$  is the probability of a collision between the human and robot at prediction time instant  $k$  (see Section IV-C), and  $P_{\text{th}} \in [0, 1]$  is a threshold value.

*Remark 1:* In general, the constraint  $\mathbb{P}_{\text{Coll}}(t + T_R) \leq P_{\text{th}}$  does not guarantee recursive feasibility. To the best of our knowledge, this is an issue in robot action planning schemes which are developed based upon probabilistic predictive human models (e.g., [22], [23], [38]). Indeed, since the robot actions are limited and due to imperfect human models, the recursive feasibility is unsurprisingly very difficult to satisfy. Future research will investigate how to ensure recursive feasibility without rendering the solution conservative.

## B. Human Model

The human can be modeled as

$$x_H(t + 1) = f_H(x_H(t), u_H(t)) \quad (4)$$

where  $x_H(t) \in \mathbb{R}^{n_H}$  and  $u_H(t) \in \mathcal{U}_H \subset \mathbb{R}^{m_H}$  are respectively the human's state and action. Note that  $\mathcal{U}_H$  is the set of admissible human actions,  $n_H$  is the dimension of human's state-space, and  $m_H$  is the dimension of human's action.

As discussed in [18], [20], and [22], human's actions can be modeled by objective-driven optimization. We can model the human's action as optimizing a combination of two additive objectives: 1) rewarding the achievement of a goal (will be referred as the goal objective function) and 2) being safe by avoiding collisions (will be referred as safety objective function). Given  $g_H \in \mathbb{R}^{n_H}$  as the human's goal state, we take inspiration from [22] and [23] to define the human's goal objective function as follows:

$$Q_g^H(x_H(t), u_H(t), g_H) = \theta_3 \|x_H(t) - g_H\|^2 + \theta_4 \|u_H(t)\|^2 \quad (5)$$

where  $\theta_3, \theta_4 \in \mathbb{R}_{>0}$  are design parameters. Regarding the safety objective function, we use the following function which has been experimentally validated in [39] and [40]:

$$Q_s^H(x_H(t), u_H(t), \hat{x}_R(t)) = \theta_5 e^{-\theta_6 \|x_H(t) - \hat{x}_R(t)\|^2} \quad (6)$$

where  $\theta_5, \theta_6 \in \mathbb{R}_{>0}$  are design parameters, and  $\hat{x}_R(t) \in \mathbb{R}^{n_R}$  is an estimation of the robot's state at time instant  $t$  computed by the human. Analogous to [41], we model the estimation as  $\hat{x}_R(t) = x_R(t) + \epsilon(t)$ , where  $\epsilon(t) \in \mathbb{R}^{n_R}$  is a zero-mean Gaussian random variable with covariance  $\Sigma \in \mathbb{R}_{\geq 0}^{n_R \times n_R}$ , i.e.,  $\epsilon(t) \sim \mathcal{N}(0, \Sigma)$ .

Given  $Q_g^H(\cdot)$  and  $Q_s^H(\cdot)$ , the human's action at time instant  $t$  is the solution of the following optimization problem<sup>2</sup>:

$$u_H^*(t) = \arg \min_{u_H \in \mathcal{U}_H} \eta_1 Q_g^H(\cdot) + \beta \eta_2 Q_s^H(\cdot) \quad (7)$$

where  $\beta$  is a binary variable (i.e.,  $\beta \in \{0, 1\}$ ) that we refer to as the *danger awareness coefficient*, and  $\eta_1, \eta_2 \in \mathbb{R}_{>0}$  are weighting parameters.

*Remark 2:* According to (7), we can interpret  $\eta_1$  and  $\eta_2$  as the weights we attach to the objective functions  $Q_g^H(\cdot)$  and  $Q_s^H(\cdot)$ , respectively. In particular, we can [42] interpret the ratio  $\eta_1/\eta_2$  as the relative weight or relative importance of the objective function  $Q_g^H(\cdot)$  compared to the objective function  $Q_s^H(\cdot)$ . Based on this insight, the weights  $\eta_1$  and  $\eta_2$  can be determined by using techniques described in, e.g., [43], [44], and [45].

A common interpretation of  $\beta = 0$  is a human who does not see the robot or for some reasons is ignoring the dangers the robot may present (one possible such reason is that human presumes that it is the robot's responsibility to keep a safe distance). Whereas,  $\beta = 1$  means that the human is aware of the danger and acts properly to reduce the risk. In the rest of the article, we will use the terms *concerned* and *unconcerned* to refer to a human with  $\beta = 1$  and  $\beta = 0$ , respectively. Note that since  $\beta$  multiplies  $Q_s^H(\cdot)$  in (7), either  $\beta = 0$  (i.e., unaware humans) or  $Q_s^H(\cdot) = 0$  (e.g., children who do not recognize the danger) will be referred as unconcerned humans. We will also use the term *initially unaware* to refer to a human initially with  $\beta = 0$  but changed to  $\beta = 1$  as a result of danger signaling (see Section III-C). See Appendix A for a game-theoretic interpretation of the danger awareness coefficient.

*Remark 3:* We assume that the robot knows the objective functions  $Q_g^H(\cdot)$  and  $Q_s^H(\cdot)$ ; this assumption is pretty standard<sup>3</sup> in robotics literature (see, for instance, [22], [23], [46], [47]). Indeed, the robot can learn these objective functions (e.g., by using inverse reinforcement learning [48], [49], [50]) or these functions can be explicitly provided by the system designers (e.g., based upon behavioral patterns [51], [52], [53], [54]).

## C. Danger Signaling System

The robot employs a pre-collision method which uses signals/indicators to alert the danger to the human. This will be referred as the danger signaling system. From a technical viewpoint, the danger signaling creates a communication channel between the agents [55], [56], which the robot utilizes to actively perturb the environment so to improve the efficiency and safety of the interaction. The planner does so by: 1) acquainting an unaware human or a human who underestimates the danger and 2) helping the human to reduce the estimation error  $\epsilon(t)$ .

We denote the on/off status of the danger signaling by the binary variable  $d_R$ , where  $d_R = 0$  if the signaling is off and  $d_R = 1$  if it is on. The robot switches the signaling

<sup>2</sup>For the sake of brevity, we denote  $Q_g^H(x_H(t), u_H(t), g_H)$  and  $Q_s^H(x_H(t), u_H(t), \hat{x}_R(t))$  by  $Q_g^H(\cdot)$  and  $Q_s^H(\cdot)$ , respectively, whenever there is no confusion.

<sup>3</sup>We understand that there might be some differences between people. However, this article does not aim to deal with the differences.

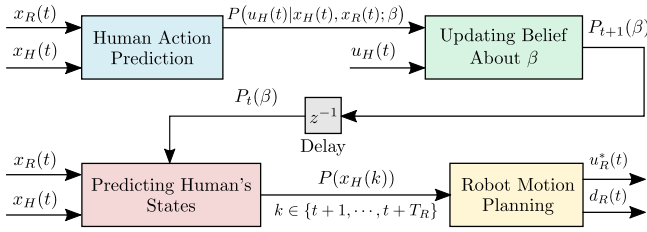


Fig. 1. General structure of the proposed planning scheme.

on if the last constraint in (3) is active (i.e., it affects the solution of the optimization problem). Note that the danger signal  $d_R(t)$  can change the danger awareness coefficient  $\beta$  by generating awareness in the human, and consequently impact human's behavior so as to reduce the probability of collision  $\mathbb{P}_{\text{Coll}}(k)$ ,  $\forall k$ ; hence, the danger signal  $d_R(t)$  impacts the robot's decision by impacting the last constraint in (3).

#### D. Problem Statement

This article considers the following problem.

**Problem 1:** Suppose one robot and one human are moving to different goal locations. Suppose that robot's model is as in (1), and the robot solves (3) to determine its next action. Suppose that the human's model is as in (4), and the human decides their next action via (7). Suppose that the robot uses the danger signaling system to alert the danger to the human. Design a robot action planning scheme to provide a plan that ensures both agents reach their goal states efficiently, while guaranteeing the safety of agents.

To solve Problem 1, assuming that the robot can observe the human's states and actions, we will develop a robot action planning scheme to provide safe robot actions to guide the robot to the goal state, without colliding with the human and rendering the solution conservative. The general structure of the proposed scheme is depicted in Fig. 1. The gist of this scheme is the development of a human predictive model whose values are computed through posterior calculations performed by the robot. This robot action planning scheme will be discussed in detail in Sections IV and V.

### IV. PREDICTIVE HUMAN MODEL

According to the optimization problem (3), the robot determines its actions by taking into account the human's future states. This section discusses how one can build a predictive human model to anticipate human's states in the future.

#### A. Human Action Prediction

The robot uses the following mixture distribution to model the human's behavior:

$$\mathbb{P}(u_H|x_H, x_R; \beta) = (1 - \omega_H) \cdot \mathbb{P}_d(u_H|x_H, x_R; \beta) + \omega_H \cdot \mathbb{P}_r(u_H) \quad (8)$$

where  $\mathbb{P}_d(u_H|x_H, x_R; \beta)$  models the human's deliberate behavior,  $\mathbb{P}_r(u_H)$  models the human's random behavior, and  $\omega_H \in [0, 1]$  is the mixture weight. Note that (8) should be computed for every  $u_H \in \mathcal{U}_H$  and  $\beta \in \{0, 1\}$ .

**1) Human's Deliberate Behavior:** The function  $\mathbb{P}_d(u_H|x_H, x_R; \beta)$  describes the probability distribution of the human's action if the human chooses the action according to the goal and safety objective functions as in (7). Assuming that the robot can observe human's states, analogous to [23], the robot can use the Boltzmann distribution to model the human's deliberate behavior, as

$$\mathbb{P}_d(u_H|x_H, x_R; \beta) \propto e^{-\gamma(Q_s^H(\cdot) + \beta Q_s^H(x_H(t), u_H(t), x_R(t)))} \quad (9)$$

where  $\gamma \gg 1$  is the temperature of the Boltzmann distribution.

**Remark 4:** We assume that the robot is equipped with high-precision sensors and employs a decent state estimator [57], [58] to accurately extract the human's state from possibly noisy sensor measurements. Thus, the robot uses the human's actual state  $x_H(t)$  in (9).

**Remark 5:** The robot uses its actual state  $x_R(t)$  to predict the human's actions as in (9), while the human uses an estimate of that (i.e.,  $\hat{x}_R(t)$ ) to decide the next action in (7). We will study the impact of the estimation error in Section VI-B.

**2) Human's Random Behavior:** In reality, a human may choose a random action by completely ignoring the objective functions for any reason. The robot makes use of a uniform distribution to model the human's random behavior, as follows:

$$\mathbb{P}_r(u_H) = \frac{1}{|\mathcal{U}_H|} \quad \forall u_H \in \mathcal{U}_H \quad (10)$$

where  $|\mathcal{U}_H|$  is the cardinality of the set  $\mathcal{U}_H$ .

#### B. Robot's Belief About the Human's Danger Awareness

Let  $\mathbb{P}_t(\beta = 1)$  be the robot's belief at time instant  $t$  about the likelihood that the human is concerned, where  $\mathbb{P}_0(\beta = 1)$  indicates the robot's prior belief. By observing the human's state  $x_H(t)$  and action  $u_H(t)$ , the robot updates its belief via the following Bayesian update rule:

$$\mathbb{P}_{t+1}(\beta = 1) = \frac{\mathbb{P}(u_H(t)|x_H(t), x_R(t); \beta = 1)\mathbb{P}_t(\beta = 1)}{\sum_{\beta} \mathbb{P}(u_H(t)|x_H(t), x_R(t); \beta)\mathbb{P}_t(\beta)} \quad (11)$$

where  $\mathbb{P}(u_H(t)|x_H(t); \beta = 1)$  can be computed via (8).

#### C. Computation of the Probability of Collision $\mathbb{P}_{\text{Coll}}(k)$

**1) Human's State in the Future:** Analogous to [22], we divide the human's state-space into  $N_c$  grid cells. Thus, the probability distribution of the human's states in the time interval  $[t + 1, t + T_R]$  can be predicted via the following recursive rule:

$$\mathbb{P}(x_H(k + 1)) \propto \sum_{x_H(k), u_H(k), \beta} \mathbb{P}(x_H(k + 1)|x_H(k), u_H(k)) \cdot \mathbb{P}(u_H(k)|x_H(k), x_R(k); \beta) \cdot \mathbb{P}_t(\beta) \quad (12)$$

where  $k = t, \dots, t + T_R - 1$ ,  $\mathbb{P}(u_H(k)|x_H(k), x_R(k); \beta)$  is as in (8),  $\mathbb{P}_t(\beta)$  is as in (11), and  $\mathbb{P}(x_H(k + 1)|x_H(k), u_H(k))$  is equal to 1 if  $x_H(k + 1)$ ,  $x_H(k)$ , and  $u_H(k)$  satisfy (4), and is equal to zero otherwise.



2) *Probability of Collision*: Let  $\lambda(x_H(k))$  be a neighborhood around the predicted human's state  $x_H(k)$ . This neighborhood should be determined according to the predefined safe distancing measures, the effect of possible modeling and tracking errors on predictions, and quantization error in gridding the human's state-space. Then, the probability of a collision event at prediction time instant  $k$  can be computed as the probability that  $x_R(k)$  is inside the neighborhood  $\lambda(x_H(k))$ , without any collisions prior to  $k$ . This probability is given in (13), as shown at the bottom of the page, which should be computed recursively, where  $\mathbb{P}(x_R(k) \in \lambda(\alpha)) = \mathbb{P}(x_H(k) = \alpha)$  with  $\alpha$  as a cell in the human's state-space, which can be computed via (12). It should be remarked that the upper-bound in (13) indicates that collision probabilities over time can be treated independently. Note that using the upper-bound given in (13) can significantly reduce the computational complexity of optimization problem (3), although it may lead to a conservative solution; this upper-bound has been used in simulation studies presented in Section VI.

## V. ROBOT ACTION PLANNING SCHEME

The robot action planning algorithm is presented in Algorithm 1. Note that since any theoretical guarantee is tied to the model it is based on, Algorithm 1 inherits the probabilistic nature of human model, and thus can only provide probabilistic safety guarantees; see Appendix B for more details. To the best of our knowledge, this is an issue in robot action planning schemes which are developed based upon probabilistic predictive human models (e.g., [22], [38], [47]).

### Algorithm 1 Robot Action Planning Scheme

**Input:** Humans' states  $x_H(t)$  and robot's state  $x_R(t)$ .

**Output:** Robot's action  $u_R^*(t)$  and belief  $\mathbb{P}_t(\beta)$ .

- 1: Compute the probability distribution  $\mathbb{P}(u_H|x_H(t), x_R(t); \beta)$  for every  $u_H \in \mathcal{U}_H$  and  $\beta \in \{0, 1\}$  as in (8).
- 2: Compute the probability distribution over human's states  $\mathbb{P}(x_H(k))$  for  $k = t+1, \dots, t+T_R$  as in (12).
- 3: Compute the probability of collision  $\mathbb{P}_{Coll}[k]$  for  $k = t+1, \dots, t+T_R$  as in (13).
- 4: Determine the robot's action  $u_R^*(t)$  by solving (3) and determine the on/off status of the danger signaling  $d_R$ .
- 5: Observe the human's action  $u_H(t)$ .
- 6: Update the belief about the human's danger awareness  $\mathbb{P}_{t+1}(\beta = 1)$  as in (11).

*Remark 6:* Note that Algorithm 1 might be computationally expensive if applied to high-order systems. In general, there are two mutually non exclusive approaches to decrease the computational complexity of Algorithm 1: 1) using a small set  $\mathcal{U}_H$  to categorize the human's actions and 2) considering relatively large grid cells in human's state-space.

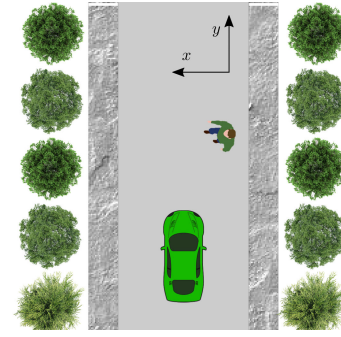


Fig. 2. Considered interaction between a self-driven car and a pedestrian.

## VI. SIMULATION STUDY—SELF-DRIVEN CAR-PEDESTRIAN INTERACTION

In this section, we investigate the effectiveness of the robot action planning scheme given in Algorithm 1 by simulating and testing it on an interaction between a self-driven car and a pedestrian, which is shown in Fig. 2. Our focus in this section is on the algorithmic and analytical aspects of the proposed framework on a simplistic, but yet realistic, scenario.

We assume that the self-driven car can only move vertically (i.e., it can only move on  $y$ -axis) to a goal position  $g_R \in \mathbb{R}$ . Note that this assumption is reasonable in the case of a narrow one-way street. We use the Ackermann steering kinematics [59], [60], [61], [62] to model the self-driven car, as:  $x_R(t+1) = x_R(t) + u_R(t)\Delta T$ , where  $x_R(t) \in \mathbb{R}$  and  $u_R(t) \in \mathbb{R}$  are the robot's vertical position and its directional velocity at time instant  $t$ , respectively, and  $\Delta T$  is the time step size. Let  $\mathcal{U}_R = \{0, v_R/2, v_R\}$ , where  $v_R$  is the basic velocity of the self-driven car, which means that the self-driven car can move at either zero speed, or half speed, or full speed.

The pedestrian is an adult who can only move horizontally (i.e., they can only move on  $x$ -axis) toward the goal position  $g_H \in \mathbb{R}$  (which is a point on the left side of the street). The pedestrian's model is then  $x_H(t+1) = x_H(t) + u_H(t)\Delta T$ , where  $x_H(t) \in \mathbb{R}$  and  $u_H(t) \in \mathbb{R}$  are the human's horizontal position and their directional velocity at time instant  $t$ , respectively, and  $\Delta T$  is the time step size. Let  $\mathcal{U}_H = \{-2v_H, -v_H, 0, v_H, 2v_H\}$ , where  $v_H$  is the pedestrian's basic walking velocity, which means that the pedestrian can stop, walk, or run in either directions.

We assume that  $g_R = 80$ ,  $g_H = 5$ ,  $v_R = 2$ ,  $v_H = 0.5$ ,  $\omega_H = 0.5$ ,  $P_{th} = 0.1$ ,  $T_R = 5$ ,  $\gamma = 1000$ ,  $\eta_1 = \eta_2 = 1$ ,  $\theta_1 = 1$ ,  $\theta_2 = 0.5$ ,  $\theta_3 = 2.5$ ,  $\theta_4 = 8 \times 10^{-3}$ ,  $\theta_5 = 300$ ,  $\theta_6 = 6 \times 10^{-3}$ ,  $\Sigma = 1$ , and  $\mathbb{P}_0(\beta = 1) = 0.5$ . We use the robot action planning scheme given in Algorithm 1 to manage the above-mentioned interaction. We define  $\lambda(x_H(k))$  as a circle centered on  $x_H(k)$  with radius  $\rho = 2$ , and use the YALMIP toolbox [63] to solve the corresponding optimization problems. In order to have a visual demonstration of the considered interaction, a simulator

$$\begin{aligned} \mathbb{P}_{Coll}(k) &= \mathbb{P}(x_R(k) \in \lambda(x_H(k)), x_R(k-1) \notin \lambda(x_H(k-1)), \dots, x_R(t+1) \notin \lambda(x_H(t+1))) \\ &\leq \mathbb{P}\left(x_R(k) \in \lambda(x_H(k)) \middle| x_R(k-1) \notin \lambda(x_H(k-1)), \dots, x_R(t+1) \notin \lambda(x_H(t+1))\right). \end{aligned} \quad (13)$$

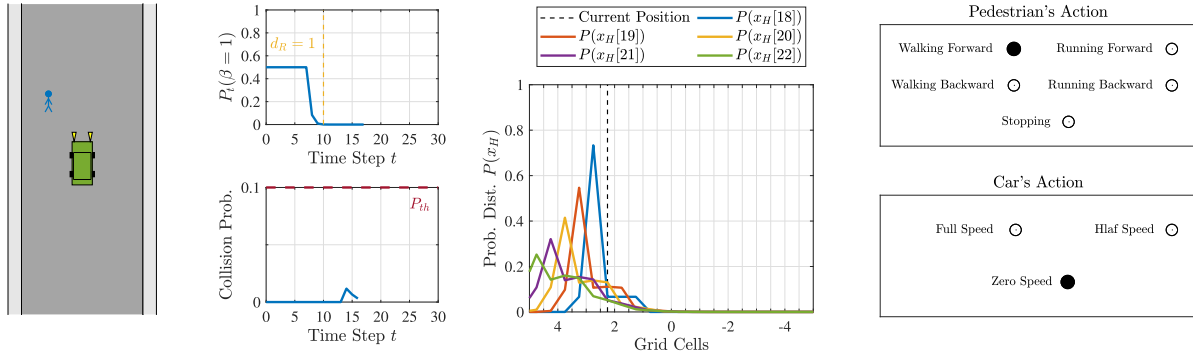


Fig. 3. Screenshot of the generated simulator shown in the accompanied video ([https://youtu.be/\\_9UjDvZYT2U](https://youtu.be/_9UjDvZYT2U)). Left figure: the interaction between a self-driven car and a pedestrian in the street; the pedestrian moves from right to left, and the car moves from south to north. Middle-left figures: the top figure shows the time profile of the robot's belief about the likelihood that the human is aware of the danger, i.e.,  $\mathbb{P}_t(\beta = 1)$ , and the bottom figure shows the probability of collision over the prediction horizon, computed at each time instant (here at time  $t = 17$ ); note that the pedestrian moves from right to left. Right figures: the top and bottom figures indicate the pedestrian and car actions at the current time instant, respectively.

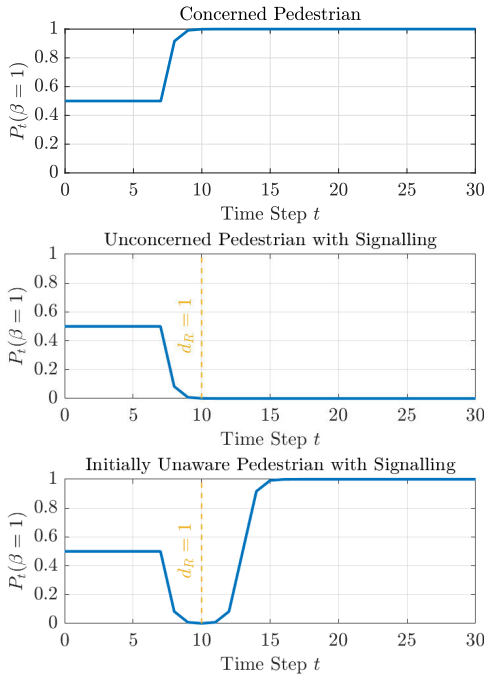


Fig. 4. Impact of the danger signaling on the robot's belief  $\mathbb{P}_t(\beta = 1)$ .

has been generated. A video of operation of the simulator is available at [https://youtu.be/\\_9UjDvZYT2U](https://youtu.be/_9UjDvZYT2U). Fig. 3 presents a snapshot of the generated simulator.

#### A. Impact of the Danger Signaling System

As discussed in Section III-C, the danger signaling system can improve the efficiency and safety by acquainting an unaware pedestrian, which is shown in Fig. 4. As seen in this figure, when the pedestrian is unconcerned (i.e.,  $\beta = 0$  and/or  $Q_s^H(\cdot) = 0$ ), the pedestrian keeps walking toward the goal position  $g_H$ . Thus, the self-driven car comes to a full stop to keep the probability of collision lower than the threshold value. Whereas, when the robot alerts the danger to an initially unaware pedestrian, they take a decent action to keep a safe distance. Thus, the car continues toward  $g_R$  without stopping.

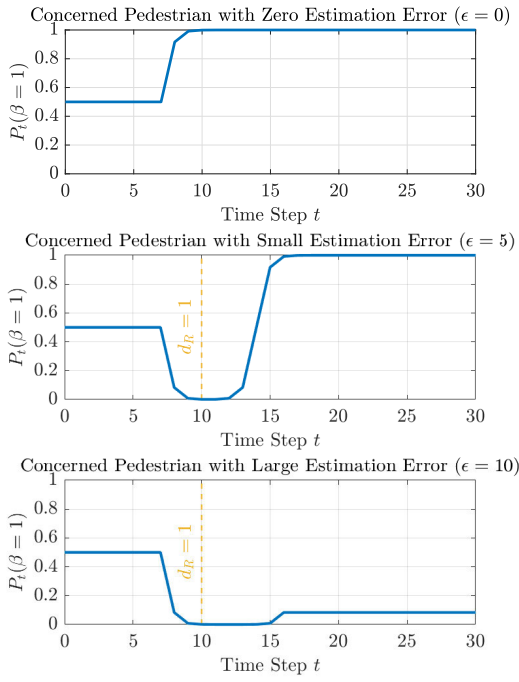
TABLE I

COMPARISON STUDY—NUMBERS OF FULL STOPS AND SLOWDOWNS FOR THE PROPOSED ROBOT ACTION PLANNING SCHEME WITH AND WITHOUT DANGER SIGNALING

	Without Signaling	With Signaling
Full Stops	469	41
Slowdowns	351	94

To provide a quantitative study to show the effectiveness of danger signaling, we consider a population of 1000 pedestrians, with 50% concerned, 25% unconcerned, and 25% initially unaware pedestrians, where the initial distance between the pedestrian and the self-driven car for each experiments is uniformly selected from the interval  $[70, 90]$ . Such a population represents a realistic set of pedestrians in a real street-crossing scenario. The numbers of full stops and slowdowns for the proposed robot action planning scheme with and without danger signaling are reported in Table I. As seen in this table, danger signaling can reduce the number of slowdowns and the number of full stops by  $\sim 90\%$  and  $\sim 70\%$ , respectively.

*Remark 7:* Suppose that the danger signal is on (i.e.,  $d_R = 1$ ), but the pedestrian misinterprets it as off. In this case, the pedestrian will act as an unconcerned pedestrian, and thus, the robot will identify the pedestrian as an unconcerned human and will determine a plan accordingly. Now, suppose that the danger signal is off (i.e.,  $d_R = 0$ ), but the pedestrian misinterprets it as on. In this case, the pedestrian assumes that there is a danger while there is none, and consequently will return/stay on the initial position (sidewalk on the right) or will run toward the goal position (sidewalk on the left). Hence, although the pedestrian's actions do not increase the risk, the robot will identify the pedestrian as an unconcerned human because their actions do not follow the objective-driven optimization given in (7). However, as the self-driven car moves toward the pedestrian, the pedestrian's actions will gradually match the the objective-driven optimization given in (7), and thus the robot will eventually identify the pedestrian as a concerned human.

Fig. 5. Impact of the estimation error on the robot's belief  $\mathbb{P}_t(\beta = 1)$ .

### B. Impact of the Estimation Error $\epsilon(t)$

As discussed in Section III-B, the pedestrian makes use of an estimation of the self-driven car's states. However, as discussed in Section IV-A, the self-driven car predicts the pedestrian's states based on its actual position. This means that in the presence of a large estimation error, even if the pedestrian is concerned (i.e.,  $\beta = 1$ ), the pedestrian may take an action which increases the risk. As a result, the self-driven car cannot learn accurately the pedestrian's danger awareness (see Remark 5), and may have unnecessary stops.

As discussed in Section III-C, the danger signaling system can impact the interaction by decreasing the estimation error. This impact has been studied in Fig. 5. We assume that the pedestrian compensates the estimation error as  $\epsilon(t) = \epsilon_0 \cdot e^{-\epsilon_1 \cdot (t-t_d)}$ , where  $\epsilon_0 \in \mathbb{R}$  is the initial error,  $\epsilon_1 \in \mathbb{R}_{>0}$  is the compensation rate, and  $t_d$  is the time instant that  $d_R$  switches from 0 to 1. Note that a deterministic error (i.e.,  $\Sigma = 0$ ) is considered in this set of simulations, as the main goal here is to understand the impact of estimation error on the performance of the proposed robot action planning scheme.

As seen in Fig. 5, even though the pedestrian is concerned, they react late due to the estimation error. When the estimation error is small (i.e.,  $\epsilon(t) = 5$ ), the pedestrian runs backward toward the sidewalk on the right, as they realize the danger when they are on the right half of the street (note that the pedestrian walks from right to left). While, when the estimation error is large (i.e.,  $\epsilon(t) = 10$ ), the pedestrian runs forward toward the sidewalk on the left, as they realize the danger when they are on the left half of the street.

### C. Impact of the Mixture Weight $\omega_H$

As discussed in Section IV-A, the mixture weight  $\omega_H$  defines the relationship between the deliberate and random components of the human's behavior. More precisely,  $\omega_H = 0$

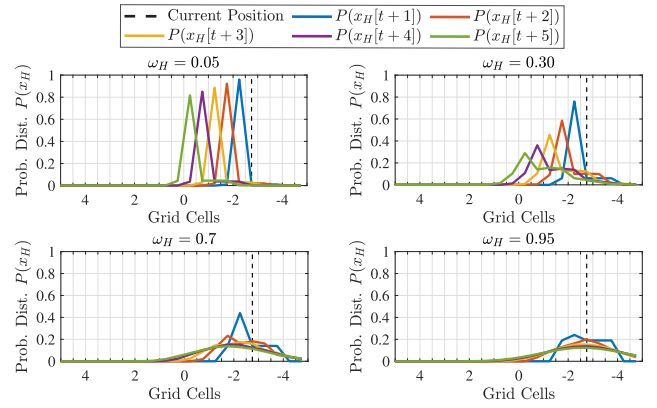
Fig. 6. Impact of the mixture weight  $\omega_H$  on the probability distribution over the human's future positions, computed at time  $t$ . Note that the pedestrian moves from right to left.

TABLE II

COMPARISON STUDY—NORMALIZED NUMBERS OF FULL STOPS AND SLOWDOWNS FOR THE PROPOSED ROBOT ACTION PLANNING SCHEME AND THE SCHEME OF [22]

	Scheme of [22]	Proposed Scheme
Full Stops (Norm.)	122	40
Slowdowns (Norm.)	181	97

means that the human is being driven only by the goal and safety objective functions as in (7), and  $\omega_H = 1$  means that the human chooses actions randomly by completely ignoring the objectives functions. This mixture weight affects the prediction of the pedestrian's future states.

For a large  $\omega_H$ , the pedestrian appears more random. Thus, the probability distribution over the pedestrian's state in the future will be wide; this means that the prediction of the pedestrian's future positions is uncertain. A small  $\omega_H$  leads to a tighter distribution, meaning that the prediction of the pedestrian's future positions becomes less uncertain. This impact is shown in Fig. 6 for four different values of  $\omega_H$ .

### D. Comparison Study

Consider a population of 1000 pedestrians, with 50% concerned, 25% unconcerned, and 25% initially unaware pedestrians, who start from uniformly distributed initial positions.

For comparison purposes, we simulate the planning scheme presented in [22]. Fisac et al. [22], first, introduces the notion of "model confidence" (which can be viewed as a time-varying indicator of the performance of the predictive human model) and designs a framework for reasoning about it, and then develops a robotic motion planner that incorporates the notion into planning. We choose Fisac et al. [22] for comparison as it addresses our problem setting; that is Fisac et al. [22] investigates the issue of uncertainty inherent in predictions of human movements that are computed by objective-driven optimization and develops a probabilistically safe plan for the robot. Numbers of full stops and slowdowns are reported in Table II. As seen in this table, compared with the scheme presented in [22], the robot action planning scheme in Algorithm 1 reduces the number of slowdowns and the number of full stops by  $\sim 50\%$  and  $\sim 70\%$ , respectively.

## VII. EXTENSION TO MULTI-HUMAN–MULTI-ROBOT INTERACTION

Consider a human–robot interaction, in which multiple robots and multiple humans are moving to different goal locations. In this case, robots should employ probabilistic predictive models to produce a distribution of states the humans may occupy in the future. However, generating accurate real-time predictions for multiple humans is an open problem [23]. The main challenge is the complexity of modeling interactive effects as the number of agents increases, such that any simplifying assumptions could result in prediction inaccuracy and may even threaten safety of agents.

The predictive human model presented in Section IV adapts prediction uncertainty online to reflect the degree to which humans' actions match an internal model. This property allows us to simplify interactive effects between agents without threatening safety, as uncertain/erroneous predictions will result in more conservative, but safe, plans. More precisely, the Bayesian update given in (11) will automatically generate more conservative predictions when humans' actions deviates from the structure given in (7).

Based upon the above-mentioned property, we will extend Algorithm 1 for the case of multi-human-multi-robot interaction. Two approaches will be pursued. In the first approach, first, each robot observes states of humans and updates its belief about their danger awareness. Then, robots use the predictive human model presented in Section IV to generate a time-indexed set of occupancy grids, and compute the probability of collision over the prediction horizon. Finally, under the reasonable assumption that robots can communicate with each other [64], [65], [66], robots will employ the sequential planning method [67], [68] to compute the optimal trajectory.

In the second approach, each robot will consider only the biggest problem at every time (i.e., the human whose probability of collision has the largest value); it is foreseen that the second approach will have a low computational cost.

## VIII. CONCLUSION AND FUTURE WORK

This article introduced the notion of danger awareness as a way of improving safety and efficiency in human–robot interaction. The danger awareness coefficient encodes whether the human sees the robot and quantifies the their intention about participating in ensuring safety. This article developed an online Bayesian method to learn the human's danger awareness by observing their behavior. This enables the robot to build a predictive human model and to predict the human's future actions and states. This article also proposed the danger signaling system as a way of generating awareness in the human. The combination of the human's danger awareness and the danger signaling contributes to the state-of-the-art by revoking the fully transfer of safety responsibilities to robots, while humans are eager to contribute in ensuring safety (note that this improves safety and efficiency in human–robot interaction). Finally, a planning scheme was proposed to provide probabilistically safe, but yet efficient, plans for robots. The proposed planning scheme was verified through intensive simulation studies on an interaction between a self-driven car and a pedestrian, and the impact of different parameters

were assessed. Future work should investigate validation and assessment of the workflow for multi-human-multi-robot interaction presented in Section VII.

## APPENDIX A

### GAME-THEORETIC INTERPRETATION OF THE DANGER AWARENESS COEFFICIENT

Extensive work in cognitive science (e.g., [69], [70]) has shown that humans exploit a hierarchical reasoning to predict others behavior/actions in decision-making processes. In other words, humans make decisions based on finite depth of reasoning. The Level- $k$  game theory [71], [72], [73] provides a very useful framework to mathematically model the above-mentioned reasoning hierarchy, where the level of an agent (also known as player) represents that agent's depth of reasoning. In this framework, an agent base its decisions on predictions about the likely actions of other agents, by assuming that they are less sophisticated (i.e., they consider a lower depth of reasoning).

The notion of danger awareness introduced in Section III-B can be interpreted based on the Level- $k$  game theory

- 1) *Level-0 Agent*: A level-0 agent that does not consider probable actions and/or reactions of other agents. This behavior resembles an unconcerned human discussed in Section III-B. In this case, the human's control policy, denoted by  $\pi_0$ , can be defined as a map from a pair of the human's state  $x_H(t)$  and the estimation of the robot's state  $\hat{x}_R(t)$  to the human's action  $u_H(t)$  when  $\beta = 0$ , i.e.,

$$\pi_0: (x_H(t), \hat{x}_R(t)) \mapsto u_H(t)|_{\beta=0} \quad (\text{A-1})$$

where  $u_H(t)|_{\beta=0}$  is determined by solving the optimization problem (7) and letting  $\beta = 0$ .

- 2) *Level-1 Agent*: A level-1 agent assumes that all other agents are level-0. Such an agent tries to provide the best response to the level-0 agents, which resembles a concerned human discussed in Section III-B. Thus, the human's control policy, denoted by  $\pi_1$ , can be defined via the following map:

$$\pi_1: (x_H(t), \hat{x}_R(t)) \mapsto u_H(t)|_{\beta=1} \quad (\text{A-2})$$

where  $u_H(t)|_{\beta=1}$  is determined by solving the optimization problem (7) and letting  $\beta = 1$ .

- 3) *Level-2 Agent*: A level-2 agent assumes that all other agents are level-1. Such an agent will behave as an unconcerned human, as it knows that all other agents are concerned, and thus it will selfishly focus on its objectives. Thus, the human's control policy, denoted by  $\pi_2$ , can be defined via the following map:

$$\pi_2: (x_H(t), \hat{x}_R(t)) \mapsto u_H(t)|_{\beta=0} \quad (\text{A-3})$$

where  $u_H(t)|_{\beta=0}$  is determined by solving the optimization problem (7) and letting  $\beta = 0$ . From (A-1) and (A-3), we have  $\pi_0 = \pi_2$ .

- 4) *Level-3 and Higher Agent*: A level-3 agent assumes that all other agents are level-2. Thus, a level-3 agent will behave as a concerned human. Therefore, the human's



control policy, denoted by  $\pi_3$ , can be defined via the following map:

$$\pi_3: (x_H(t), \hat{x}_R(t)) \mapsto u_H(t)|_{\beta=1} \quad (\text{A-4})$$

where  $u_H(t)|_{\beta=1}$  is determined by solving the optimization problem (7) and letting  $\beta = 1$ . From (A-2) and (A-4), we have  $\pi_1 = \pi_3$ . Note that higher level of agents can be modeled using the same logic.

According to (A-1)–(A-4), we observe that: 1) the human's control policy in different situations can be described by means of a binary variable and 2) level-3 and higher humans do not introduce a control policy that is different from lower humans (this observation has been reported in prior work [70], [74]).

## APPENDIX B

### CONNECTION TO FORWARD REACHABILITY SET

Although it will not constitute a formal safety guarantee, this appendix analyzes the safety properties of Algorithm 1 by examining how predicted state distributions over time relate to forward reachable sets. First, we define the forward reachable set in the following.

**Definition 1:** For the dynamical model 4, the forward reachable set  $\mathcal{F}(x_H(t_0), t + t_0)$  of a state  $x_H(t_0)$  after time  $t$  has elapsed is

$$\begin{aligned} \mathcal{F}(x_H(t_0), t + t_0) \\ = \{x'_H | \exists u_H(t_0), u_H(t_0 + 1), \dots, \\ u_H(t + t_0 - 1) \text{ such that } x_H(t + t_0) = x'_H\}. \end{aligned} \quad (\text{A-5})$$

□

According to Definition 1, to formally prove safety, it would suffice to ensure that the robot never comes too close to the human's forward reachable set, i.e.,  $x_R(t + t_0) \cap \mathcal{F}(x_H(t_0), t + t_0) = \emptyset$  for every  $x_R(t + t_0)$  along a motion plan generated when the human was at state  $x_H(t_0)$ . This condition is very restrictive as it requires avoiding the full forward reachable set. This induces a fundamental tradeoff [22] between safety and *liveness*. In the submitted manuscript, the threshold probability  $P_{th} \in [0, 1]$  is a design parameter that defines the tradeoff between safety and liveness. To understand this, first, we define the probabilistic forward reachable set.

**Definition 2:** For the dynamical model (4) with human's behavior given in (8), the forward reachable set with probability  $P_{th}$  of a state  $x_H(t_0)$  after time  $t$  has elapsed is

$$\begin{aligned} \mathcal{PF}(x_H(t_0), t + t_0, P_{th}) \\ = \{x'_H | \exists u_H(t_0), \dots, \\ u_H(t + t_0 - 1) \text{ s.t. } \mathbb{P}(x_H(t + t_0) = x'_H) \geq P_{th}\}. \end{aligned} \quad (\text{A-6})$$

□

As shown in [47] and [75], we have  $\mathcal{PF}(x_H(t_0), t + t_0, P_{th} = 1) \subseteq \mathcal{PF}(x_H(t_0), t + t_0, P_{th} = 0.9) \subseteq \dots \subseteq \mathcal{PF}(x_H(t_0), t + t_0, P_{th} = 0.1) \subseteq \mathcal{PF}(x_H(t_0), t + t_0, P_{th} = 0) = \mathcal{F}(x_H(t_0), t + t_0)$ , which implies that the smaller the threshold probability  $P_{th}$  is, the more closely our approach satisfies the sufficient condition for safety. Based on this argument, Algorithm 1 will determines a robot action plan when it is

predicted to be *sufficiently safe* with respect to the threshold probability  $P_{th}$ .

## REFERENCES

- [1] C. Brooks and D. Szafir, "Balanced information gathering and goal-oriented actions in shared autonomy," in *Proc. 14th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2019, pp. 85–94.
- [2] M. El-Shamouty, X. Wu, S. Yang, M. Albus, and M. F. Huber, "Towards safe human-robot collaboration using deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 4899–4905.
- [3] R. Peddi, C. D. Franco, S. Gao, and N. Bezzo, "A data-driven framework for proactive intention-aware motion planning of a robot in a human environment," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 5738–5744.
- [4] C. I. Mavrogiannis, V. Blukis, and R. A. Knepper, "Socially competent navigation planning by deep learning of multi-agent path topologies," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 6817–6824.
- [5] C. I. Mavrogiannis, W. B. Thomason, and R. A. Knepper, "Social momentum: A framework for legible navigation in dynamic multi-agent environments," in *Proc. 13th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2018, pp. 361–369.
- [6] J. Roh, C. Mavrogiannis, R. Madan, D. Fox, and S. S. Srinivasa, "Multimodal trajectory prediction via topological invariance for navigation at uncontrolled intersections," 2020, *arXiv:2011.03894*.
- [7] P. Trautman, J. Ma, R. M. Murray, and A. Krause, "Robot navigation in dense human crowds: Statistical models and experimental studies of human-robot cooperation," *Int. J. Robot. Res.*, vol. 34, no. 3, pp. 335–356, Mar. 2015.
- [8] S. Nikolaidis, D. Hsu, and S. Srinivasa, "Human-robot mutual adaptation in collaborative tasks: Models and experiments," *Int. J. Robot. Res.*, vol. 36, nos. 5–7, pp. 618–634, Jun. 2017.
- [9] M. Huber, A. Knoll, T. Brandt, and S. Glasauer, "When to assist—Modelling human behaviour for hybrid assembly systems," in *Proc. ISR (41st Int. Symp. Robotics) ROBOTIK (6th German Conf. Robotics)*, Jun. 2010, pp. 1–6.
- [10] Y. Shi, Y. Huang, D. Minnen, A. Bobick, and I. Essa, "Propagation networks for recognition of partially ordered sequential action," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2004, pp. 862–869.
- [11] M. Albanese et al., "A constrained probabilistic Petri net framework for human activity detection in video," *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 982–996, Oct. 2008.
- [12] J. Kinugawa, A. Kanazawa, S. Arai, and K. Kosuge, "Adaptive task scheduling for an assembly task coworker robot based on incremental learning of human's motion patterns," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 856–863, Apr. 2017.
- [13] D. Vasquez, T. Fraichard, and C. Laugier, "Incremental learning of statistical motion patterns with growing hidden Markov models," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 403–416, Sep. 2009.
- [14] B. T. Morris and M. M. Trivedi, "Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2287–2301, Nov. 2011.
- [15] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Jun. 2014, pp. 2831–2837.
- [16] H. S. Koppula and A. Saxena, "Anticipating human activities for reactive robotic response," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, p. 2071.
- [17] K. Li, J. Hu, and Y. Fu, "Modeling complex temporal composition of actionlets for activity prediction," in *Proc. Eur. Conf. Comput. Vis.*, Florence, Italy, Oct. 2012, pp. 286–299.
- [18] C. L. Baker, J. B. Tenenbaum, and R. R. Saxe, "Goal inference as inverse planning," in *Proc. Annu. Meeting Cognit. Sci. Soc.*, vol. 29, 2007, pp. 779–784.
- [19] J. V. Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 2007.
- [20] R. D. Luce, *Individual Choice Behavior: A Theoretical Analysis*. Chelmsford, MA, USA: Courier Corporation, 2012.
- [21] B. D. Ziebart et al., "Planning-based prediction for pedestrians," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 3931–3936.
- [22] J. Fisac et al., "Probabilistically safe robot planning with confidence-based human predictions," in *Proc. Robot., Sci. Syst. XIV*, Jun. 2018.

- [23] A. Bajcsy et al., "A scalable framework for real-time multi-robot, multi-human collision avoidance," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 936–943.
- [24] K. Hawkins and P. Tsiotras, "Anticipating human collision avoidance behavior for safe robot reaction," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 6301–6306.
- [25] S. Pellegrinelli, H. Admoni, S. Javdani, and S. Srinivasa, "Human-robot shared workspace collaboration via hindsight optimization," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 831–838.
- [26] Y. Cheng, L. Sun, C. Liu, and M. Tomizuka, "Towards efficient human-robot collaboration with robust plan recognition and trajectory prediction," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2602–2609, Apr. 2020.
- [27] R. Wilcox, S. Nikolaidis, and J. Shah, "Optimization of temporal dynamics for adaptive human-robot interaction in assembly manufacturing," in *Proc. Robot., Sci. Syst. VIII*, Jul. 2012, pp. 441–448.
- [28] H. Ding, M. Schipper, and B. Matthias, "Optimized task distribution for industrial assembly in mixed human-robot environments—case study on IO module assembly," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2014, pp. 19–24.
- [29] K. P. Hawkins, N. Vo, S. Bansal, and A. F. Bobick, "Probabilistic human action prediction and wait-sensitive planning for responsive human-robot collaboration," in *Proc. 13th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, Oct. 2013, pp. 499–506.
- [30] K. P. Hawkins, S. Bansal, N. N. Vo, and A. F. Bobick, "Anticipating human actions for collaboration in the presence of task and sensor uncertainty," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 2215–2222.
- [31] Y. Tanaka, J. Kinugawa, and K. Kosuge, "Motion planning with worker's trajectory prediction for assembly task partner robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Vilamoura, Portugal, Oct. 2012, pp. 1525–1532.
- [32] A. Kanazawa, J. Kinugawa, and K. Kosuge, "Adaptive motion planning for a collaborative robot based on prediction uncertainty to enhance human safety and work efficiency," *IEEE Trans. Robot.*, vol. 35, no. 4, pp. 817–832, Aug. 2019.
- [33] K. Baizid, A. Yousnadj, A. Meddahi, R. Chellali, and J. Iqbal, "Time scheduling and optimization of industrial robotized tasks based on genetic algorithms," *Robot. Comput.-Integr. Manuf.*, vol. 34, pp. 140–150, Aug. 2015.
- [34] F. Belkhouche, "Reactive path planning in a dynamic environment," *IEEE Trans. Robot.*, vol. 25, no. 4, pp. 902–911, Aug. 2009.
- [35] G. S. Aoude, B. D. Luders, J. M. Joseph, N. Roy, and J. P. How, "Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns," *Auto. Robots*, vol. 35, no. 1, pp. 51–76, Jul. 2013.
- [36] S. K. Jayaraman, L. P. Robert, X. J. Yang, A. K. Pradhan, and D. M. Tilbury, "Efficient behavior-aware control of automated vehicles at crosswalks using minimal information pedestrian prediction model," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2020, pp. 4362–4368.
- [37] E. F. Camacho and C. B. Alba, *Model Predictive Control*, 2nd ed. Springer-Verlag London, 2007.
- [38] M. P. Chapman et al., "A risk-sensitive finite-time reachability approach for safety of stochastic dynamic systems," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2019, pp. 2958–2963.
- [39] R. Tian, N. Li, I. Kolmanovsky, Y. Yildiz, and A. R. Girard, "Game-theoretic modeling of traffic in unsignalized intersection network for autonomous vehicle control verification and validation," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2211–2226, Mar. 2022.
- [40] D. Sadigh, A. Dragan, S. Sastry, and S. Seshia, "Active preference-based learning of reward functions," in *Proc. Robot., Sci. Syst. XIII*, Jul. 2017.
- [41] N. Fairfield, "Localization, mapping, and planning in 3D environments," Ph.D. dissertation, Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, 2009.
- [42] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [43] X. Yang, *Nature-Inspired Optimization Algorithms*. Amsterdam, The Netherlands: Elsevier, 2014.
- [44] M. A. Gennert and A. L. Yuille, "Determining the optimal weights in multiple objective function optimization," in *Proc. 2nd Int. Conf. Comput. Vis.*, Dec. 1988, pp. 87–89.
- [45] R. T. Marler and J. S. Arora, "The weighted sum method for multi-objective optimization: New insights," *Struct. Multidisciplinary Optim.*, vol. 41, no. 6, pp. 853–862, Jun. 2010.
- [46] H. Bai, S. Cai, N. Ye, D. Hsu, and W. S. Lee, "Intention-aware online POMDP planning for autonomous driving in a crowd," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 454–460.
- [47] D. Fridovich-Keil et al., "Confidence-aware motion prediction for real-time collision avoidance," *Int. J. Robot. Res.*, vol. 39, nos. 2–3, pp. 250–265, Mar. 2020.
- [48] C. You, J. Lu, D. Filev, and P. Tsiotras, "Advanced planning for autonomous vehicles using reinforcement learning and deep inverse reinforcement learning," *Robot. Auto. Syst.*, vol. 114, pp. 1–18, Apr. 2019.
- [49] J. Jara-Ettinger, "Theory of mind as inverse reinforcement learning," *Current Opinion Behav. Sci.*, vol. 29, pp. 105–110, Oct. 2019.
- [50] S. Arora and P. Doshi, "A survey of inverse reinforcement learning: Challenges, methods and progress," *Artif. Intell.*, vol. 297, Aug. 2021, Art. no. 103500.
- [51] J. Zacharias, "Pedestrian behavior pedestrian behavior and perception in urban walking environments," *J. Planning Literature*, vol. 16, no. 1, pp. 3–18, Aug. 2001.
- [52] B. J. Campbell, C. V. Zegeer, H. H. Huang, and M. J. Cyneck, "A review of pedestrian safety research in the United States and abroad," Univ. North Carolina, Chapel Hill, NC, USA, Tech. Rep. FHWA-RD-03-042;HRDS-06/01-04(1M)E, Nov. 2003.
- [53] V. Mehta, "Walkable streets: Pedestrian behavior, perceptions and attitudes," *J. Urbanism, Int. Res. Placemaking Urban Sustainability*, vol. 1, no. 3, pp. 217–245, Nov. 2008.
- [54] D. McAslan, "Walking and transit use behavior in walkable urban neighborhoods," *Michigan J. Sustainability*, vol. 5, no. 1, pp. 51–71, May 2017.
- [55] H. Admoni, T. Weng, B. Hayes, and B. Scassellati, "Robot nonverbal behavior improves task performance in difficult collaborations," in *Proc. 11th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2016, pp. 51–58.
- [56] D. Whitney, E. Rosen, J. MacGlashan, L. L. S. Wong, and S. Tellex, "Reducing errors in object-fetching interactions through social feedback," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 1006–1013.
- [57] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking system with Kalman filtering," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2004, pp. 584–589.
- [58] U. Scheuerner, H. Cramer, B. Fardi, and G. Wanielik, "Multi sensor based tracking of pedestrians: A survey of suitable movement models," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2004, pp. 774–778.
- [59] R. F. Carpio et al., "A navigation architecture for Ackermann vehicles in precision farming," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1103–1110, Apr. 2020.
- [60] H. Zhang, Y. Zhang, C. Liu, and Z. Zhang, "Energy efficient path planning for autonomous ground vehicles with Ackermann steering," *Robot. Auto. Syst.*, vol. 162, Apr. 2023, Art. no. 104366.
- [61] J. Peng, Y. Chen, Y. Duan, Y. Zhang, J. Ji, and Y. Zhang, "Towards an online RRT-based path planning algorithm for Ackermann-steering vehicles," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 7407–7413.
- [62] A. D. Luca, G. Oriolo, and C. Samson, "Feedback control of a nonholonomic car-like robot," in *Robot Motion Planning Control*, J. P. Laumond, Ed. Cham, Switzerland: Springer, 1998, pp. 171–253.
- [63] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Conf. Robot. Autom.*, Sep. 2004, pp. 284–289.
- [64] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das, *Cooperative Control of Multi-Agent Systems: Optimal and Adaptive Design Approaches*. London, U.K.: Springer-Verlag, 2014.
- [65] A. Torreno, E. Onaindia, A. Komenda, and M. Stolba, "Cooperative multi-agent planning: A survey," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–32, Nov. 2018.
- [66] T. Mylvaganam, M. Sassano, and A. Astolfi, "A differential game approach to multi-agent collision avoidance," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 4229–4235, Aug. 2017.
- [67] M. Chen, J. F. Fisac, S. Sastry, and C. J. Tomlin, "Safe sequential path planning of multi-vehicle systems via double-obstacle Hamilton-Jacobi-Isaacs variational inequality," in *Proc. Eur. Control Conf. (ECC)*, Jul. 2015, pp. 3304–3309.
- [68] M. Chen, S. Bansal, J. F. Fisac, and C. J. Tomlin, "Robust sequential trajectory planning under disturbances and adversarial intruder," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 4, pp. 1566–1582, Jul. 2019.
- [69] T. Hedden and J. Zhang, "What do you think I think you think?" *Cognition*, vol. 85, no. 1, pp. 1–36, Aug. 2002.

- [70] M. A. Costa-Gomes, V. P. Crawford, and N. Iriberri, "Comparing models of strategic thinking in van Huyck, Battalio, and Beil's coordination games," *J. Eur. Econ. Assoc.*, vol. 7, nos. 2–3, pp. 365–376, Apr. 2009.
- [71] D. O. Stahl, "Evolution of smart<sub>n</sub> players," *Games Econ. Behav.*, vol. 5, no. 4, pp. 604–617, Oct. 1993.
- [72] D. O. Stahl and P. W. Wilson, "Experimental evidence on players' models of other players," *J. Econ. Behav. Org.*, vol. 25, no. 3, pp. 309–327, Dec. 1994.
- [73] Y. Yildiz, A. Agogino, and G. Brat, "Predicting pilot behavior in medium scale scenarios using game theory and reinforcement learning," *J. Guid. Control Dyn.*, vol. 37, no. 4, pp. 1335–1343, 2014.
- [74] N. Li, D. W. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and A. R. Girard, "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1782–1797, Sep. 2018.
- [75] M. P. Chapman, "Risk-sensitive safety analysis and control for trustworthy autonomy," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, Berkeley, CA, USA, 2020.



**Mehdi Hosseinzadeh** (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2016.

From 2017 to 2019, he was a Post-Doctoral Researcher at Université Libre de Bruxelles, Brussels, Belgium. From October 2018 to December 2018, he was a Visiting Researcher at the University of British Columbia, Vancouver, BC, Canada. From 2019 to 2022, he was a Post-Doctoral Research Associate at Washington University in St. Louis, MO, USA. He is currently an Assistant Professor at the School of Mechanical and Materials Engineering, Washington State University, Pullman, WA, USA. His research focuses on safety, resilience, and long-term autonomy of autonomous systems, with applications to autonomous robots, energy systems, video streaming, and drug delivery systems.



**Bruno Sinopoli** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of California at Berkeley, Berkeley, CA, USA, in 2005.

After a postdoctoral position at Stanford University, Stanford, CA, he was the Faculty at Carnegie Mellon University, Pittsburgh, PA, USA, from 2007 to 2019, where he was a Full Professor at the Department of Electrical and Computer Engineering with courtesy appointments in mechanical engineering and in the Robotics Institute and the Co-Director of the Smart Infrastructure Institute. In 2019, he joined Washington University in St. Louis, where he is the Chair of the Electrical and Systems Engineering Department. His research interests include the modeling, analysis and design cyber-physical systems with applications to energy systems, interdependent infrastructures and internet of things.



**Aaron F. Bobick** (Fellow, IEEE) received the B.S. degree in mathematics and computer science and the Ph.D. degree in cognitive science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1981 and 1987, respectively.

In 1987, he joined the Perception Group of the Artificial Intelligence Laboratory, SRI International, Menlo Park, CA, USA, and soon after was jointly named a Visiting Scholar at Stanford University, Stanford, CA, USA. From 1992 to July 1999, he served as an Assistant and, then, an Associate Professor at the Vision and Modeling Group of the MIT Media Laboratory. From 1999 to 2015, he was with the College of Computing, Georgia Institute of Technology, Atlanta, Georgia. In 2015, he joined Washington University in St. Louis, St. Louis, MO, USA, as the Dean of the School of Engineering and Applied Science and the James M. McKelvey Professor. His research primarily focuses on action recognition by computer vision—a field in which he is a pioneer—and robot perception for human-robot collaboration.