Interaction-Aware Planning via Nash Equilibria for Manipulation in a Shared Workspace

Shray Bansal[1], Mustafa Mukadam[1] and Charles L. Isbell[1]

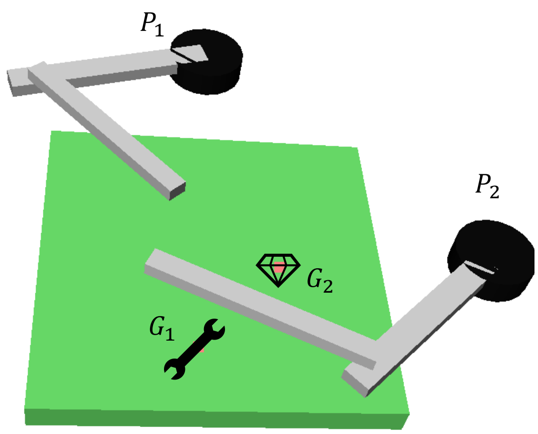

Fig. 1: Pick-and-Place scenario studied in this work. Here, two independently controlled agents P_1 and P_2 are each assigned the task of reaching an object, G_1 and G_2 respectively, and then to move back to the initial position. The agents aim to complete it without collisions. Actions of one agent can severely limit the possible actions for the other in this small shared workspace.

***Abstract*— Our goal is to control a robot to act efficiently with other intentional agents (robots or humans) to successfully perform manipulation tasks in a shared space. We model this as a noncooperative game between the robot and other agents. Our decentralized planner finds actions that achieve a Nash Equilibrium and then iteratively replans based on the observations of the other agents. We test this approach on a pick-and-place task with two agents in a 2D workspace with the goal to minimize task completion time and collisions. Our results show that agents acting in consideration of the goals and interaction-awareness of others achieve higher efficiency as well as improved safety.**

I. Introduction

Situations involving multiple agents acting independently are common in everyday life. Most of these do not require pure cooperation or competition, but a shared resource like workspace or tools can lead to interaction. For *e.g.*, in driving, people have independent destinations and a common goal to avoid collisions. The shared resource here is the road, which in situations like lane changing in traffic leads to complicated negotiations due to conflicting optimal paths for the agents. Humans frequently perform such activities with high success by considering not only the state, but also the goals and preferences of the other agents. Robots are usually confined to acting in relative isolation and lack the capability to efficiently perform tasks in the presence of mutually-interfering independent agents.

We study a manipulation scenario in this paper inspired by activities like parts assembly in factories. Two agents perform pick-and-place in close proximity as depicted in Figure I. Here, actions of any one agent can limit the ability of the other in achieving its goal. Our aim is to design a planner to enable a robot to efficiently complete this task with other agents without having to explicitly communicate its plans. Our experiments include two agents but our approach is designed for any set of rational agents (including humans).

Work in Human-Robot Interaction (HRI) has studied the close interaction between independent agents in manipulation activities and our problem is modeled similarly to those in [1], [2]. While the human fluency in interactive scenarios is attributed to their modeling of others as intentional agents, the focus in HRI has been to build predictive models of the human and plan trajectories to avoid collision [1]. We want our approach also to consider the others as rational agents acting to achieve their goals while avoiding interference.

Game Theory provides us with a framework for modeling the decision-making of rational agents [3]. Recently, it has been used in HRI for high-level task planning to decide the order of objects to pick in a pick-and-place task by [2]. [4] produced human-like motions in navigation amongst crowds by planning with Nash Equilibria (NE). Inspired by their success, we model our problem as a finite noncooperative game between two agents and plan actions that achieve a NE. We do this by first sampling a set of plans for each agent independently and evaluating their cartesian product to find the NE. The corresponding action for our agent from the NE solution is performed and then iterative replanning after observing the others' actions provides the next NE plan.

We make the following contributions.

- Introduce an interaction-aware algorithm for planning in close-proximity manipulation tasks.
- Demonstrate the utility of considering agents' goals and preferences during planning through experiments in the two agent pick-and-place scenario.

II. PROBLEM

Each agent is assigned with a task of reaching an object on the table and then returning to their initial position while avoiding collisions. We formulate this as an N-player game G, according to the tuple, $G = (P, A, c)$ [3]. Where $P = \{P_1, ..., P_N\}$ is a finite set of N players, $A = A_1 \times ... \times A_N$ where A_i is the set of actions available to player i, also each vector $a \in A$ where $a = (a_1, ..., a_N)$ will be referred to as an action profile and $c = (c_1, ..., c_N)$ contains the real-valued cost corresponding to each agent, $c_i : A \mapsto \mathbb{R}$.

In our scenario, each agent P_i is a robot, each action set A_i is the set of valid trajectories, each a sequence of jointspace positions and velocities, and the cost $c_i(a)$ is equal to ∞ for collisions and the task completion time t otherwise.

III. APPROACH

The goal of an agent is to find an action that minimizes its cost c_i. Since c_i is parameterized by an action profile including all agents, we compute the NE instead of optimizing independently. This gives us an action profile a^{Nash} such that no single agent can achieve higher reward by unilaterally altering their action,

$$a_i^{Nash} = \arg\min_{a_i} c_i(a_1^{Nash}, ., a_i, ., a_N^{Nash}) \quad \forall i \in N. \quad (1)$$

By finding a_i in this way we implicitly consider other agents as acting rationally to achieve their goals. Since agents are decentralized, the actions of other agents can be considered as predictions of $a_{j \neq i}$ by P_i conditioned on a_i.

The finite set A_i is constructed for each agent by planning k collision-free trajectories in the static environment since the dynamic agents are considered through the NE. Doing so for all agents gives us k^N action profiles from which we find a^{Nash} according to Equation 1. We are guaranteed at least one NE [3]. However, in case more are found, we select only Pareto-optimal equilibria from the set a^{Nash} and choose one uniformly at random. From this action profile, the actions for i are executed for a fixed time and then the procedure is repeated after observing the actions of the other agents until the agent has reached its goals or the time limit is reached.

IV. EXPERIMENTS

We use OpenRave[1] to simulate the robots and the Bi-directional RRT from OMPL[2] for generating trajectories. We sample 5 plans for each agent, and include a copy at $0.5\times$ speed and a no-action plan, $|A_i| = 5 \times 2 + 1$. We chose the placement of agents and goals to induce interaction between optimal plans. The simulation time-step is $0.01s$ and replanning is performed every $0.5s$. Each trial has a time-limit of $10s$. To ensure safe operation, actions leading to agents being within $1cm$ of each other cause a *Safety Stop*.

In the first experiment we conducted 5 trials each with both robots being controlled by different algorithms. **Unreactive** picks an action profile a that minimizes its cost $c_i(a)$ while assuming other agents as being fixed, **Nash** picks a randomly from the Pareto-optimal set of NE, **Cooperative** picks a that minimize the sum of costs for over agents, *i.e.*$\sum_{i=1}^{N} c_i(a)$. Table I shows the mean and standard deviation of the time for task completion by the first(t_1) and second(t_2) robots. We see that both agents take similar time for the same algorithm ($t_1 \approx t_2$) and *Unreactive* takes the most time, while *Cooperative* takes the least and *Nash* is in the middle. The *Unreactive* agent also triggered a large number of safety stops while the other two did not. The *Unreactive* ignores the other agent altogether, while *Cooperative* assumes both agents act in concert for a common goal. Thus, these results seem to affirm our hypothesis that more knowledge about

[1] http://openrave.org/
[2] http://ompl.kavrakilab.org

	t_1	t_2	Safety Stops
Unreactive	$4.9s \pm 2.0$	$4.8s \pm 1.9$	$23.4s \pm 12.2$
Nash	$2.8s \pm 0.8$	$2.7s \pm 0.9$	$0.0s \pm 0.0$
Cooperative	$2.1s \pm 0.3$	$2.2s \pm 0.2$	$0.0s \pm 0.0$

TABLE I: Interaction between agents of the same kind.

	t_1	t_2	Safety Stops
Selfish	$8.4s \pm 3.2$	$8.3s \pm 3.3$	$9.0s \pm 5.5$
Random	$2.8s \pm 0.8$	$2.7s \pm 0.9$	$0.0s \pm 0.0$
Coordinated	$2.4s \pm 0.8$	$2.3s \pm 0.9$	$0.0s \pm 0.0$

TABLE II: Different strategies for selecting equilibria.

the other agents lead to safer, more efficient interactions. We believe the **Cooperative**'s coordination assumption is unreasonable for independent agents and that *Nash* is able to strike a good balance by assuming that the agents act rationally and that they may not have a common cost.

Table II compares three strategies for selecting amongst Pareto-optimal NE. Similar to above, **Random** selects one uniformly at random, while **Selfish** chooses the equilibrium that minimizes its own cost function, and **Coordinated** chooses an equilibrium at random but assigns it to both agents. The *Coordinated* strategy performs the best with *Random* a little worse and *Selfish* being significantly worse. These results indicate that strategies leading to agents selecting the same equilibria are better. This is guaranteed for *Coordinated* and is probable but less likely for *Random* and least likely for *Selfish*. Higher completion times and more safety stops for agents optimizing selfish costs might be due to conflicting optimal plans in this problem. Here, for an action profile to succeed one of the agents will have to compromise on their individual cost. Since the *Coordinated* NE approach is not possible with independent agents, these results indicate that inferring the NE during the interaction should lead to better performance than choosing it randomly. We plan to pursue this in future work.

V. DISCUSSION AND LIMITATIONS

Although, preliminary results support the utility of planning while considering the rational behavior of other agents, more extensive experiments including planning in 3D scenarios is necessary to make the claim stronger. Presently, our approach assumes knowledge of the other agents' goals. However, this assumption is limiting and future work should consider a distribution over goals updated with task progression. Other kinds of agents also should be compared, including those using predictive models to test the importance of the mutual-influenceability in this work.

REFERENCES

[1] J. Mainprice, R. Hayne, and D. Berenson, "Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 897–908, 2016.

[2] V. Gabler, T. Stahl, G. Huber, O. Oguz, and D. Wollherr, "A game-theoretic approach for adaptive action selection in close proximity human-robot-collaboration," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 2897–2903.

[3] K. Leyton-Brown and Y. Shoham, "Essentials of game theory: A concise multidisciplinary introduction," *Synthesis lectures on artificial intelligence and machine learning*, vol. 2, no. 1, pp. 1–88, 2008.

[4] A. Turnwald and D. Wollherr, "Human-like motion planning based on game theoretic decision making," *International Journal of Social Robotics*, vol. 11, no. 1, pp. 151–170, 2019.