Designing Algorithms for Socially Competent Robotic Navigation

Christoforos I. Mavrogiannis
Sibley School of Mechanical & Aerospace
Engineering
Cornell University
Ithaca NY, USA
cm694@cornell.edu

Ross A. Knepper
Department of Computer Science
Cornell University
Ithaca NY, USA
rak@cs.cornell.edu

ABSTRACT

Despite the great progress in robotic navigation in the past decades, navigating a human environment remains a hard task for a robot, due to the lack of formal rules guiding traffic, the lack of explicit communication among agents and the unpredictability of human behavior. Inspired by the efficiency of human navigation, we employ the insights of sociology studies on pedestrian behavior and psychology studies on action interpretation to design an online planning framework that leverages the power of implicit communication to generate legible robot behaviors in pedestrian environments. The foundation of our approach is a novel topological representation, based on braid groups. Preliminary results demonstrate the efficiency of our approach in simulation, whereas planned experiments with human subjects are expected to enable us to extract realistic predictive models and get user feedback. Finally, we plan on evaluating our approach by running our algorithms on our social robot (Fig. 1) in crowded environments.

1. INTRODUCTION

Despite the great success of existing approaches in simulating realistically pedestrian flows (e.g. [8]) and the development of effective algorithms for robotic navigation in dynamic environments (e.g. [9]), robots still tend to confuse humans and draw undesired attention. In particular, the robot motion is often hard to read, resulting in unpredictable human reactions to which the robot in turn reacts to, contributing to an oscillatory joint behavior that hinders humans' paths. We argue that the root of the problem lies in the failure from the robot's part to convey consistently its intentions to human observers. To address this issue, we

Figure 1: Suitable Tech. Beam Pro.

draw from studies on pedestrian behavior [10] and action interpretation [2] and employ a topological representation, based on braid groups [1], to develop a principled algorithm design for closed-loop, intent expressive navigation in crowded environments. Our approach leverages the power of implicit communication, which we argue is of particular importance for human-robot tasks [4].

2. MODELS & METHODS

2.1 Human-Inspired Inference Mechanism

According to Wolfinger [10], the social order in pedestrian environments is established through the *Pedestrian Bargain*, a set of foundational rules regulating human behavior: (1) *people must behave like competent pedestrians* and (2) *people must trust copresent others to behave like competent pedestrians*. The bargain requires an agreement on what is considered as *competent* behavior. This is ensured through context-specific inference mechanisms that humans develop from experience, as they grow up. Csibra and Gergely [2] pointed out that these mechanisms are *teleological* in nature, i.e., humans tend to interpret observed actions as goal-directed, by attributing potential *goals* to observed *actions*. This observation is particularly interesting for robotics applications in human environments, where it is important to incorporate an understanding of the effect of robot actions to human observers at the planning stage (see e.g. Dragan and Srinivasa [3]).

Inspired by the aforementioned works, our approach contributes a principled inference mechanism that associates observed *actions* with potential *goals* in a given multi-agent context. This mechanism allows an agent i to form a belief

$$bel_i = P(G|A, M) \tag{1}$$

where G is a goal from a set of goals $\mathcal{G}$, A is the sequence of actions of all agents so far and M is the context of the scene. By context M, we refer to information that is explicitly or implicitly (via secondary inference mechanisms) available to all agents (e.g. the map, agents' destinations etc), whereas A can be intuitively defined as the collection of all agents' trajectories so far. The definition of G is not straightforward in a multi-agent setting, as it should reflect the superposition of all agents' intentions. We model G by taking a topological perspective, employing the formalism of *braids* [1] to characterize the collective behavior of multiple agents.
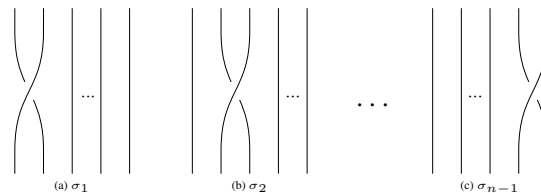

Figure 2: The generators of the Braid Group $\mathcal{B}_n$.

2.2 Topological Model of Collective Behavior

Braids are topological objects with geometric and algebraic representations that form a group $\mathcal{B}_n$, which can be generated by a set of $n-1$ primitive elements, called generators (shown in Fig. 2), denoted as $\sigma_1, ..., \sigma_{n-1}$. The generators, along with their inverses

can synthesize arbitrarily long and complex patterns through the group composition operation. Each braid can also be algebraically represented as a word, synthesized as a sequence of the symbols of the elements that produced it.

In this work, we model the set of potential goals $\mathcal{G}$ to be the braid group $\mathcal{B}_n$. As the agents avoid each other on their way to their destinations, their trajectories form a spatiotemporal pattern. This pattern carries quantitative properties such as velocities
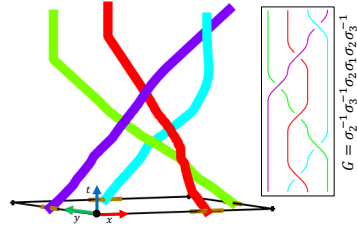
Figure 3: A collection of trajectories (left) and a *joint strategy* (right), represented as a braid diagram and a braid word, upon projecting the trajectories onto the x-t plane.

and distances but also qualitative properties, such as right/left agent passings. For our purposes, those qualitative properties are particularly significant, as they represent an abstract form of a *joint strategy* in which all agents engaged to resolve collisions. We make use of the braid formalism to characterize topologically this joint strategy: upon projecting a collection of trajectories to a selected space-time plane, we label any trajectory crossings that emerge, as elementary braids and form a corresponding braid word-goal G (see Fig. 3). Conversely, given agents' current locations and a prediction of their destinations, we may represent the pattern that their future trajectories will form (joint strategy) as a braid word.

2.3 Planning Framework

We have introduced two different approaches that make use of our model of collective behavior and inference mechanism to generate robot motion that implicitly communicates intent in a multi-agent context, in discrete and continuous configuration settings.

In a discretized setting, we have presented a utility-based decision making mechanism that enables the robot to incorporate its understanding regarding the emerging joint strategy to its action selection process [5]. The utility function comprises an individual efficiency term and a consensus term (the entropy of the distribution $P(G|A, M)$), quantifying the status of the consensus with respect to a joint strategy-braid among all agents. The robot thus decides on an action by compromising between progress to destination and communication of compliance with a joint strategy. We tested this framework in simulation by considering different scenarios involving multiple agents navigating a discretized workspace (see Fig. 4a) and demonstrated that by employing the consensus term, the agents were able to agree on a joint strategy much faster. For more details see [5].

In a continuous setting, we have presented a planning framework, based on trajectory optimization, for generating trajectories of selected topological braiding [6, 7]. The framework allows an agent to score a set of candidate joint strategies, with respect to their likelihood of occurrence, generate smooth corresponding geometric representations for each one and conclude to a final trajectory that best communicates the agent's understanding with respect to the emerging joint strategy. We tested the approach by considering scenarios involving multiple agents navigating a common workspace (see Fig. 4b) and showed that our method accelerates the convergence to a joint strategy. For more details, see [6].

3. FUTURE WORK

We are currently working on several research directions. Our main focus involves learning a humanlike model of $P(G|A, M)$
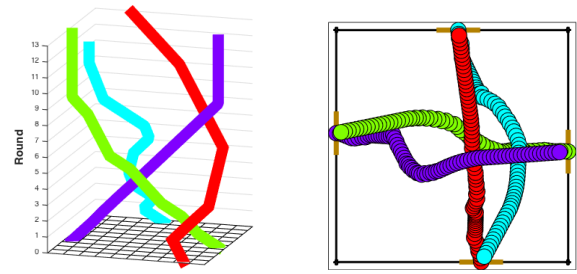
(a) Trajectories extracted by running our algorithm from [5].

(b) Swept volumes extracted by running our algorithm from [6, 7].

Figure 4: Simulations extracted by running our planning algorithms.

from human demonstrations. To this end, we plan on conducting lab experiments, recording subjects' trajectories with our motion capture system. Another direction involves generalizing our approach to more realistic problem settings and optimizing our algorithm for real-time operation. Furthermore, we are working on the development of our experimental setup: a Suitable Technologies Beam Pro Telepresence robot (Fig. 1), which we have augmented with on-board computation and a camera for tracking pedestrians in real-time. Our goal is to use this platform to test the efficiency of our algorithms around humans and study their effect on pedestrians and on users of telepresence systems.

Acknowledgment

References

[1] J. S. Birman. *Braids Links And Mapping Class Groups*. Princeton University Press, 1975.

[2] G. Csibra and G. Gergely. 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124(1):60–78, 2007.

[3] A. D. Dragan and S. Srinivasa. Integrating human observer inferences into robot motion planning. *Auton. Robots*, 37(4): 351–368, 2014.

[4] R. A. Knepper, C. I. Mavrogiannis, J. Proft, and C. Liang. Implicit communication in a joint action. In *12th ACM international conference on Human-robot interaction*, 2017.

[5] C. I. Mavrogiannis and R. A. Knepper. Decentralized multi-agent navigation planning with braids. In *Workshop on the Algorithmic Foundations of Robotics*, 2016.

[6] C. I. Mavrogiannis and R. A. Knepper. Towards socially competent navigation of pedestrian environments. In *Robotics: Science and Systems 2016 Workshop on Social Trust in Autonomous Robots*, 2016.

[7] C. I. Mavrogiannis and R. A. Knepper. Interpretation and communication of pedestrian intentions using braid groups. In *Human Robot Interaction Workshop on Intention Recognition in HRI*, 2016.

[8] M. Moussaïd, D. Helbing, and G. Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proc. of the Nat. Academy of Sciences*, 108(17):6884–6888.

[9] 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. of Robotics Res.*, 34(3):335–356, 2015.

[10] N. H. Wolfinger. Passing Moments: Some Social Dynamics of Pedestrian Interaction. *Journal of Contemporary Ethnography*, 24(3):323–340, 1995.