Risk-Calibrated Human-Robot Interaction via Set-Valued Intent Prediction

Justin Lidard, Hang Pham, Ariel Bachman, Bryan Boateng, Anirudha Majumdar
Department of Mechanical and Aerospace Engineering
Princeton University, Princeton, New Jersey 08540
Email: jlidard@princeton.edu

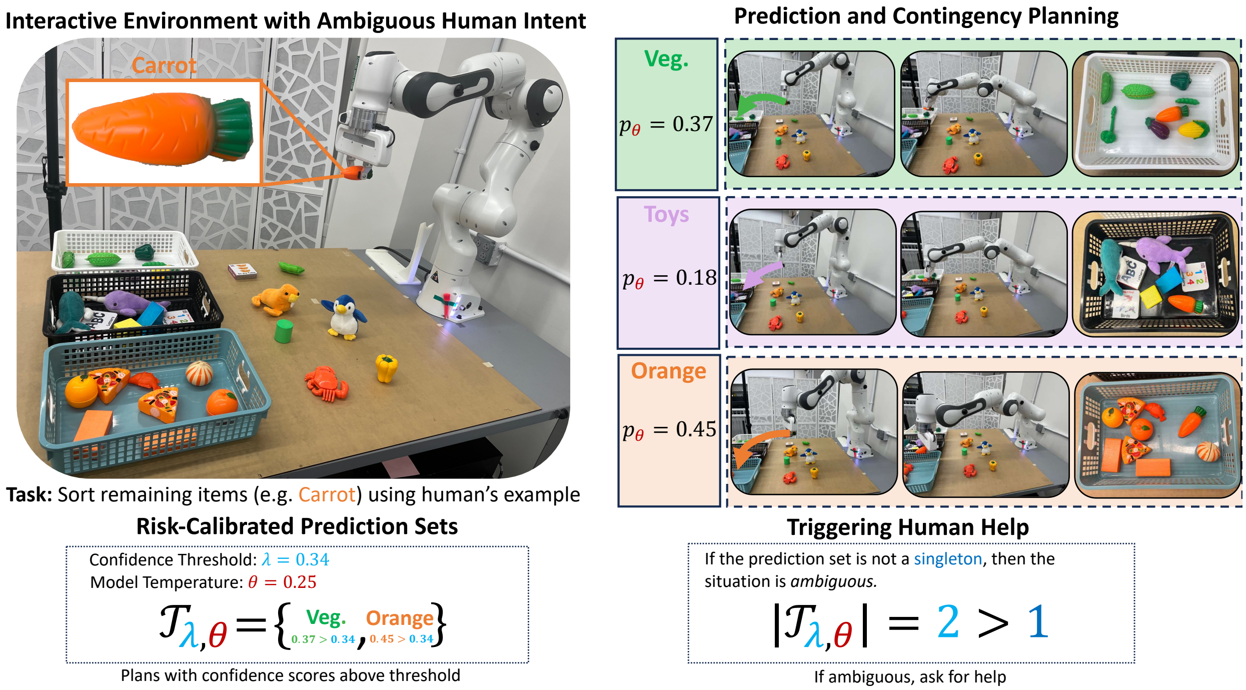

Fig. 1: Risk-Calibrated Interactive Planning (RCIP) statistically calibrates risk for human-robot interaction. Given a set of possible human intents and confidence scores, a planner generates a weighted set of actions. The set of actions from each plan are collected in a set according to a threshold on the predicted intents. If there is more than one action in the set, the robot asks for help.

I. Introduction

Predicting and understanding human intent is a critical task for robotics, specifically for safe interaction with humans in cluttered, close-quarters environments. However, human intent prediction is challenging because no two humans may have the same preferences, and intents may differ depending on the specific environment. As an example, a robot is tasked with sorting items into three bins based on an example provided by the human (see Fig. 1). While the bins have a ground-truth sorting criterion known by the human (vegetables, children's toys, and miscellaneous orange items), the robot must infer the human's intent in order to sort new items. Given the provided context, the robot should be able to sort some unambiguous items (e.g. the crab) autonomously, while other items (e.g. the carrot) may be placed into multiple bins, resulting in *situational ambiguity*. If asked to operate fully autonomously, the robot must take a *risk* and guess the correct placement for the carrot. However, the robot may also *ask for help* if it is unsure, guaranteeing the correct action but potentially burdening the human.[1]

[1]Website with additional information, videos, and code: https://risk-calibrated-planning.github.io/

Statement of contributions. In this work, we introduce RCIP, a framework for measuring and calibrating risk in situations that involve interactions with humans with potentially ambiguous action choices. Our approach utilizes deep-learned human intent prediction models (e.g. [1, 2]) for understanding interactivity, and rigorously quantifies the uncertainty of these models in order to decide when to ask for help. As shown in Fig. 1 (middle), we produce a limited set of human intents based on the prediction model's confidence scores. By reasoning about the human's desired task outcome in the space of *intents*, we efficiently plan safe actions in the face of diverse, multi-modal human behavior, and ask for help when necessary. We make the following contributions: **(1)** We demonstrate how to use statistical risk control (SRC) to control the planning error rate across a set of model hyper-parameters, allowing flexible but provably safe levels of autonomy. **(2)** We prove theoretical guarantees for multi-dimensional risk control for both single-step and multi-step planning problems: with a set of user-specified risk budgets $(\alpha_1, \ldots, \alpha_K)$ for different measures of risk (e.g., probability of failure and probability that the robot asks for help) and the robot performs the task correctly (with high probability) by asking for help if any of the K risk budgets will be violated. **(3)** We evaluate RCIP in both simulation and hardware with a suite of human-robot interactive planning tasks with various styles of situational ambiguity. Experiments across multiple platforms and human uncertainty showcase the ability of RCIP to provide statistically guaranteed task success rates while providing more flexible autonomy levels than baseline approaches. RCIP reduces the amount of human help by $5-30\%$ versus baseline approaches.

II. Related Work

RCIP brings together techniques from contingency planning, human intent prediction, and conformal prediction and empirical risk control. We discuss related work in each area below.

A. Contingency Planning and Priviledged Learning

Contingency planning [3] is a growing literature on planning for multi-agent interactive scenarios where future outcomes are diverse. Recent approaches [4]–[7] typically favor a predict-then-plan approach, wherein multi-modal motion predictions are first generated and then used to produce a set of safe plans conditioned on each prediction. The authors of [8] formulate a multi-agent contingency planning problem as a generalized Nash equilibrium problem, thereby assuming that agents are non-cooperative. In this work, we assume that the human and robot act in good faith (i.e., they are cooperative).

In this work, we provide the robot with additional information about the internal state of the "human" during the planning phase. We eliminate the need for a separate distillation procedure by instead using a set-valued prediction strategy, introduced in the following sections. By allowing the robot to ask for help when it is uncertain, we statistically quantify risk associated with the robot acting optimally, even when it is uncertain.

B. Human Intent Prediction

Predicting intent of humans for downstream planning has been widely applied in autonomous driving [9]–[11], social navigation [2, 12, 13], and game theory [14]. Several works [10, 15] use a discrete latent variable to capture qualitative behaviors in human motion. In this work, we use intent prediction *to bound directly the risk associated with downstream planning*. We use *set-valued* prediction to compute a set of possible intents, from which a planning module can compute a conditional plan.

C. Conformal Prediction and Empirical Risk Control

Conformal prediction [16]–[18] has recently gained popularity in a variety of machine learning and robotics applications due to its ability to rigorously quantify and calibrate uncertainty. A recent line of works [19]–[21] has extended the theory from prediction of labels (e.g. actions) to sequences (e.g. trajectories), and other works [22, 23] have extended conformal prediction theory to handle more general notions of risk. Our work differs in three key ways: (i) we provide a separate calibration stage in which the robot can adjust its parameterization of prediction sets through a modest-size dataset of interactive scenarios, reducing the number of "unrecoverable" scenarios in which the robot exceeds its risk budget early on in a rollout, and (ii) we provide a way to synthesize from scratch risk-averse control policies, and (iii) we reason about human uncertainty in the space of *intents*, permitting a more natural way to capture diverse interactive behaviors than other representations (e.g. trajectories).

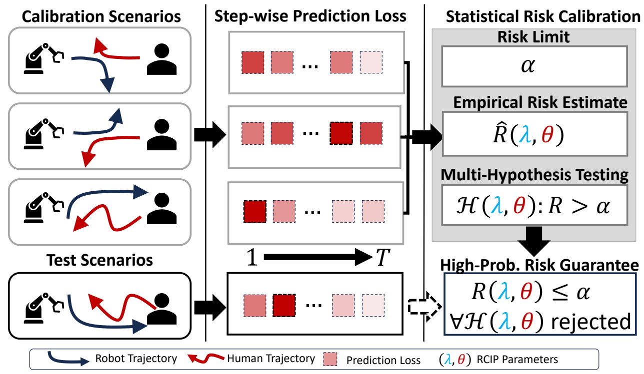

Fig. 2: RCIP formulates interactive planning as a multi-hypothesis risk control problem. Using a small set of calibration scenarios, RCIP computes step-wise prediction losses to form an aggregate emperical risk estimate. Using a risk limit, for each pair (λ,θ) of prediction thresholds and model hyperparameters, RCIP evaluates the hypothesis that the test set risk is above the limit. Thus, for all hypotheses that are rejected, the test set risk satisfies the threshold (with high probability).

III. Problem Formulation and Approach

In this section, we pose the problem of human-robot cooperation with intent uncertainty as a partially observable Markov decision process (POMDP). We present a brief overview of the prediction-to-action pipeline and our goals of risk specification and flexible autonomy.

A. Dynamic Programming with Intent Uncertainty

Environment Dynamics. We consider an interaction between a robot R and human H in environment e, governed by a nonlinear dynamical system with time horizon T:

$$x_{t+1} = f_e(x_t, u_t) \quad \forall t \in [T]. \tag{1}$$

Let $\pi = (\pi^R, \pi^H)$ denote the joint control policy.

Intent Dynamics. We assume that the human's (potentially unknown) policy π^H is parameterized by a discrete latent variable with the following dynamics:

$$z_{t+1} \sim q(\cdot | x_t, z_t), \tag{2}$$

where $z_t \in \mathcal{Z} = [N]$ characterizes the human's intent at time t, and N is the number of high-level human behaviors.

Planning Objective. Each agent $i \in \{R, H\}$ has the goal to minimize their corresponding cost function J^i in finite-horizon T with running cost l^i. The cumulative cost of a policy π^i starting from initial state x and a *known* human intent z is:

$$\begin{aligned} \min_{\pi^i} \quad & J^i(x, z, \pi^i) \\ \text{s.t.} \quad & h^i(x_t, z_t) \leq 0 \quad \forall t \in [T] \\ & (x_1, z_1) = (x, z), \end{aligned} \tag{3}$$

where constraints h^i capture satisfaction of the human intent.

Conditional Action Selection. The optimal action for both agents is given by the minimizer of their *intent-conditioned* Q-function:

$$u^{i*}(z) = \underset{u}{\operatorname{argmin}}\, Q^i(x,u,z) \quad \forall i \in \{R,H\}. \tag{4}$$

B. Risk-Calibrated Interactive Planning

Predicted Action Set. Let λ be a scalar prediction threshold (defined next) and θ be a model hyperparameter, such as temperature. We aggregate the intent predictor's confidence scores from g_θ into a set $\mathcal{S}_{\lambda,\theta} \subseteq \mathcal{Z}$ of predicted intents via the rule

$$\mathcal{S}_{\lambda,\theta}(\bar{x}_t) = \{z \in \mathcal{Z} : g_\theta(\bar{x}_t, z) \geq \lambda\}, \tag{5}$$

where λ is a confidence threshold. Since human intent uncertainty alone may not alter the optimal robot plan, we compute a set of predicted actions from the set of predicted intents as

$$\mathcal{T}_{\lambda,\theta}(\bar{x}_t) = \{u \in \mathcal{U} : \exists z \in \mathcal{S}_{\lambda,\theta} \text{ s.t. } u = u^{R*}(z) \text{ and } g^*_\theta(\bar{x}_t, z) \geq \lambda\}, \tag{6}$$

where we define g^*_θ as the sum of all intent-based confidence scores that lead to the same action, i.e., $g^*_\theta(\bar{x}_t, z) := \sum_{z' \in \mathcal{Z}} g_\theta(\bar{x}_t, z') \mathbb{1}\{u^{R*}(z) = u^{R*}(z')\}$. To simplify notation, we define $g^*_\theta(z) := g^*_\theta(\bar{x}_t, z)$.

Policy Deployment. We now define our overall robot policy Π^R. Given the predicted action set $\mathcal{T}_{\lambda,\theta}(\bar{x}_t)$ defined in Eqn. (6), the robot has two behaviors:

1) **Autonomy.** If $\mathcal{T}_{\lambda,\theta}(\bar{x}_t)$ is a singleton, then the robot is confident in the predicted action, and the action is executed.
2) **Triggering Help.** If $\mathcal{T}_{\lambda,\theta}(\bar{x}_t)$ is not a singleton, then the robot triggers human help, and the human reveals their true intent, z^*. The robot executes the action $u^{R*}(z^*)$.

If λ and θ are chosen such that $\mathcal{T}_{\lambda,\theta}(\bar{x}_t)$ is empty, the task is failed. Fig. 2 depicts the empirical risk calibration procedure used in RCIP. We prove in both single- and multi-step settings that using our procedure, a set of parameter pairs (λ, θ) control an *arbitrary* notion of risk with high probability.

IV. Experiments

We demonstrate RCIP in three multi-step, interactive domains, which exhibit three ways in which a robot planner can be integrated with an intent predictor. Each interactive prediction task is defined below.

Scenario Distribution and Calibration Dataset. RCIP can be used to obtain risk guarantees for an *unknown* scenario distribution — that is, of environments and human partners — if can can collect i.i.d. samples from it for calibration. We envision that RCIP will enable a robot to interact with an end user (or set of users) through interactive data collection. Each calibration dataset is generated by random sampling from the environment distribution and from the distribution over human intents.

Baselines. We compare RCIP against similar set-valued prediction approaches. A simple but naive approach for approximated $1-\alpha_1$ coverage of optimal actions is **Simple Set**, which ranks actions according to a $1-\alpha_1$ threshold using the predictor's raw confidence scores. Actions are sorted by greatest to least confidence, and actions are added to the prediction set in order of the sorted action set until the threshold is reached. To measure the effect of *overall uncertainty* rather than individual scores, we compare against **Entropy Set**, which includes the highest overall prediction if the entropy of the distribution predicted actions is below a threshold; if not, then all actions are included in the prediction set, and the robot must ask for help. To evaluate the performance of vanilla conformal prediction against the richer hypothesis space of RCIP, we report results for **KnowNo** [21]. Similar in spirit but different from our work, KnowNo seeks to maximize coverage of optimal actions but without any guarantees on the human help rate, and assumes model parameters are fixed. Instead of maximizing coverage outright, RCIP balances prediction of optimal actions with limits on the human help rate, providing flexible performance guarantees depending on model parameters. Lastly, we consider **No Help** as an option, where the predicted action set always contains the predictor's most-confident action, and the human help rate is identically zero.

Metrics. For all environments, we report the task-level risks of (i) plan success rate and (ii) human help rate, on the test set. We also report the instantaneous risks — measured as an average over time — of plan success and human help.

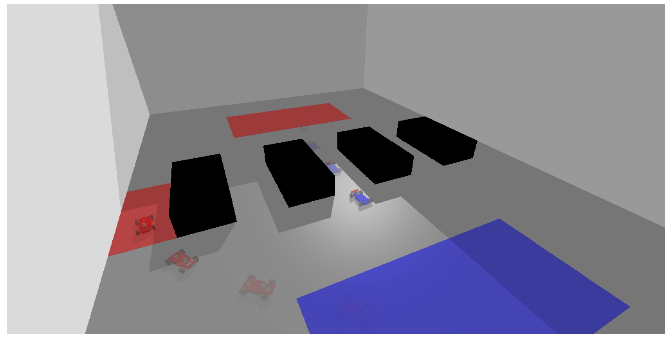

Fig. 3: Multi-step RCIP is applied in **Hallway Navigation**. The robot car (blue) and human car (red) are tasked with navigating to their respective goal states (large blue and red rectangles). The human car is constrained via its intent to pass through one of the five hallways (highlighted in red). The blue car does not observe the human's intent during evaluation.

A. Simulation: Hallway Navigation

Autonomous navigation around other autonomous decision-making agents, including humans, requires the robot to recognize scenario uncertainty (whether another agent will turn right or left) with task efficiency (energy spent braking or taking detours). While safety can almost always be guaranteed if each vehicle declares their intent at all times, such communication can be costly, especially if human prompting is involved. In this example (see Fig. 3), the robot is asked to navigate to the human vehicle's initial condition without colliding while the human does the same. The set of intents is $\mathcal{Z} = \{1,2,3,4,5\}$, where each intent corresponds to one of the five hallways. The confidence scores for each

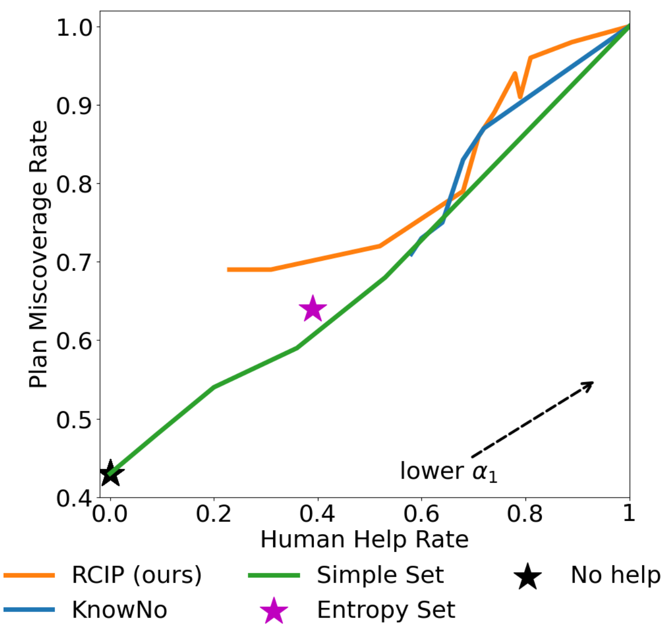

Fig. 4: Baseline comparison for RCIP verus other set-valued predictors for **Hallway Navigation**. RCIP provides a framework for tuning model parameters to achieve risk control, versus other methods that assume that model parameters are held fixed: KnowNo [21], Simple Set, Entropy Set, and No Help.

Method	$1-\alpha_1$	Plan Succ.↑	Plan Help↓	Step Succ.↑	Step Help↓
RCIP	0.85	0.86	**0.34**	**0.95**	**0.24**
KnowNo [21]	0.85	0.86	0.48	0.92	0.42
Simple Set	0.98	0.85	0.48	0.92	0.42
Entropy Set	–	0.75	0.07	0.86	0.02
No Help	–	0.73	0	0.86	0

TABLE I: Results for **Cooperative Navigation**. The optimal action miscoverage rate is held fixed between RCIP, KnowNo, and Simple Set for comparing the other metrics.

intent are computed by taking the temperature-weighted softmax scores for each hallway. The final action probabilities are computed according to Eqn. (6). The robot interacts with the human over $T=200$ environment time steps and predicts the human's intent every $T_z=20$ time steps. Fig. 4 provides a comparison between RCIP and other baseline approaches that employ set-valued prediction.

B. Simulation: Cooperative Navigation in Habitat

Habitat [24] is a photo-realistic simulator containing a diverse set of scenes, objects, and humans models for human robotics tasks. In this experiment, a Boston Dynamics Spot robot and human are jointly tasked with navigating to a set of goal objects in sequence, to simulate cleaning up a house (i.e., grabbing various items, such as crackers, cans of soup, etc). Each scene contains $5-10$ objects of interest. Although the human may initially be out of view of the robot, the robot must find the human and maintain a safe distance of one meter at all times. We simulate the human's decision making by choosing a high-level intent from the set of

Method	$1-\alpha_1$	Plan Succ.↑	Plan Help↓	Step Succ.↑	Step Help↓
RCIP	0.85	0.85	**0.95**	**0.98**	**0.65**
KnowNo [21]	0.85	0.85	0.96	**0.98**	0.66
Simple Set	0.97	0.85	1.00	1.00	1.00
Entropy Set	–	0.66	0.46	0.45	0.06
No Help	–	0.62	0	0.94	0

TABLE II: Results for **Cooperative Manipulation**. The optimal action miscoverage rate is held fixed between RCIP, KnowNo, and Simple Set for comparing the other metrics.

objects; here, $\mathcal{Z}=[N_o]$, where N_o is the number of objects in the scene. The confidence scores for each intended object are computed by taking the temperature-weighted softmax scores for each goal object. The final action probabilities are computed according to Eqn. (6). The robot interacts with the human over $T=600$ environment time steps and selects a new goal object every $T_z=100$ time steps. We present results for cooperative navigation in Table I.

C. Hardware: Cooperative Manipulation

In this example (Fig. 1), each scenario tasks the robot with helping a human to sort a set of objects by inferring the sorting category for each object. We assume that the human's intent set $\mathcal{Z}$ is represented in the (high-dimensional) space of natural language descriptions, such as "the color orange", "children's toys", and "vegetables", and that intents and actions are one-to-one. We first use GPT-4V (`gpt-4-vision-preview`) [1] to process the image by asking for a description of each scene and use a language-only model (`gpt-3.5-turbo`) to rank a set of possible plans via multiple-choice question and answering (MCQA) [25, 26]. The temperature-weighted softmax scores for each bin give the final action probabilities. The robot interacts with the human over $T=8$ environment time steps, and the human selects a new sorting plan every $T_z=1$ time step. We show in Table II that RCIP reduces the plan-wise help rate by 5% and the step-wise help rate by 35% in cooperative manipulation. We use a Franka Emika Panda arm for the robotic manipulation portion of the task.

V. Conclusion

We propose Risk-Calibrated Interactive Planning (RCIP), a framework that applies statistical multi-hypothesis risk control to address the problem of risk calibration for interactive robot tasks. We formalize RCIP as providing a statistical guarantee on an arbitrary number of user-specified risks, such as prediction failures and the amount of human help, subject to a bound on the rate at which the robot fails to predict the optimal actions. By optimizing preferences over a small number of model parameters, RCIP is able to achieve higher flexibility in aligning to user preferences than fixed-paramter methods. Experiments across a variety of simulated and hardware setups demonstrate that RCIP does not exceed user-specified risk levels. Moreover, RCIP reduces user help $5-30\%$ when compared to baseline approaches that lack formal assurances.

References

[1] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, *et al.*, "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.

[2] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone, "Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16*. Springer, 2020, pp. 683–700.

[3] J. Hardy and M. Campbell, "Contingency planning over probabilistic obstacle predictions for autonomous road vehicles," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 913–929, 2013.

[4] W. Zhan, C. Liu, C.-Y. Chan, and M. Tomizuka, "A non-conservatively defensive strategy for urban autonomous driving," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 459–464.

[5] Y. Chen, U. Rosolia, W. Ubellacker, N. Csomay-Shanklin, and A. D. Ames, "Interactive multi-modal motion planning with branch model predictive control," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5365–5372, 2022.

[6] S. H. Nair, V. Govindarajan, T. Lin, C. Meissen, H. E. Tseng, and F. Borrelli, "Stochastic mpc with multi-modal predictions for traffic intersections," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 635–640.

[7] A. Cui, S. Casas, A. Sadat, R. Liao, and R. Urtasun, "LookOut: Diverse multi-future prediction and planning for self-driving," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Oct. 2021.

[8] L. Peters, A. Bajcsy, C.-Y. Chiu, D. Fridovich-Keil, F. Laine, L. Ferranti, and J. Alonso-Mora, "Contingency games for multi-agent interaction," *arXiv preprint arXiv:2304.05483*, 2023.

[9] S. Shi, L. Jiang, D. Dai, and B. Schiele, "Motion transformer with global intention localization and local movement refinement," *Advances in Neural Information Processing Systems*, vol. 35, pp. 6531–6543, 2022.

[10] X. Huang, G. Rosman, I. Gilitschenski, A. Jasour, S. G. McGill, J. J. Leonard, and B. C. Williams, "Hyper: Learned hybrid trajectory prediction via factored inference and adaptive sampling," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 2906–2912.

[11] Z. Zhou, L. Ye, J. Wang, K. Wu, and K. Lu, "Hivt: Hierarchical vector transformer for multi-agent motion prediction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 8823–8833.

[12] S. Liu, P. Chang, Z. Huang, N. Chakraborty, K. Hong, W. Liang, D. L. McPherson, J. Geng, and K. Driggs-Campbell, "Intention aware robot crowd navigation with attention-based interaction graph," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 12 015–12 021.

[13] P. Agand, M. Taherahmadi, A. Lim, and M. Chen, "Human navigational intent inference with probabilistic and optimal approaches," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 8562–8568.

[14] M. F. A. R. D. T. (FAIR)†, A. Bakhtin, N. Brown, E. Dinan, G. Farina, C. Flaherty, D. Fried, A. Goff, J. Gray, H. Hu, *et al.*, "Human-level play in the game of diplomacy by combining language models with strategic reasoning," *Science*, vol. 378, no. 6624, pp. 1067–1074, 2022.

[15] J. Gu, C. Sun, and H. Zhao, "Densetnt: End-to-end trajectory prediction from dense goal sets," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 15 303–15 312.

[16] V. Vovk, V. Fedorova, I. Nouretdinov, and A. Gammerman, "Criteria of efficiency for conformal prediction," in *Conformal and Probabilistic Prediction with Applications: 5th International Symposium, COPA 2016, Madrid, Spain, April 20-22, 2016, Proceedings 5*. Springer, 2016, pp. 23–39.

[17] V. Vovk, "Conditional validity of inductive conformal predictors," in *Asian conference on machine learning*. PMLR, 2012, pp. 475–490.

[18] M. Sadinle, J. Lei, and L. Wasserman, "Least ambiguous set-valued classifiers with bounded error levels," *Journal of the American Statistical Association*, vol. 114, no. 525, pp. 223–234, 2019.

[19] K. Stankeviciute, A. M Alaa, and M. van der Schaar, "Conformal time-series forecasting," *Advances in neural information processing systems*, vol. 34, pp. 6216–6228, 2021.

[20] K. J. Strawn, N. Ayanian, and L. Lindemann, "Conformal predictive safety filter for rl controllers in dynamic environments," *arXiv preprint arXiv:2306.02551*, 2023.

[21] A. Z. Ren, A. Dixit, A. Bodrova, S. Singh, S. Tu, N. Brown, P. Xu, L. Takayama, F. Xia, J. Varley, *et al.*, "Robots that ask for help: Uncertainty alignment for large language model planners," *arXiv preprint arXiv:2307.01928*, 2023.

[22] S. Bates, A. Angelopoulos, L. Lei, J. Malik, and M. Jordan, "Distribution-free, risk-controlling prediction sets," *Journal of the ACM (JACM)*, vol. 68, no. 6, pp. 1–34, 2021.

[23] J. Lekeufack, A. A. Angelopoulos, A. Bajcsy, M. I. Jordan, and J. Malik, "Conformal decision theory: Safe autonomous decisions from imperfect predictions," *arXiv preprint arXiv:2310.05921*, 2023.

[24] X. Puig, E. Undersander, A. Szot, M. D. Cote, T.-Y. Yang, R. Partsey, R. Desai, A. W. Clegg, M. Hlavac, S. Y. Min, *et al.*, "Habitat 3.0: A co-habitat for humans, avatars and robots," *arXiv preprint arXiv:2310.13724*, 2023.

[25] A. Srivastava, A. Rastogi, A. Rao, A. A. M. Shoeb, A. Abid, A. Fisch, A. R. Brown, A. Santoro, A. Gupta, A. Garriga-Alonso, *et al.*, "Beyond the imitation game: Quantifying and extrapolating the capabilities of language models," *arXiv preprint arXiv:2206.04615*, 2022.

[26] D. Hendrycks, C. Burns, S. Basart, A. Zou, M. Mazeika, D. Song, and J. Steinhardt, "Measuring massive multitask language understanding," *arXiv preprint arXiv:2009.03300*, 2020.