Scalable Confidence-Aware Safety Analysis for Robot Planning around Multiple Humans

Jaime F. Fisac, David Fridovich-Keil, Andrea Bajcsy, Sylvia Herbert, Sampada Deglurkar,
Claire Tomlin, and Anca Dragan
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
{ jfisac, dfk, abajcsy, sylvia.herbert, sampada_deglurkar, tomlin, anca }@berkeley.edu

Abstract*—In order to safely share the space with people, it is important for robots to reason about future human motion. While predictive models often perform acceptably for a single human, safe navigation among multiple people presents two serious challenges. First, interaction with others is an inherently complex aspect of human behavior, making reliable predictions difficult to obtain. Second, reasoning jointly about many interacting agents faces a combinatorial explosion in computation, making joint distributions of multiple human trajectories highly intractable. Making simplifying assumptions may seem like the only feasible approach, yet inaccurate predictions can compromise the safety of a robot's motion plan. In recent work, we proposed a Bayesian framework to reason about the reliability of model predictions in real time, allowing the robot to quickly adapt its uncertainty about future human actions when its model performed poorly; combining its *confidence-aware* predictions with a robust motion planning and control scheme the robot could successfully plan and execute probabilistically safe plans around a single human. In this work, we leverage a natural strength of this method to tackle multi-human predictions in a way that allows safe robot navigation: if the robot's predictive model fails to accurately capture a human's behavior while interacting with others, predictions will naturally become uncertain when interactions significantly affect her motion, and will regain confidence once the effect diminishes. The robot can then use simple but highly scalable predictive models that simplify or even fully neglect the interaction component between multiple humans, naturally maneuvering more conservatively around humans who are substantially deviating from their predicted behavior due to interaction. In the case of non-interactive models, our algorithm becomes fully parallel between humans, enabling prediction and planning of the same time complexity as the single-human case. We foresee that efficiently scaling robust motion plans to multiple humans with an adaptive amount of conservativeness can enable safe robot motion in a variety of human-populated environments.

As the capabilities of autonomous robotic systems continue to advance, their domain of application will further extend from the traditional controlled factory settings into more complex and unstructured environments. Crucially, at least some of these new domains will require robots to coexist with human beings in the same space, be it service robots helping in the home or autonomous drones joining the civilian airspace [7], making safety a paramount consideration. Safety analysis, however, becomes substantially more challenging in the presence of human actors, due to the difficulty in providing reliable quantitative guarantees when these are contingent on human behavior. While predictive approaches in recent years, often combining learning techniques with models from cognitive science [3], [9], [12], have enabled robots to form useful predictions for a single human, these techniques cannot be readily applied to the problem of safe robotic navigation among multiple humans.

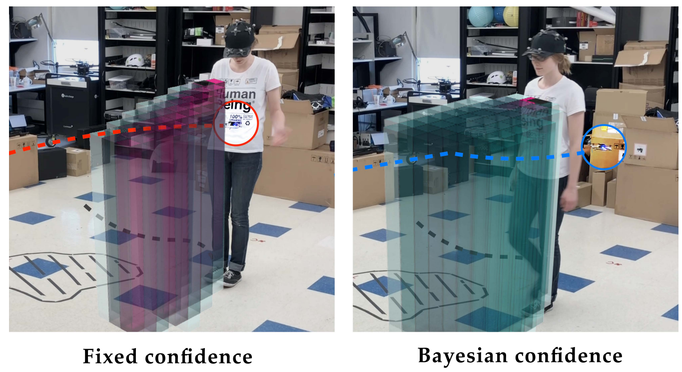

Fig. 1: When planning around humans, predictive models can enable robots to reason about future motions the human might take. These predictions rely on human motion models, but such models will often be incomplete and lead to inaccurate predictions and even collisions (left). Our method addresses this by updating its *human model confidence* in real time (right). Video available online: https://youtu.be/2ZRGxWknENg.

Two main difficulties emerge when considering the multi-human prediction problem. On the one hand, there is a modeling challenge: while certain modes of human behavior, such as pedestrian navigation in isolation [13], have allowed some success in predictions from low-dimensional state and utility representations, this seems much less likely to be the case in scenarios including interaction. For example, two humans currently walking towards each other may sidestep to avoid colliding and continue on their paths, but they may also stop to greet or converse with each other; or they may hug, effectively colliding. Forming an accurate prediction in such scenarios would require a model representation considerably richer and more subtle than for a single pedestrian. These situations, however, are not uncommon in human behavior, and autonomous robots will need to handle them gracefully if they are to be deployed in human-populated environments.

On the other hand, there is a computational challenge: even assuming that a reliable human interaction model is available, it is then necessary to use this model to generate predictions about the future actions of nearby humans. Existence of interactions implies that the predictions for different humans will in general be coupled with each other. Reasoning over possible joint trajectories therefore carries a combinatorial computational complexity, making real-time implementation highly impractical. For example, assume a simple state-space representation of each human as being given by only their X and Y coordinates on the ground; reasoning about the uncertain future location of a single human in isolation requires propagating a distribution over a 2-dimensional space, which may be implemented as a discretized 100×100 probability grid, with $10,000$ entries. However, if there are 5 interacting humans, then the distribution over their *joint* positions will be over a 10-dimensional space, therefore requiring a grid with $100^{10} = 10^{20}$ (100 quintillion) entries: even if each entry was represented as a single-precision floating-point number, this exceeds the memory capacity of any computer in existence today by several orders of magnitude.

In order to maintain tractability, some work has proposed using heuristic methods to account for interactions between multiple humans without needing to evaluate probabilities in the joint hypothesis space [10]. While these heuristics can certainly provide useful approximations in limited settings, they will eventually perform poorly, especially in scenarios that deviate from the conditions for which the heuristic was designed. In these cases, the system may make incorrect yet equally confident predictions about humans' future actions. Making poor predictions *and* failing to quickly acknowledge them as such can lead an autonomous robot to act overconfidently and safety may be compromised.

Proposed method. While using a simplified model of interaction can be a useful strategy to provide computable probabilistic predictions of joint human trajectories, we claim that the robot should also have the ability to reason in real time about how confident it can afford to be on its model's predictions. In recent work, we proposed a Bayesian approach to allow the robot to carry out a fast but effective meta-analysis of its own predictions and infer, in real time, the current reliability of its predictive model [4]. This, combined with a novel robust motion planning and control scheme [5], allowed the robot to compute trajectories with a quantitative safety guarantee in the presence of a single human. The framework was demonstrated on quadcopter flight avoiding a walking pedestrian (Fig. 1).

In the work presented here, we apply this method to the problem of safe robot navigation in the presence of multiple humans, as a principled means to mitigate the inevitable prediction inaccuracies resulting from the use of tractable but limited human predictive models. We believe that our framework [4] can be used in conjunction with approximate prediction methods [6], [10], fulfilling a complementary role. Since these approximations will occasionally produce inaccurate predictions for any particular human, our real-time Bayesian inference on the underlying model confidence (based on how "rationally" this human appears to be following the model) will naturally detect anomalous behavior and accordingly increase the uncertainty in the model's predictions. Conversely, when the human's behavior becomes more likely according to the model, prediction uncertainty decreases.

It is important to note here that our approach leverages the inverse temperature parameter in Boltzmann-rational models, which are standard in the inverse optimal control (also inverse reinforcement learning) literature [3], [9], [12]. However, we stress that the central principle in this work can more generally be applied to probabilistic models in which a small number of parameters determine the entropy of the distribution; investigating such an extension lies beyond the scope of this paper and we will consequently assume a Boltzmann-rational model class.

As in [4], our proposed method combines probabilistic human predictions with robust robot motion planning. The confidence-aware prediction component continually keeps track of how well its model is describing the motion of *each* of the humans in the environment, and produces real-time probabilistic predictions based on this confidence. These improved human motion predictions are then combined, in conjunction with the robust motion planning scheme [5] to obtain marginal collision probabilities. The result are *probabilistically safe* robotic motion plans that are conservative when the model cannot confidently predict the motion of nearby humans, but efficient whenever model predictions perform well.

Formulation. We consider a single robot moving to a preset goal location in a space shared with N humans. Let $x_R \in \mathbb{R}^{n_R}$ be the robot state for planning purposes, and let $x_i \in \mathbb{R}^{n_H}$ be the state of the i-th human, with $i \in \{1, .., N\}$. These states could represent the locations of a mobile robot and a number of human pedestrians in a shared environment, the kinematic configurations a robotic manipulator and several humans' arms in a workspace, or the state of an autonomous vehicle and a number of human-driven vehicles on the road.[1] In addition let $x = [x_R, x_1, \ldots, x_N] \in \mathbb{R}^{n_R + N n_H}$ denote the joint state of all agents for planning purposes; we assume x is observable to all agents. The dynamics are

$$x_R^{\tau+1} = f_R(x_R^\tau, u_R^\tau), \qquad x_i^{\tau+1} = f_H(x_i^\tau, u_i^\tau), \tag{1}$$

where $u_R \in \mathbb{R}^{m_R}$ and $u_i \in \mathbb{R}^{m_H}$ are the control actions of the robot and the i-th human, respectively. Further, let $s_R \in \mathbb{R}^{n_S}$ denote the state of higher-fidelity dynamical robot model, capturing e.g. inertias and actuator delays ignored by x_R.

The robot needs to plan a trajectory that, when tracked by the physical system, will reach a goal state as efficiently as possible while avoiding collisions with high confidence, based on an informed prediction of each human's future motion. Since safety guarantees will inevitably inherit the probabilistic nature of this prediction, our goal is to find efficient plans that will keep collisions with humans below an acceptable probability. Formally, given a current state $x_R^{\text{now}} \in \mathbb{R}^{n_R}$, a cumulative cost $c : \mathbb{R}^{n_R} \times \mathbb{R}^{m_R} \to \mathbb{R}$, a probability threshold $P_{\text{th}} \in [0, 1]$ and a final time T, we define the constrained planning problem:

[1] We can more generally combine different types of human agents (e.g. pedestrians, cyclists, drivers) each with their own dynamics class. For clarity of exposition, however, we consider identical dynamics for all humans.

$$\min_{u_R^{t:T}} \sum_{\tau=t}^{T} c(x_R^\tau, u_R^\tau) \tag{2a}$$

$$\text{s.t. } x_R^t = x_R^{\text{now}} \tag{2b}$$

$$x_R^{\tau+1} = f_R(x_R^\tau, u_R^\tau), \quad \tau \in t, ..., T-1 \tag{2c}$$

$$P_{\text{coll}}^{t:T} := P\big(\exists \tau \in \{t, ..., T\} : \text{coll}(x_R^\tau, \{x_i^\tau\}_{i=1}^N)\big) \leq P_{\text{th}} \tag{2d}$$

The term $\text{coll}(x_R^\tau, \{x_i^\tau\}_{i=1}^N)$ is a Boolean variable indicating whether the robot is in collision with at least one human at time τ. That is:

$$\text{coll}(x_R^\tau, \{x_i^\tau\}_{i=1}^N) = \text{coll}(x_R^\tau, x_1^\tau) \vee \ldots \vee \text{coll}(x_R^\tau, x_1^\tau) \ . \tag{3}$$

To compute $P_{\text{coll}}^{0:T}$, the robot must account for the human agents' future movements, as well as the deviations of the physical robot from the ideal motion plan at execution time.

Confidence-aware prediction. The analysis in this work follows the formulation in [4] and is based on Boltzmann-rational models of human decision-making (also referred to as maximum entropy or soft-max models). A large body of work in econometrics and cognitive science has investigated and validated the modeling of human behavior by utility-driven optimization [1], [8], [11]; these models have also been widely adopted within the robotics community [3], [9]. The robot models the i-th human as attempting to optimize a reward function, $r_i(x, u_i; \theta_i)$, that can generally depend on the joint state and her own action, as well as a set of parameters θ_i (e.g. the human's intended goal location). Given r_i, the robot can compute the i-th human's policy, defined as a probability distribution over actions conditioned on the state. The robot models human agents as likely to choose actions with high expected utility, in this case the (state-action) Q-value:

$$P(u_i^t \mid x^t; \beta_i, \theta_i) \propto e^{\beta_i Q_i(x^t, u_i^t; \theta_i)} \ . \tag{4}$$

The *inverse temperature* term β_i is traditionally called the *rationality coefficient* and it determines the degree to which the robot expects to observe human actions aligned with its model of the utility. A common interpretation of $\beta_i = 0$ is a human who appears "irrational", choosing actions uniformly at random and completely ignoring the modeled utility, while $\beta_i \to \infty$ corresponds a "perfectly rational" human. We observed in [4] that β_i can instead be leveraged as an indicator of the model's predictive capabilities, and refer to it as *model confidence*. In practice, the same model will perform differently over time, across changing situations and for each human individual. Maintaining a Bayesian belief over each human's β_i as a *hidden state*, the robot can dynamically adapt its predictions, and thereby its motion plan, to the changing reliability of its human model for each human individual. This belief over β_i is compatible (and can in fact be updated jointly) with a Bayesian belief over the parameters θ_i of the human's reward function. The joint belief over (θ_i, β_i) allows rich multimodal distributions and is compatible with mixture models.

At runtime, every new observed action u_i^t of the i-th human provides the robot with a "measurement update" to its belief $b^t(\cdot)$ about the model confidence β_i, following Bayes' rule:

$$b^{t+1}(\beta_i \mid \theta_i) = \frac{P(u_i^t \mid x^t; \beta_i, \theta_i) b^t(\beta_i \mid \theta_i)}{\sum_{\hat{\beta}_i} P(u_i^t \mid x^t; \hat{\beta}_i, \theta_i) b^t(\hat{\beta}_i \mid \theta_i)} \ , \tag{5}$$

with prior $b^t(\beta_i \mid \theta_i) = P(\beta_i \mid x^{0:t}, \theta_i)$ for $t \in \{0, 1, ...\}$, and likelihood $P(u_i^t \mid x^t; \beta_i, \theta_i)$ given by (4). It is critical to be able to perform this update extremely fast, which would be difficult to do in the original continuous hypothesis space $\beta_i \in [0, \infty)$. Fortunately, we observe that maintaining a Bayesian belief over a relatively small set of β_i values ($M_\beta \approx 10$ on a log-scale) achieves significant improvement relative to maintaining a fixed precomputed value. Marginalizing over β_i (and possibly θ_i), we can now proceed as we would have with the standard (fixed-β_i) predictive model. Combining (4) with dynamics (1) gives $P(x_i^{\tau+1} \mid x^\tau; \beta_i, \theta_i)$ for each human, which allows us to recursively propagate the humans' motion over time and compute a joint probabilistic prediction of their state occupancy.

Complexity. Notice that determining the Q-value used in (4) for each human may be a nontrivial problem, depending on the interaction model. For example, if the humans are assumed to behave in a strategic way accounting for each other's future decisions, the full solution to $\{Q_i\}_{i=1}^N$ is game-theoretic and, if at all tractable, may need to be numerically precomputed offline [2]. Other interaction models may propose simpler formulations of Q_i, or might learn it directly using inverse optimal control [9]. For example, if humans are modeled as being reactive to other agents in acting but oblivious of them in planning, their subjective Q-value has the form $Q_i(x, u_i; \theta_i) = r_i(x, u_i; \theta_i) + V_i(x_i; \theta_i)$, where V_i can more tractably be computed for each human separately. The computation of the Q-value is specific to the choice of human interaction model, and is independent of whether or not our proposed framework is being used. We stress that our framework is agnostic to the concrete choice of interaction model and does not affect its core computations.

Once the Q-value from the chosen model is available, the Bayesian update step introduced by our framework is extremely scalable with the number of humans N. Crucially, the Bayesian update (5) can be performed independently for each i-th human given the observed state x, which means the total amount of computation is linear in N and, if parallel computation is available, computation *time* is constant in N. Therefore, any Boltzmann-rational predictive model of multi-human motion that is suitable to use in real time can be made confidence-aware and continue to be suitable for real time.

Collision probability. The final step is to integrate the joint human prediction into the robot's motion planning to obtain a safety guarantee not only for the *planned* state x_R, but for the *physical* state s_R at execution time. The recent FaSTrack framework [5] uses Hamilton-Jacobi robust optimal control to compare the dynamics of x_R and s_R (possibly under bounded disturbances) and compute a guaranteed tracking error bound $\mathcal{E} \subset \mathbb{R}^{n_R}$ between the planned and realized state at any point along the executed trajectory, as well as an optimal tracking control policy to enforce the guarantee.

For any planned state x_R^τ, and given the tracking error bound $\mathcal{E}$, let $\mathcal{H}_\mathcal{E}(x_R^\tau)$ be the set of human states that would be in collision with *some* robot state s_R^τ allowed by $\mathcal{E}$ around x_R^τ. We then have $\text{coll}(x_R^\tau, x_i^\tau) \iff x_i^\tau \in \mathcal{H}_\mathcal{E}(x_R^\tau)$. We can thus write the probability of a collision with at least one human as

$$P\big(\text{coll}(x_R^\tau, \{x_i^\tau\}_{i=1}^N)\big) = \\ 1 - \prod_{i=1}^{N} P\big(\neg\text{coll}(x_R^\tau, x_i^\tau) \mid \neg\text{coll}(x_R^\tau, \{x_j^\tau\}_{j=1}^{i-1})\big) \ , \quad (6)$$

where each of the terms in the right-hand product can in principle be obtained by iteratively integrating the conditional distribution $P(x_i^\tau \mid x_1, ..., x_{i-1})$ on the compact set $\mathcal{H}_\mathcal{E}(x_R^\tau)$.

These operations can be expected to become extremely computationally intensive using a joint probability distribution, which would in general have to be discretized as a $N \times n_H$-dimensional grid. Operating with (or even storing) such a grid becomes rapidly impractical as N grows beyond trivial numbers due to the infamous "curse of dimensionality".

Instead, tractable approximations can be computed by only storing the marginal predicted distribution of each human at every future time step τ. This way, the robot need only operate with N grids of n_H dimensions. Neglecting the effect of conditioning, the distributions in (6) are treated as independent, giving a computationally simple noisy-OR operation:

$$P\big(\text{coll}(x_R^\tau, \{x_i^\tau\}_{i=1}^N)\big) \approx 1 - \prod_{i=1}^{N} \Big(1 - P\big(\text{coll}(x_R^\tau, x_i^\tau)\big)\Big) \ . \quad (7)$$

The planner can then reject plans for which, at any time $\tau > t$, $P_{\text{coll}}^\tau \geq P_{\text{th}}$, returning real-time motion plans that are (approximately, when exact computation is intractable) P_{th}-safe for the physical robot.

Results. Preliminary simulation results using the marginal probability distributions and the noisy-OR approximation show promising predictions and robot behavior. As expected, the increase in uncertainty when two humans deviate to avoid each other is comparable to the one observed in [4] for a single human avoiding a static obstacle (Fig. 1, right). Further simulations and physical experiments are in preparation.

Discussion. This work has made a first effort to generalize and extend the confidence-aware probabilistically safe planning methodology recently presented in [4] to the multi-human case. It has presented the theoretical analysis as well as important practical considerations regarding tractability. One of the key strengths of our confidence-aware prediction framework is that it is naturally resilient to model misspecification, by quickly adapting conservativeness of predictions to gradual or sudden degradation in the model's ability to capture human behavior. As a result, it can be a strong tool with which to address the intractability of exact combinatorial predictions of multiple humans: since exact predictions cannot be made, our framework allows simplified interaction models to produce approximate probabilistic predictions of human motion, quickly increasing uncertainty when the actual interaction fails to match the models.

References

[1] Chris L Baker, Joshua B Tenenbaum, and Rebecca R Saxe. Goal inference as inverse planning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 29, 2007.

[2] Tamer Basar and Alain Haurie. Feedback Equilibria in Differential Games with Structural and Modal Uncertainties. *Advances in Large Scale Systems*, 1984.

[3] Chelsea Finn, Sergey Levine, and Pieter Abbeel. Guided cost learning: Deep inverse optimal control via policy optimization. *International Conference on International Conference on Machine Learning*, 2016.

[4] Jaime F. Fisac, Andrea Bajcsy, Sylvia L. Herbert, David Fridovich-Keil, Steven Wang, Claire J. Tomlin, and Anca D. Dragan. Probabilistically safe robot planning with confidence-based human predictions. *Robotics: Science and Systems (RSS)*, 2018.

[5] Sylvia L. Herbert*, Mo Chen*, SooJean Han, Somil Bansal, Jaime F. Fisac, and Claire J. Tomlin. Fastrack: a modular framework for fast and guaranteed safe motion planning. *IEEE Conference on Decision and Control*, 2017.

[6] Boris Ivanovic, Edward Schmerling, Karen Leung, and Marco Pavone. Generative modeling of multimodal multi-human behavior. *International Conference on Intelligent Robots and Systems (IROS)*, 2018.

[7] Parmial Kopardekar. Unmanned Aerial System (UAS) Traffic Management (UTM): Enabling Low-Altitude Airspace and UAS Operations. Technical report, NASA, 2014.

[8] R. Duncan Luce. *Individual choice behavior: A theoretical analysis*. Wiley, 1959.

[9] Deepak Ramachandran and Eyal Amir. Bayesian inverse reinforcement learning. *Urbana*, 51(61801), 2007.

[10] Peter Trautman and Andreas Krause. Unfreezing the robot: Navigation in dense, interacting crowds. *International Conference on Intelligent Robots and Systems (IROS),*, 2010.

[11] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton University Press Princeton, NJ, 1945.

[12] Brian D. Ziebart, Andrew L. Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. *AAAI*, 2008.

[13] Brian D. Ziebart, Nathan Ratliff, Garratt Gallagher, Christoph Mertz, Kevin Peterson, J. Andrew Bagnell, Martial Hebert, Anind K Dey, and Siddhartha Srinivasa. Planning-based prediction for pedestrians. *International Conference on Intelligent Robots and Systems*, 2009.