

Interpretation and Communication of Pedestrian Intentions Using Braid Groups

Christoforos I. Mavrogiannis and Ross A. Knepper

Abstract— We present a novel framework for socially competent pedestrian navigation based on understanding pedestrians’ intentions and planning intent-expressive robot motion. We model pedestrians’ intentions as combinations of intended topological routes and intended destinations. The core of this approach is a novel topological representation of a pedestrian scene, based on braid groups. This representation is used as a basis for the development of an online navigation algorithm, designed according to conclusions of recent psychology studies on action interpretation and sociology studies on human pedestrian behavior. Simulation results demonstrate the potential of our approach for use in real world pedestrian environments.

I. INTRODUCTION

Typical pedestrian environments are characterized by a high level of uncertainty, imposed by the lack of formal rules to control traffic and the lack of explicit communication among pedestrians. Nonetheless, humans are capable of traversing pedestrian workspaces with remarkable efficiency without hindering each other’s paths (see Fig. 1). Interacting pedestrians relax uncertainty by mutually agreeing on a joint motion plan. This joint motion plan is the result of *negotiation* taking place among them via implicit communication of their intentions through their motion.

Sociology studies [18] refer to this type of negotiation as the *pedestrian bargain* and specify that it is based on two foundational principles: (1) *people must behave like competent pedestrians* and (2) *people must trust copresent others to behave like competent pedestrians*. The enforcement of the pedestrian bargain imposes a form of trust among pedestrians, which is the foundation of socially competent behavior. Consequently, engineering the pedestrian bargain is a key for generating socially competent robot motion in pedestrian environments. Towards this goal, Knepper and Rus [11] proposed a navigation algorithm enabling robots to behave as competent pedestrians. In this work, we aim at moving a step further: beyond generating human-like robot motion, we also incorporate in the robot’s motion plan the trust that others will behave as competent pedestrians.

This requires answering the following two questions: (1) how is social competence defined in a pedestrian environment? and (2) how do pedestrians use this notion of competence to interpret observed behaviors? We approach pedestrian competence as a combination of three features: a) moving efficiently towards intended destinations, b) following trajectories that comply with socially acceptable standards of motion and c) clearly communicating intentions



Fig. 1: Smooth collision avoidance in a hallway.

to others. Regarding interpretation of observed behavior, our work based on the conclusions of recent psychology studies on the mechanisms of action interpretation [5], suggesting that humans tend to adopt a teleological inference mechanism to attribute *goals* to observed *actions*. We also make use of the relevant framework of Dragan and Srinivasa [6] for incorporating observer inferences into motion planning.

In order to incorporate the aforementioned principles in robot motion, we need a proper model of the pedestrian scene. The evolution of a pedestrian scene, from the beginning to the end, is the result of several pairwise *negotiations* among all agents. We argue that the main component of a *negotiation* between an interacting pair of agents concerns the decision over the passing sides (right or left?) that the agents will be following. Therefore, reaching to a decision over passing sides, requires that each agent (1) clearly expresses its *intention* of passing side and (2) correctly identifies the *intention* of passing side of each other. At the same time, it is important that each agent has a global understanding of the scene, involving all future interactions that it might have.

To properly capture and encode all these features, we derive a topological representation of the pedestrian scene, based on braid groups [4]. This representation allows us to symbolically characterize the interactions among agents and consequently globally and collectively characterize a pedestrian scene from start to end. We demonstrate how braids

can be used by robots to express their own intentions and infer the intentions of others (either humans or robots) and how this model can be used to generate socially competent behavior in a pedestrian context. The core of our approach is inherently generalizable to any workspace with any number of agents.

II. RELATED WORK

Initial attempts in the area of pedestrian navigation, typically treating other agents as obstacles, have resulted in simulations of a variety of pedestrian environments and scenarios (e.g. [8, 15]). However collision avoidance is not sufficient in the real world setting; human navigation is a much more complex process, comprising decisions related to human comfort and context-specific social conventions. Sisbot et al. [13] identified the lack of such components in existing planners and presented a costmap-based planner, incorporating costs for safety, visibility and surprise. Nonetheless, without a model of pedestrian cognitive processes, it is impossible to realistically anticipate pedestrian behaviors or pedestrians' reactions to robot motion.

To address the need for realistically modeling pedestrian decision making, roboticists have lately employed data-driven techniques in an effort to learn from humans and perform online trajectory prediction in order to plan in an informed way [3, 9, 16, 19]. Such approaches were engineered towards imitating observed pedestrian behavior in specific, well-defined contexts, but lack underlying cognitive models of pedestrians, particularly the mechanisms that guide their interactions, thus failing to provide a robust, generalizable solution. To this end, the works of [2, 7] focused on learning models for predicting pedestrians' intentions in urban environments to plan safe motion for autonomous cars. However, navigating a car on streets is a highly structured problem compared to pedestrian navigation, as the former can leverage the existence of well established, formal rules and structured signals. Pedestrian interactions have not evolved such rigid rules and structure, thus a more sophisticated infrastructure is needed to identify pedestrian intentions.

A key to simplifying our problem is the concept of cooperation among pedestrians. A few works have proposed approaches that explicitly leverage the fact that human agents tend to engage in joint cooperative collision avoidance [10–12, 17]. Our work is conceptually close to this class of works in that it models pedestrian interactions. We explicitly make use of conclusions from sociology regarding pedestrian navigation and psychology studies regarding action interpretation to model pedestrian cognitive processes and incorporate them into the motion planning. We do that based on a novel pedestrian scene representation based on braid groups that symbolically encodes topologically distinct classes of global pedestrian scene evolutions. The braid representation, unlike most existing models for pedestrian navigation, offers an inherent generalization to different environments and arbitrary numbers of agents and allows us to simultaneously reason about a set of topologically distinct

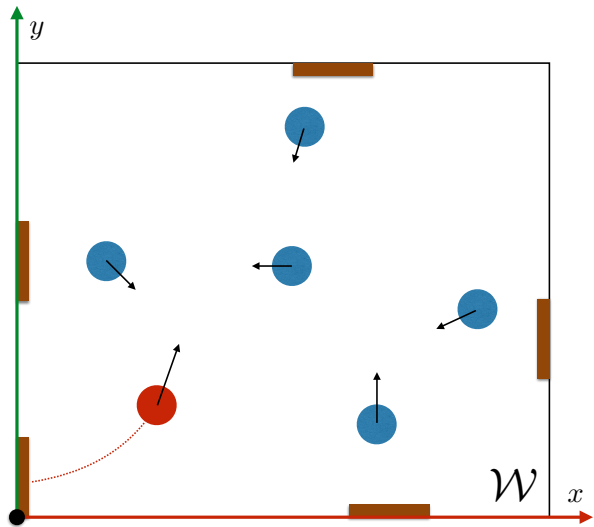


Fig. 2: A workspace \mathcal{W} with a robot (red color) and a set of human pedestrians (shown in blue color), each moving towards its intended destination (one of the doors shown in brown color). In such a scenario, the robot needs to gain a global understanding of the scene, which is equivalent to inferring the *pedestrian scenario*.

scene evolutions. This way enables us to avoid committing to a single imperfect prediction and achieve smooth, socially competent navigation.

III. FOUNDATIONS

A. Pedestrian Scene Model

Consider a set \mathcal{A} of n agents moving in a pedestrian workspace $\mathcal{W} \subset \mathbb{R}^2$. Each agent $a \in \mathcal{A}$ moves from an initial position $s_a \in \mathcal{W}$ towards an intended destination $d_a \in \mathcal{D}$, where $\mathcal{D} \subset \mathcal{W}$ is a finite set of possible destinations in the scene (see Fig. 2). Let us collect the starting locations and the intended destinations of all agents in the tuples $\Sigma = \langle s_1, \dots, s_n \rangle$ and $\Delta = \langle d_1, \dots, d_n \rangle$ respectively. Denote by $t \in I = [0, 1]$ a normalized timing parameter, with $t = 0$ corresponding to the beginning of observations and $t = 1$ to the time when the last agent is reaching its destination.

Define the state of each agent $a \in \mathcal{A}$ at time t by $q_a(t) \in \mathcal{W}$. Its trajectory is a continuous function $\xi_a : I \rightarrow \mathcal{W}$ with $\xi_a(0) = s_a$ and $\xi_a(1) = d_a$, lying in a corresponding Hilbert space of trajectories Ξ .

Likewise, we represent the state of the system of all n agents as a tuple $Q = \langle q_1, q_2, \dots, q_n \rangle \in \mathcal{C}_n$, where the configuration space $\mathcal{C}_n = \mathcal{W}^n \setminus E$ with $E = \{Q : \|q_i - q_j\| \leq \epsilon\}$ containing the set of system states in collision. The system trajectory is a continuous function $\zeta : I \rightarrow \mathcal{C}_n$ with $\zeta(0) = \Sigma$ and $\zeta(1) = \Delta$, lying in a corresponding Hilbert space of trajectories \mathcal{Z} . As the system of agents is traversing the workspace from Σ to Δ , each one of them makes discrete decisions of passing from the right or left side of each other to avoid leading the system to a state in E . These decisions are reflected in their trajectories, which entangle in the space-time domain, forming a pattern β .

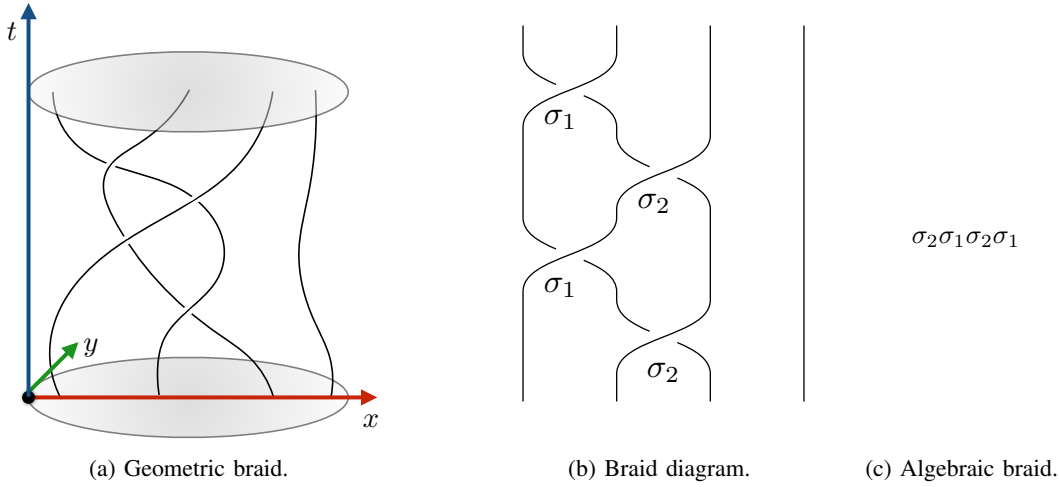


Fig. 3: Braid representations.

B. Braid Theory Primer

We abstract a pedestrian scene by defining a *Pedestrian Scenario* (see Fig. 2) to be a tuple $S = \langle \Sigma, \Delta, \beta \rangle \in \mathcal{S}$, where \mathcal{S} is the set of all possible Pedestrian Scenarios that can be defined in a given scene. We refer to Σ as the *Initialization Scenario*, Δ as the *Destination Scenario* and β as the *Entanglement Pattern*. This modeling decision is motivated by the observation that the main feature that qualitatively describes a collection of pedestrian trajectories is the strategy (right or left) that each agent employs to avoid each other on its way to an intended destination. In this paper, we formally model entanglement patterns β using braid theory [1, 4]. The next subsections introduce the fundamentals of braid theory and connect them with our pedestrian scene model.

A braid on $n \geq 1$ strands is a system of n curves embedded in \mathbb{R}^3 , called the strands of the braid, such that each strand i intersects each plane $\{x, y, t'\}$ only once for any $t' \in I$ and connects the point $(i, 0, 0)$ with the point $d_{p(i)} = (p(i), 0, 1)$, where the sequence $p(1), \dots, p(n)$ is a permutation of the set $\{1, 2, \dots, n\}$.

A geometrical braid representation is commonly referred to as a *geometric braid* (see Fig. 3a). Any geometric braid can be represented with a *Braid Diagram*. A braid diagram (see Fig. 3b) is a projection of a geometric braid to $\mathbb{R} \times \{0\} \times I$, including indications of which string goes “under” the other at each crossing.

The set of all braids on n strands form a group \mathcal{B}_n . The group \mathcal{B}_n is generated from a set of $n-1$ elementary braids, called the generators of \mathcal{B}_n (see Fig. 4). A generator σ_i , $i \in \{1, 2, \dots, n-1\}$ can be described as the pattern that emerges upon exchanging the i -th string (counted from left to right) with the $(i+1)$ -th string, such that the left string passes over the right. For any two braids $B_1, B_2 \in \mathcal{B}_n$, the group operation $B_1 \cdot B_2$ can be described as a concatenation of the braids, resulting in B_2 being placed at the bottom of B_1 , by attaching the top endpoints of B_2 to the bottom endpoints of B_1 and shrinking each braid by a factor of 2, along the t axis. The following axioms are also satisfied, completing the group definition:

- **Closure:** The product $B_1 \cdot B_2$ is a new braid in \mathcal{B}_n .
- **Associativity:** for any braids $B_1, B_2, B_3 \in \mathcal{B}_n$, $(B_1 \cdot B_2) \cdot B_3 = B_1 \cdot (B_2 \cdot B_3)$.
- **Identity Element:** The identity braid corresponds to the trivial case of a braid without any strand crossings, with each strand i , starting from the point $(i, 0, 0)$ and ending at the point $(i, 0, 1)$.
- **Inverse Element:** The inverse of B_i , denoted as B_i^{-1} , is defined by reversing the order of all crossings in B_i .

Any braid can be written as a product of generators and their inverses. This product is commonly referred to as *braid word* and this form of representation as an *algebraic braid*. For an example of transitioning among braid representations, see Fig. 3: the generators that composed the geometric braid in Fig. 3a are denoted on the braid diagram depicted in Fig. 3b and the corresponding braid word is shown in Fig. 3c.

C. Entanglement Patterns as Elements of Braid Groups

We adopt the aforementioned braid representations to characterize collections of pedestrian trajectories. In particular, we model the entanglement pattern of a trajectory collection as a braid word composed of generators describing the crossings taking place between all pairs of neighboring pedestrian trajectories, upon their projection onto a selected plane. We represent an entanglement pattern as a tuple $\beta = \langle b_1, b_2, \dots, b_K \rangle$, where $b_k \in \mathbb{Z}$, $k \in \{1, \dots, K\}$ is an integer describing the k -th crossing as time increases (e.g. the integer $-i$ is used to denote the generator σ_i^{-1}), while K is the total number of crossings from the start to the end of time.

For example, Fig. 5 depicts a set of trajectories of five agents moving towards their intended destinations along with their entanglement pattern and corresponding algebraic braid (upon projection onto the $x-t$ plane). The algebraic representation of this braid can be found to be the word $\sigma_2 \sigma_3^{-1} \sigma_4^{-1} \sigma_1^{-1} \sigma_2^{-1} \sigma_3^{-1} \sigma_4^{-1}$ and consequently the entanglement pattern for this scenario can be represented as $\beta = \langle 2, -3, -4, -1, -2, -3, -4 \rangle$.

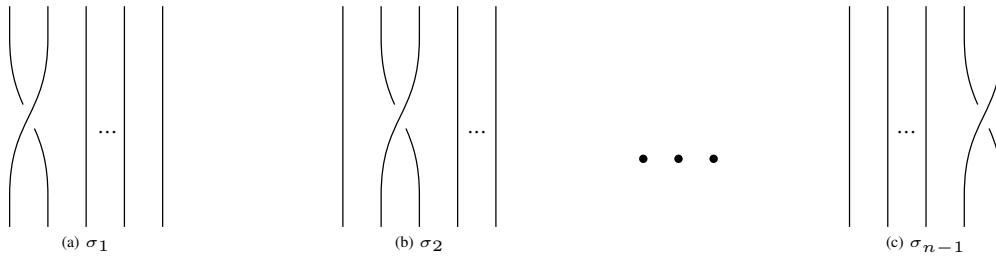


Fig. 4: The generators of the Group \mathcal{B}_n .

IV. INFERENCE OF PEDESTRIAN INTENTIONS

The pedestrian bargain [18] is based on a notion of competence, which determines pedestrians' behaviors but also shapes their expectations from others. We model pedestrian competence to be the result of three different motion specifications: 1) move efficiently towards intended destinations, 2) follow trajectories that comply with socially acceptable standards of motion and 3) clearly express own intentions. At the same time, pedestrians expect others to behave competently, according to the same definition of competence. In particular, pedestrians associate observed past behaviors with potential future behaviors. The mechanisms under which humans tend to interpret observed actions have lately been studied from a psychology perspective by Csibra and Gergely [5], who highlighted the tendency of humans to interpret any observed *action* as *goal-directed*. In the context of a pedestrian environment, we consider a trajectory collection ζ to represent the *action*, while a *pedestrian scenario* S to be the *goal*.

In this respect, we model a pedestrian's inference as a prediction of a pedestrian scenario S based on an observed collection of trajectories ζ , i.e., $P(S|\zeta)$ as:

$$P(S|\zeta) = P(\Sigma, \Delta, \beta|\zeta). \quad (1)$$

Upon applying the multiplication rule, and given that Σ and ζ are known, we have:

$$P(S|\zeta) = P(\Delta|\zeta)P(\beta|\Delta, \zeta), \quad (2)$$

where the distributions $P(\Delta|\zeta)$ and $P(\beta|\Delta, \zeta)$ represent the pedestrian's beliefs over the emerging *destination scenario* Δ (i.e., where are all the agents going?) and the *entanglement pattern* β (i.e., how will they get there?) respectively. Therefore $P(S|\zeta)$ represents a pedestrian's belief over the ensemble of all agents' *intentions*. Thus, prediction of overall human or robot pedestrian motion is tantamount to modeling just these two distributions, with countably-many outcomes. Since humans are already modeling these distributions, we can leverage them to generate intent-expressive robot behavior.

V. MOTION GENERATION IN PEDESTRIAN ENVIRONMENTS

In order to generate motion in pedestrian environments, we employ the aforementioned inference mechanism and incorporate it in the motion planning. A principled way of integrating human observer inferences in motion planning was recently proposed by Dragan and Srinivasa [6], who, motivated by the conclusions of Csibra and Gergely [5], mathematically formalized the properties of *predictability*

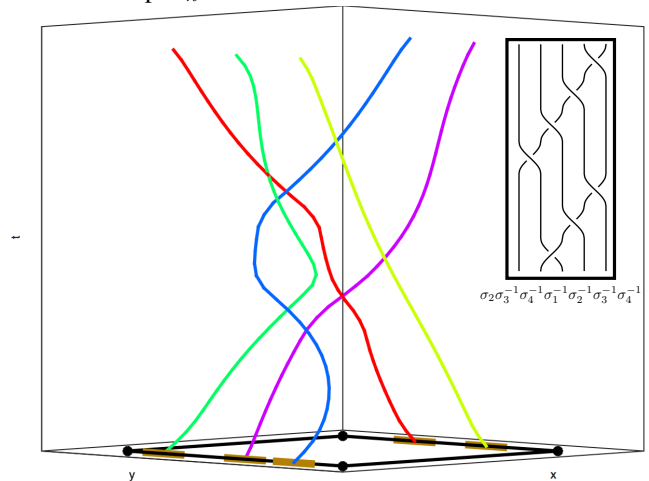


Fig. 5: A trajectory collection (left) and its corresponding braid representation (upon a projection onto the x - t plane).

and *legibility* of motion. These properties become important in environments where robots are operating close to humans and can serve as cost functions for planning intent expressive robot motion. Intuitively, an observer expects to see a *predictable* motion from an agent whose goal is known, whereas the generation of *legible* motion allows an observer to quickly, confidently and accurately predict the agent's goal. In pedestrian environments, such concepts appear to be especially meaningful if we consider the foundations of the pedestrian bargain [18].

A. Predictability of Pedestrian Motion

Predictable pedestrian motion can be assumed to be the result of optimization in terms of *pedestrian competence*. In a scene with n agents and a set of \mathcal{D} destinations, given a specific pedestrian scenario $S^* = \langle \Sigma^*, \Delta^*, \beta^* \rangle$, the predictability score of a trajectory collection ζ can be defined as

$$Predictability(\zeta, S^*) = \exp(-C(\zeta, S^*)), \quad (3)$$

where $C : \mathcal{Z} \times \mathcal{S} \rightarrow \mathbb{R}$ represents *pedestrian competence* of a trajectory collection for a given scenario S^* . Minimizing the pedestrian competence cost leads to a trajectory collection that an observer would expect, given S^* . The exponential mapping in the predictability score accounts for the enforcement of the Principle of *Maximum Entropy*, that results to the following effect: more *competent* trajectory collections are considered as exponentially more predictable, while more *competent* trajectory collections are not severely penalized. This is important, since occasionally, pedestrian motion can be suboptimal.

B. Legibility of Pedestrian Motion

Legible pedestrian motion allows an observer to correctly, confidently and quickly infer the pedestrian scenario S^* that is emerging as a result of the superposition of all agents' individual intentions. The legibility score for a collection of trajectories ζ of total duration T , with respect to a pedestrian scenario S^* can be defined as:

$$\text{Legibility}(\zeta, S^*) = \frac{\int_0^T P(S^*|\zeta)f(t)dt}{\int_0^T f(t)dt}, \quad (4)$$

where $P(S^*|\zeta)$ is the probability that the observer correctly predicts the emerging scenario S^* at time t . The function $f(t)$ serves as a weighting factor that favors trajectories that allow early inference of the actual emerging scenario S^* .

C. Generating Socially Competent Pedestrian Motion

We incorporate the aforementioned formulations in a framework for generating socially competent robot motion in pedestrian environments, based on trajectory optimization. The framework comprises of two main steps: (1) prediction and (2) generation.

In the prediction step, a set of likely, topologically distinct Pedestrian Scenarios \mathcal{S} is derived and represented in a *predictable* way, using the Predictability criterion. In particular, for each pedestrian scenario $S_j \in \mathcal{S}$, we derive a corresponding predictable representation by solving the following trajectory optimization problem:

$$\zeta_j^p = \arg \max_{\zeta \in \mathcal{Z}} \text{Predictability}(\zeta, S_j), \quad (5)$$

and conclude to a tuple $Z^p = \langle \zeta_1^p, \dots, \zeta_m^p \rangle$, containing m collections of predictable trajectories.

We generate robot motion that is aware of all likely scene evolutions contained in \mathcal{S} , in an effort to relax the complication of prediction uncertainties and ensure continuity of robot motion between successive replanning cycles. This is achieved by optimizing the robot's trajectory with respect to a weighted sum of the legibilities of all considered scenarios, where the weight of each scenario j is equal to its respective probability of occurrence $P(S_j|\zeta)$:

$$\xi_R^l = \arg \max_{\xi_R \in \Xi} \sum_{j=1}^m P(S_j|\zeta) \text{Legibility}(\zeta_j^p, S_j), \quad (6)$$

with ξ_R denoting the robot trajectory (which is a part of a trajectory collection ζ).

In the beginning, uncertainty is typically high and thus robot motion will be the result of a weighted synthesis of legible motion reactions to all scenarios. As the scene evolves in time, the emerging scenario will gradually be taken into higher consideration, whereas the rest of them will gradually be rendered as more unlikely and eventually disregarded. This way, the robot motion will neither hinder other pedestrians' paths nor impose unpredictable and unrealistic interactions among them, thus appearing natural and socially competent, even in case of events considered as unlikely.

The outlined planning architecture implements the principal components of the pedestrian bargain as we identified

them earlier, as it leads to (1) efficient, (2) socially compliant and (3) legible motion generation.

Remark 1. It should be noted that the robot is not able to enforce a scenario by itself, as by definition, a scenario is an emergent pattern, synthesized by conscious individual choices of all agents in the scene. In this respect, legibility is used to bias observers' belief towards the scenario that appears to be the most likely at a given moment.

D. Online Algorithm for Socially Competent Navigation

Our online algorithm for generating socially aware robot motion for navigation in pedestrian environments can be summarized as follows:

- 1) Record the state history of all agents for some predefined amount of time.
- 2) Derive a set of possible pedestrian scenarios.
- 3) Score each pedestrian scenario according to its probability of occurrence.
- 4) Derive a *predictable* representation for each pedestrian scenario.
- 5) Synthesize a *legible* robot trajectory by considering all pedestrian scenarios, each weighted according to its score.
- 6) Execute plan.
- 7) Replan frequently (a rate of 10Hz is necessary to ensure satisfactory real-time operation).

VI. IMPLEMENTATION

In this section we state implementation details and assumptions for producing the simulation results of Section VII.

A. Pedestrian Competence

Pedestrian motion is the result of an interplay of geometrical/energy considerations and socially imposed, context-specific specifications. In practice, pedestrians tend to follow short and smooth trajectories while ensuring reasonable clearance from other pedestrians and obstacles. We encode the aforementioned properties into robot motion through a cost function that we refer to as *Pedestrian Competence*, $C : \mathcal{Z} \times \mathcal{S} \rightarrow \mathbb{R}$, which we define as a weighted sum of the following cost functions, each representing desirable trajectory properties for agents operating in pedestrian environments:

1) *Pairwise Clearance Among Agents:* In order to ensure proper clearance for each pair of agents, we use the cost

$$C_{cl}(\zeta, S) = \frac{1}{2} \sum_{t=2}^{T-1} \sum_{ij} a_{ij}(t)^2, \quad (7)$$

with

$$a_{ij}(t) = \begin{cases} d_{ij}(t) - d_{crit}, & \text{if } d_{ij}(t) < d_{crit} \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

where $d_{ij}(t)$ is the distance between agents i and j at time t and d_{crit} is a desired distance threshold.

2) *Clearance from Workspace Bounds*: Respecting the workspace bounds is enforced by considering the cost function

$$C_w(\zeta, S) = \frac{1}{2} C_{rep} \sum_{t=2}^{T-1} \sum_{i=1}^n \gamma_i(t)^2, \quad (9)$$

where C_{rep} is a gain expressing repulsiveness and

$$\gamma_i(t) = \begin{cases} \delta(\xi_i(t), \partial\mathcal{W}) - h_{crit}, & \text{if } \delta(\xi_i(t), \partial\mathcal{W}) < h_{crit} \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

with $\delta(\xi_i(t), \partial\mathcal{W})$ being the signed distance between agent's state at time t and the closest workspace boundary $\partial\mathcal{W}$ and h_{crit} a selected distance threshold.

3) *Path Shortness*: The specification for short path generation represented by the cost

$$C_{sh}(\zeta, S) = \frac{1}{2} \sum_{t=2}^T \sum_{i=1}^n \|q_i(t) - q_i(t-1)\|^2. \quad (11)$$

4) *Trajectory Smoothness*: The specification for smooth trajectory generation is enforced through the cost

$$C_{sm}(\zeta, S) = \frac{1}{2} \sum_{t=2}^{T-1} \sum_{i=1}^n \|q_i(t+1) - 2q_i(t) + q_i(t-1)\|^2, \quad (12)$$

which emerged by considering central finite differences.

Pedestrian Competence is finally defined to be a weighted sum of the aforementioned costs:

$$C = w_{cl}C_{cl} + w_w C_w + w_{sh}C_{sh} + w_{sm}C_{sm}, \quad (13)$$

where w_{cl} , w_w , w_{sh} , w_{sm} denote weights corresponding to the cost functions defined above.

Remark 2. It should be noted that the considered Pedestrian Competence Cost is used to provide a proof of concept for our framework and not as a definitive model of a pedestrian reward function.

B. Intention Expression and Understanding

The entanglement pattern β of each pedestrian scenario encodes a sequence of crossings, each involving a pair of agents. Being legible with respect to a considered scenario corresponds to a robot trajectory refinement towards communicating its intended passing sides in the crossings that involve it in this particular scenario. In other words, given a braid word representing β , the effect of Legibility is the reinforcement of the observer's belief towards expecting the generators specified by β , in which the robot is involved.

In the absence of a well tested, data-driven model of the distribution $P(\beta|\Delta, \zeta)$, in order to engineer legible behaviors as described above, we use the cost function

$$C_{\mathcal{B}}(\zeta, S) = \frac{1}{2} \sum_{k=1}^N \sum_{t=2}^{T-1} f(t)\lambda_k^2(t), \quad (14)$$

where

$$\lambda_k(t) = \begin{cases} \delta_b(b_k, \xi_i(t)) - \delta_{b,crit}, & \text{if } \delta_b(b_k, \xi_i(t_k)) < \delta_{b,crit} \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

and $\delta_b(b_k, \xi_i(t_k))$ represents the signed distance between the generator specified by the scenario in consideration and the corresponding generator emerging by the current trajectory collection at time t_k , while $\delta_{b,crit}$ is a desired

threshold distance and $f(t)$ is a function that favors legible trajectory modifications that take place in the beginning of the trajectory.

Finally, to ensure that the emerging legible trajectory will be sufficiently smooth, we combine the aforementioned cost with the smoothness functional to conclude to the following legibility score

$$\text{Legibility}(\zeta, S) = -(w_{\mathcal{B}}C_{\mathcal{B}} + w_{sm}C_{sm}). \quad (16)$$

C. Reasoning about the Destinations of Other Agents

Under the assumption that the destinations of all agents are independent given their trajectories so far, we model the probability $P(\Delta|\zeta)$ as

$$P(\Delta|\zeta) = \prod_{a \in \mathcal{A}} P(d_a|\zeta) = \prod_{a \in \mathcal{A}} P(d_a|\xi_a). \quad (17)$$

To model the probability $P(d_a|\xi_a)$, we take an approach similar to the one followed by Dragan and Srinivasa [6]. For every possible destination $D_i \in \mathcal{D}$, given a partial trajectory of agent a , ξ_a , from its starting position s_a to its current position q_a we can approximate $P(d_a = D_i|\xi_a)$ as

$$P(d_a = D_i|\xi_a) = \frac{1}{H} \frac{\exp(-C_{sh}(\xi_a) - V_{D_i}(q_a))}{\exp(-V_{D_i}(s_a))} P(D_i), \quad (18)$$

where $V_{D_i}(q_a) = \arg \min_{\xi'_a \in \Xi} C_{sh}(\xi'_a)$, with ξ'_a representing a trajectory starting from q_a and ending at d_a , H is a normalizer across all possible destinations \mathcal{D} , C_{sh} is a cost function quantifying path shortness and $P(D_i)$ is a prior on destinations.

D. Reasoning about Trajectory Entanglements

The distribution $P(\beta|\Delta, \zeta)$ represents how likely each considered entanglement pattern is, given a destination scenario Δ and an observed trajectory collection ζ . This distribution incorporates elements such as social conventions and context particularities. It is in our current research plans to learn such a model from human demonstrations. In this paper, we approach it as follows. For each destination scenario Δ_j , $j \in \{1, \dots, m\}$, we first derive a predictable representation ζ_j^p and then a corresponding nominal braid word β_o , encoding all crossings taking place among the trajectories of all agents of this predictable representation. This braid word is assigned a probability $P(\beta_o|\zeta_j^p, \Delta_j) = 1$, while for any other braid β_k , with $k \neq o$, we consider $P(\beta_k|\zeta_j^p, \Delta_j) = 0$.

VII. DEMONSTRATION

We demonstrate the potential of our approach for intention expression and understanding in pedestrian environments through a series of simulated examples. Trajectory optimization computations were implemented using CHOMP [20], whereas trajectory collections were mapped to algebraic braids using BraidLab [14].

Fig. 5 demonstrates the output of trajectory optimization with respect to predictability for a pedestrian scenario involving five agents moving to different destinations in a common workspace. Fig. 6 depicts the trajectory optimization process

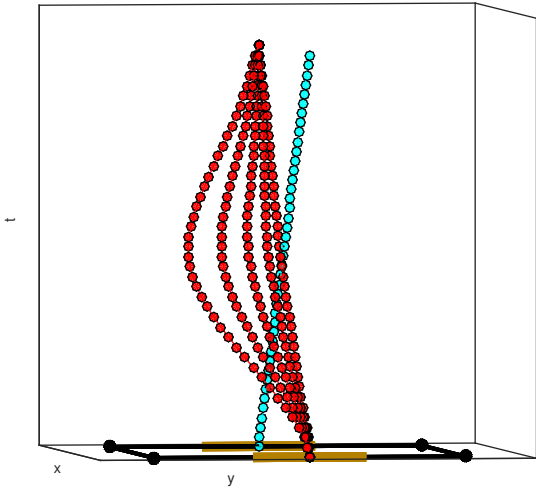


Fig. 6: Two agents moving in a hallway: Demonstration of the trajectory Optimization process for legibility for the agent in red.

for legibility. Two agents are moving towards opposing directions in a rectangular hallway. One of them (cyan) moves straight towards its destination without reacting, while the other one (red) runs the legibility optimizer. Identifying the cyan agent’s intention of passing from its right side, the red agent, initializing from a straight line, reshapes its trajectory to express its cooperation/compliance with the intention of the cyan agent. The figure shows progressive stages of the optimization process (from right to left). The final, leftmost, legible trajectory was derived after 50 iterations.

Fig. 7 depicts the outcome of an online execution of our algorithm by two agents moving to opposing directions in a common rectangular workspace. Fig. 7a shows the swept volumes of the agents as they traverse the workspace (the red agent initializes from the right side whereas the cyan agent initializes from the left side). Each agent identifies the intention of each other (destination and passing preference), derives a predictable representation of the corresponding pedestrian scenario and expresses its own intention by being legible with respect to the predicted scenario. Legibility optimization favors intent expression from early on in the trajectory, thus effectively facilitating a consensus in the collision avoidance, which is in compliance with the pedestrian bargain [18]. Fig. 7b is a plot of the agents’ trajectories in the space-time domain, which corresponds to the braid σ_1 .

Finally, Fig. 8 depicts the result of an online execution of our algorithm for a scenario involving four agents moving to different sides of a square workspace. Fig. 8a shows the swept volumes of all agents. The cyan agent is moving from the top to the bottom, whereas the red agent is moving from the bottom to the top. At the same time, the purple agent is moving from the left side to the right side, whereas the green agent is moving from right to left. Fig. 8b depicts their corresponding trajectories in the space-time domain. The trajectory of each agent is essentially a superposition of legible reactions to their beliefs over possible emerging pedestrian scenarios. The outcome of the execution corresponds to the

braid diagram depicted in Fig. 8c and the algebraic braid $\sigma_1^{-1}\sigma_3^{-1}\sigma_2\sigma_1^{-1}\sigma_1^{-1}\sigma_1^{-1}$.

VIII. DISCUSSION

We presented a novel framework for generating socially competent motion in pedestrian environments. The foundation of our approach is a topological pedestrian scene representation based on braid groups. We used this representation to develop an online navigation algorithm, relying on a mechanism of intention recognition and expression, supported by psychology studies [5]. Proof of concept simulations demonstrated a socially competent behavior that is in line with sociology studies on pedestrian locomotion [18]. Our results show that contrary to conventional wisdom, computationally expensive or sophisticated trajectory predictions are not necessary for navigating among other agents in a safe, socially compliant manner. Ongoing work involves learning predictive models from human demonstrations and conducting a user study as well as real world experiments to validate the efficiency of our algorithm in realistic everyday life situations.

REFERENCES

- [1] E. Artin. Theory of braids. *Annals of Mathematics*, 48 (1):pp. 101–126, 1947. ISSN 0003486X.
- [2] T. Bandyopadhyay, K. S. Won, E. Frazzoli, D. Hsu, W. S. Lee, and D. Rus. Intention-aware motion planning. In *Workshop on Algorithmic Foundations of Robotics (WAFR)*, 2012.
- [3] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun. Learning motion patterns of people for compliant robot motion. *International Journal of Robotics Research*, 24: 31–48, 2005.
- [4] J. S. Birman. *Braids Links And Mapping Class Groups*. Princeton University Press, 1975.
- [5] G. Csibra and G. Gergely. ‘Obsessed with goals’: Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124(1):60–78, Jan. 2007.
- [6] A. D. Dragan and S. Srinivasa. Integrating human observer inferences into robot motion planning. *Auton. Robots*, 37(4):351–368, 2014.
- [7] S. Ferguson, B. Luders, R. C. Grande, and J. P. How. Real-time predictive modeling and robust avoidance of pedestrians with uncertain, changing intentions. *CoRR*, abs/1405.5581, 2014.
- [8] D. Helbing and P. Molnár. Social force model for pedestrian dynamics. *Phys. Rev. E*, 51:4282–4286, May 1995.
- [9] P. Henry, C. Vollmer, B. Ferris, and D. Fox. Learning to navigate through crowded environments. In *In ICRA*, 2010.
- [10] S. Hoogendoorn and P. H. Bovy. Simulation of pedestrian flows by optimal control and differential games. *Optimal Control Applications and Methods*, 24(3):153–172, 2003.

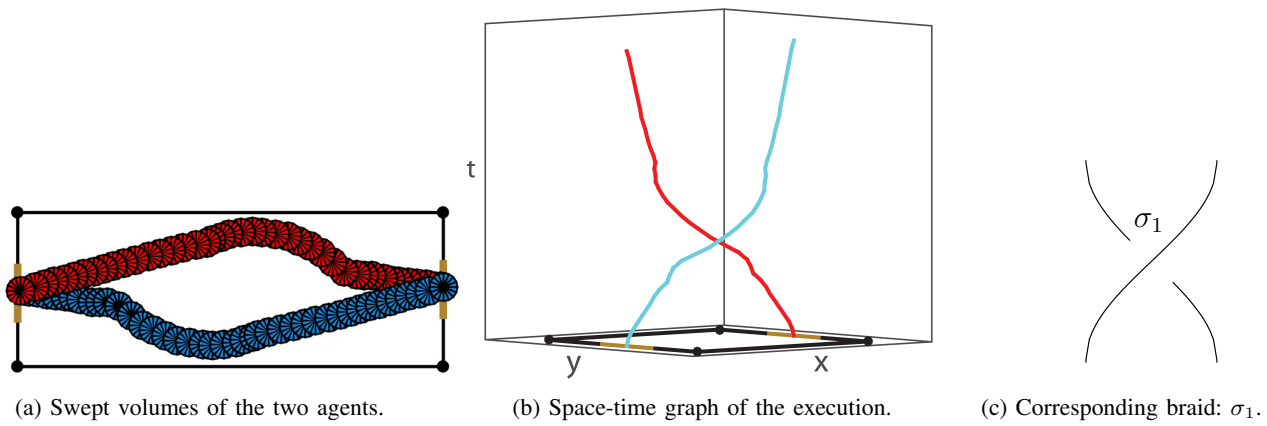


Fig. 7: Online execution: Two agents moving to opposite directions in a narrow corridor, both executing our algorithm.

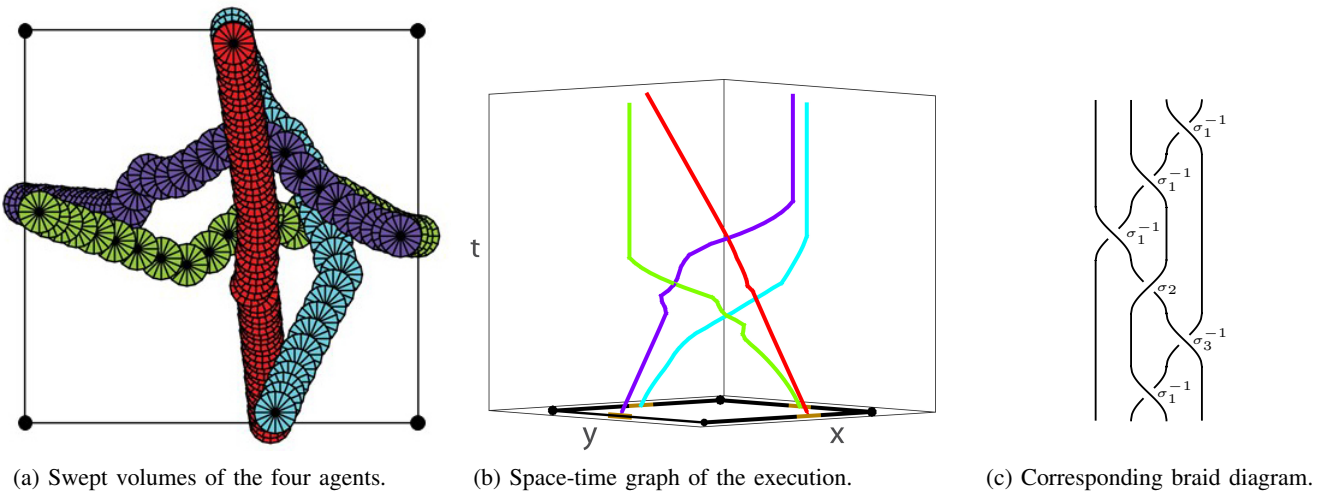


Fig. 8: Online execution: Four agents, executing our algorithm are moving to different sides of a square workspace.

- [11] R. A. Knepper and D. Rus. Pedestrian-inspired sampling-based multi-robot collision avoidance. In *RO-MAN*, pages 94–100. IEEE, 2012.
- [12] M. Kuderer, H. Kretzschmar, C. Sprunk, and W. Burgard. Feature-based prediction of trajectories for socially compliant navigation. In *Proc. of Robotics: Science and Systems (RSS)*, Sydney, Australia, 2012.
- [13] E. A. Sisbot, L. F. Marin-Urias, R. Alami, and T. Siméon. A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 23(5):874–883, 2007.
- [14] J.-L. Thiffeault and M. Budisic. Braidlab: A Software Package for Braids and Loops. *ArXiv e-prints*, Oct. 2014.
- [15] A. Treuille, S. Cooper, and Z. Popović. Continuum crowds. *ACM Transactions on Graphics*, 25(3):1160–1168, July 2006.
- [16] V. V. Unhelkar, C. Pérez-D’Arpino, L. Stirling, and J. A. Shah. Human-robot co-navigation using anticipatory indicators of human walking motion. In *IEEE International Conference on Robotics and Automation, ICRA 2015, Seattle, WA, USA, 26-30 May, 2015*, pages 6183–6190, 2015.
- [17] J. van den Berg, S. J. Guy, M. C. Lin, and D. Manocha. Reciprocal n -body collision avoidance. In *Robotics Research - The 14th International Symposium, ISRR 2009, August 31 - September 3, 2009, Lucerne, Switzerland*, pages 3–19, 2009.
- [18] N. H. Wolfinger. Passing Moments: Some Social Dynamics of Pedestrian Interaction. *Journal of Contemporary Ethnography*, 24(3):323–340, 1995.
- [19] B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *Proc. of the International Conference on Intelligent Robots and Systems*, 2009.
- [20] M. Zucker, N. Ratliff, A. D. Dragan, M. Pivtoraiko, M. Klingensmith, C. M. Dellin, J. A. Bagnell, and S. S. Srinivasa. Chomp: Covariant hamiltonian optimization for motion planning. *The International Journal of Robotics Research*, 32(9-10):1164–1193, 2013.