



Human Robot Motion: A Shared Effort Approach

José Grimaldo da Silva Filho, Thierry Fraichard

► **To cite this version:**

José Grimaldo da Silva Filho, Thierry Fraichard. Human Robot Motion: A Shared Effort Approach. European Conference on Mobile Robotics, Sep 2017, Paris, France. hal-01565873

HAL Id: hal-01565873

<https://hal.inria.fr/hal-01565873>

Submitted on 20 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Human Robot Motion: A shared effort approach

Grimaldo Silva¹ and Thierry Fraichard¹

Abstract— This paper is about Human Robot Motion (HRM), i.e. the study of how a robot should move among humans. This problem has often been solved by considering persons as moving obstacles, predicting their future trajectories and avoiding these trajectories. In contrast with such an approach, recent works have showed benefits of robots that can move and avoid collisions in a manner similar to persons, what we call human-like motion. One such benefit is that human-like motion was shown to reduce the planning effort for all persons in the environment, given that they tend to solve collision avoidance problems in similar ways. The effort required for avoiding a collision, however, is not shared equally between agents as it varies depending on factors such as visibility and crossing order. Thus, this work tackles HRM using the notion of motion effort and how it should be shared between the robot and the person in order to avoid collisions. To that end our approach learns a robot behavior using Reinforcement Learning that enables it to mutually solve the collision avoidance problem during our simulated trials.

I. Introduction

Human Robot Motion (HRM) is the study of how a robot should move among persons. In this context, robot motion must be safe and appropriate. While safety relates to guaranteeing collision-free motion [1], the term appropriate relates to respecting concepts such as social spaces [2], legibility and perceived safety [3].

Many recent studies have focused on tackling HRM by teaching a robot human-like behavior, such as in [4] and [5]. The justification for this approach is that it allows a robot to follow the flow of the persons [4], and also allows for better behavior legibility to persons around the robot. Legibility is important because it was shown that persons tend to solve collision avoidance problems in stereotypical ways under repeated conditions [6], which implies that a robot behaving in an uncommon way forces the person to actively plan its motion instead of relying on already learned motion plans, this means that human-like motion reduces planning effort for all the persons in the environment [7]. Furthermore, another argument is that unexpected motions can be perceived as unsafe by nearby persons even though in practice they may be collision free [5].

In order to create human-aware robots capable of navigating among persons, most current approaches in HRM, such as [8] and [9], operate in two steps. First the probable future behavior of the persons is predicted without considering the robot. Then the future robot motion is computed taking this prediction into account. As a result, the robot always yields,

that is, it avoids to the best of its ability regions where a person is expected to go through. Collision avoidance among persons is, however, mutually solved [10]. This means that, depending on the current disposition of nearby persons, each person is expected to contribute a certain amount of what we call *effort* to avoid a collision. The amount of effort expected from each person and in which manner this effort is represented, as speed or path changes for example, depends on many factors [10], [11], [12], such as: who is first, angle of approach, speed and visibility.

In order to replicate human collision avoidance behavior, our objective is to allow the robot to share collision avoidance effort with people, when necessary, in a safe and appropriate manner, that is, in a way that is expected by its human peers. To that end our approach accounts for two facts: visibility and crossing order. The former represents the understanding of the robot regarding what nearby persons can see, while the latter represents which agent in crossing scenarios should give way to the other. Note, however, that in situations where the person is unwilling or unable to follow a stereotypical motion the robot in our approach will still be able to take full responsibility for avoiding collisions. An important aspect is how the effort needs to be shared between persons and robot. In some situations the person does not expect the robot to yield, such as when the person is behind the robot but intending to overtake. Whereas in other cases the person expects the other agent to give him priority and also to be responsible for most of the collision avoidance [10], as is the case when the front of the robot would collide into the side of a person during perpendicular crossing scenarios.

Predicting human behavior in reaction to a given robot motion in our approach depends on a human-like model (HLM), which unlike many works in HRM such as [8] and [5] does not use the Social Force Model (SFM) which was introduced in [13]. Instead we rely on a slightly modified version of Optimal Reciprocal Collision Avoidance (ORCA) [14], also called RVO 2. This HLM was chosen as it can be directly modified to accommodate different degrees of participation from a particular agent during collision avoidance.

Based on the persons' reaction to a given robot motion, we intend to use this information to avoid collisions with persons in a human-like way. To that end, our approach relies on reinforcement learning (RL) [15] to learn such behaviors, this technique was chosen for its ability to explore the state space and also to learn behaviors that can be recalled even in real-time situations [16].

¹ Grimaldo Silva jose.jgrimaldo@gmail.com and Thierry Fraichard thierry.fraichard@inria.fr are with Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

A. Outline of the Paper

This work is divided into six sections. Section II describes works with related concepts. Afterwards, in Sec. III a formal description of our approach is presented and also how to measure the additional effort required for collision avoidance. This is followed by Sec. IV where this additional effort measure is used to build a human-like collision avoidance strategy. Experimental results of our approach are presented in Sec. V. Finally, a discussion of our results, future works and final remarks are presented in Sec. VI.

II. Background and Contributions

Initial concepts in HRM focused mainly on allowing a robot to respect social spaces, which can be defined in a general sense as regions that for whatever reason a person considers as belonging to them [2].

There are many other concepts that have an influence in HRM, such as comfort. Comfort relates to the subjective feeling of a person that the body is relieved of negative stimuli [17]. Many factors affect comfort, one such factor is the visibility which has been tackled in [17] using a multi-layer costmap that factors the cost of visibility into a costmap in order to calculate the optimal trajectory of an autonomous wheelchair. A definition of comfortable motion that is more related to HRM was made in [5], it can be summarized as the perception of a person being able to walk in their preferred velocity and if their path felt collision free.

Among the several human-like models (HLM) that can approximate human behavior in these cases, we highlight the Extended Social Force Model [13], a method based on modeling each person as being attracted to their goal (in a preferred velocity) and being repulsed by other agents and also static objects in the environment. Another tool used in simulation of pedestrians, particularly in crowd simulation [18], [19], is the reciprocal velocity objects (RVO) [14] which is based on finding velocity choices for agents that guarantee collision avoidance.

Given one such HLM, its possible to calculate the reaction of a given person to a robot motion. This contrasts with many current approaches where the planned human motion is static [8] or probabilistic [9]. That is, in these works the robot avoids regions where persons are predicted to go in order to avoid disrupting their plans.

Another concept, defined in [9], was hindrance. This term relates to situations where a person natural behavior is disrupted due to a robot's proximity. To that end, a human-like planner using Markov Decision Process associates a probability for each of the several possible person trajectories to the goal (a distribution over trajectories), this planner is trained by observing human trajectories. Thus, a robot is able move towards its goal while avoiding high hindrance regions.

Our approach brings novel contributions in relation to those works as we focus on reproducing how persons share collision avoidance effort. To this end, it is necessary to forecast short-term human motion plan in reaction to a given robot action, which we accomplish with a modified ORCA.

III. Overview of the problem

A robot is tasked with reaching a given goal, in-between his current and desired positions any number of persons may cross his path. It is evident that collisions with persons have to be avoided whenever possible. However, persons have certain expectations about how this collision avoidance should take place. To solve this problem it is important to model how the collision avoidance effort should be distributed.

A. Formalization of the problem

Consider that W represents the environment, with $W \subset \mathbb{R}^2$. In this environment, each person p and robot r that belong to the set of dynamic objects D have positional properties: $\mathbf{q}_p = (x_p, y_p, \theta_p) \in \mathbb{R}^2 \times S^1$. Thus we define the state of a given person as $\mathbf{s}_p = (\mathbf{q}_p, \dot{\mathbf{q}}_p)$, where each person also has a goal, which is known a priori, $\mathbf{g}_p = (x_p^*, y_p^*, \theta_p^*) \in \mathbb{R}^2 \times S^1$. Additionally, the robot r is also an agent in this environment and as such also has positional properties \mathbf{s}_r and a goal \mathbf{g}_r .

Although human behavior can be the result of large cognitive effort, recent studies showed that realistic trajectories can be generated with simple models where an agent solely avoids local collisions [5]. Thus, our choice to utilize a reactive HLM to evaluate human reaction to a given robot motion over n time steps is reasonable.

One possible approach to the robot-person collision avoidance problem can be posed in terms of minimizing additional human effort. First, let $\pi_{p,r} = \{\mathbf{q}_p(0), \dots, \mathbf{q}_p(n)\}$ be the predicted trajectory of person p after interaction with a robot r trajectory within a prediction window of n time steps ahead. Moreover, consider that the additional effort of a given trajectory is represented by a value $\Gamma(\pi) \rightarrow \mathbb{R}^*$ (detailed in Sec. III-B). Finally, consider one possible formulation to this problem

$$\pi_{r*} = \arg \min_{\pi_r \in \Pi_r} \sum_{p \in P} \Gamma(\pi_{p,r}) \quad (1)$$

where Π_r is the set of admissible robot motions to the goal. In this model the robot avoids causing additional effort to the person whenever possible, that is, it will minimize the disruption of the person's motion plan while still reaching its goal. This approach is necessary in case the person is unaware of the robot or either unwilling or incapable of changing his motion plan. Conversely, in real scenarios, a person does not always yield. The additional effort required for collision avoidance is shared between the persons involved. In such context, a robot that acts unlike other persons can generate scenarios where, for example, persons are forced to actively think about the robot motion plan instead of relying on already learned stereotypical trajectories. As such, to achieve HRM it is also necessary for the robot to replicate the ability of persons to share necessary changes in planning between themselves in a socially aware manner in order to solve collision avoidance situations in stereotypical situations.

To account for the effort sharing between person and robot, the problem of collision avoidance is posed as an optimization problem in this manner

$$\boldsymbol{\pi}_{r^*} = \arg \min_{\boldsymbol{\pi}_r \in \Pi_r} \sum_{p \in P} |(1 - \alpha_{r,p}) \cdot \Gamma(\boldsymbol{\pi}_{p,r}) - \alpha_{r,p} \cdot \Gamma(\boldsymbol{\pi}_r)| \quad (2)$$

where $\alpha_{r,p} \in [0, 1]$ is the effort distribution coefficient (EDC) between p and r . This coefficient indicates, at each time step, what is the relative cost of the robot's deviation from its baseline goal in relation to the person, a higher proportion engenders less deviation, this is detailed in the section IV.

B. Human trajectory cost function

Anticipating the human effort necessary to execute a given trajectory is a necessary step in order to properly divide effort between person and robot. Many models exist to measure this effort. One such function is the path length and also total time to the goal [20]. Another approach, is given by [21], which describes the cost of a trajectory as a combination of weighted acceleration controls.

Our work relies on the concept of understanding how collision avoidance requires additional effort in relation to the robot baseline motion. Baseline motion represents the trajectory that does not account for the presence of other agents in the environment. The interaction with other agents, however, requires change in the motion plan. To measure this change, the first step is calculating the distance of an agent r to the goal at time t using $d_t(r, g_r) = \sqrt{(x_r(t) - x_r^*)^2 + (y_r(t) - y_r^*)^2}$ where $x_r(t)$ and $y_r(t)$ are, respectively, the x and y coordinates of the agent r at time t . Thus, we can define the change in distance to the goal as $\Delta d_t(r, g_r) = d_t(r, g_r) - d_{t-1}(r, g_r)$. In our approach, at each time step, a baseline change in distance to the goal is estimated, that is, the agent plans its motion without accounting for other agents. This baseline change in distance to the goal at the current time step is represented by $\Delta \mathbf{B}_t(r, g_r)$ and can be understood as the desired progression to the goal.

However, interaction with other agents require additional effort, which impose changes into the baseline motion of an agent. Given this concept, we can define the additional effort of r for a given trajectory as

$$\Gamma(\boldsymbol{\pi}_r) = \sum_{t=1}^n \max\{0, \Delta d_t(r, g_r) - \Delta \mathbf{B}_t(r, g_r)\} \quad (3)$$

This cost function calculates its result based on the difference from the baseline motion to the actual motion. In this formulation, a given motion can only have an equal or smaller cost than the baseline motion at any time step. This definition guarantees that $\Gamma(\boldsymbol{\pi}) \rightarrow \mathbb{R}^*$, which is a property that is important in Sec. IV-B, when using it as part of a reward function during optimization.

IV. Presentation of the Approach

Given the initial state of the person and the robot (including position, goal and velocity), the robot wishes to find a trajectory $\boldsymbol{\pi}_{r^*}$ that shares collision avoidance effort among them in a similar way as another person would. Thus, in this

section we divide our approach to solve the optimization problem of shared effort presented in Eq. 1 and Eq. 2 in five main steps:

- 1) Receive information from sensors (world model/state)
- 2) Find $\forall p \in D$ the $\alpha_{r,p}$ based on current state
- 3) Plan collision avoidance actions up to n steps ahead
- 4) Send planned velocity (action) to wheels
- 5) Stop if goal reached, go to step 1 otherwise

As the robot receives input from its sensors it builds a representation of the world including position of the goal, position and velocity of nearby persons and also his own. This information can be used to generate what is called a model of its environment – a world model.

Information about position and velocity of nearby persons enables the robot to calculate the amount of effort it should share with each one for human-like collision avoidance. The effort distribution coefficient (EDC) and the steps necessary to calculate it are described in details in Sec. IV-A.

Given the world model and the EDC, the motion plans for future timesteps can be calculated. To that end, Reinforcement Learning (RL) is used to learn a motion plan capable of reaching a given goal while avoiding collision with a nearby person. Our formulation of this problem as RL problem is described in Sec. IV-B.

Based on this overview of our approach to solve the shared effort collision avoidance problem, in the upcoming subsections the aforementioned steps are detailed and some advantages and limitations of our approach are discussed.

A. Sharing effort

The proportion of effort shared during collision avoidance between person and a robot varies depending on crossing order and crossing angle. It is known that the person that is giving way has to contribute more to the avoidance than the one passing first [10]. One possible explanation for this comes from difference in visual stimuli that both agents have, as the person that gives way can more easily obtain visual information about the person passing first [10]. In our current formulation these two factors are taken into account to decide shared effort: crossing order and visibility.

The point of potential collision, which is the position where both agents would collide on in case they continue in their current velocity, forms an angle $\zeta_{r,p} \in [0, 2\pi]$ between the current position of the robot r and of person p . Henceforth, when analyzing angles of crossing scenarios, the angle that is being referenced is $\zeta_{r,p}$. Furthermore, the angle $\beta_{r,p}$ is formed from the bearing-angle of r in relation to the position of p . The derivative of the bearing angle $\dot{\beta}$ can be a strong indicator of potential collision and also of crossing order [22]. These angles are shown in Fig. 1.

Based on results found [10] through analysis of the perpendicular crossing scenarios, it was found that the person crossing first has a maximum of 40% contribution in collision avoidance effort, while the one crossing last has a maximum of 40%. Furthermore, it is intuitive that in most situations of head-on collision with similar velocities or when both person and robot see each other but have no clear crossing order,

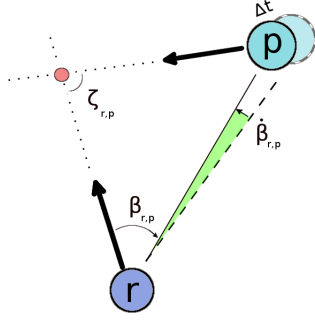


Fig. 1: Collision situation between r and p , where crossing angle ζ , bearing angle β and its derivative $\dot{\beta}$ are shown.

the effort is shared equally between participants. Conversely, in scenarios where one agent is potentially unaware of the other i.e. the passing agent is coming from behind; the responsibility shifts to the agent that sees the other. Recent results also indicate that agents are still able to avoid collisions against obstacles in peripheral vision [23].

This background allows us to correctly distribute effort during collision avoidance between a person and a robot. Thus let $\alpha_{r,p}$ represent the effort sharing coefficient between r in relation to p , which we define as a proportion that weights crossing order and visibility into the relative cost of the robot's deviation from its baseline motion in relation to the person. That is, the higher the proportion, the less deviations from baseline motions of the robot are done in comparison to the person.

The notion that agents do not react to other agents that are outside their field of view, which span around 180° (with both eyes) when looking ahead [24], is translated into our model as a function $\Psi : \mathbb{R} \rightarrow [0, 1]$. This model is used for the robot in order to find trajectories that respect humans expectations. Thus, Ψ is defined as

$$\Psi(\beta_{r,p}) = \begin{cases} 0 & \text{for } |\beta_{r,p}| \geq \frac{\pi}{2} \\ 1 - e^{-\lambda_1(|\beta_{r,p}| - \frac{\pi}{2})} & \text{otherwise} \end{cases} \quad (4)$$

where λ_1 is 15. Based on this model of visibility, the shared effort coefficient of r in relation to p that also accounts for the passing order can be defined as

$$\alpha_{r,p} = (1 - \Psi(\beta_{r,p})) + (0.5 + f(\beta_{r,p})) \cdot \Psi(\beta_{r,p}) \quad (5)$$

$$f(\beta_{r,p}) = \text{sgn}(-\beta_{r,p})(1 - \delta(\beta)) \left(A + \frac{K - A}{1 + \exp(-\lambda_2 \dot{\beta}_{r,p})} \right) \quad (6)$$

where the constants A , K and λ_2 are, respectively, 0.1, -0.1 and 30. Furthermore, $\dot{\beta}$ is the rate of change of β , sgn is the standard sign function that extracts the sign of a real number and $\delta : \mathbb{R} \rightarrow [0, 1]$ is a function that resembles a smooth approximation of the dirac delta distribution that maps β into $1 - |\tanh(\lambda_3 \beta)|$ in which $\lambda_3 = 8$ was chosen to appropriately control the rate of convergence from one to zero. The dirac-like distribution was used to guarantee that the effort is always shared evenly during head-on (or near

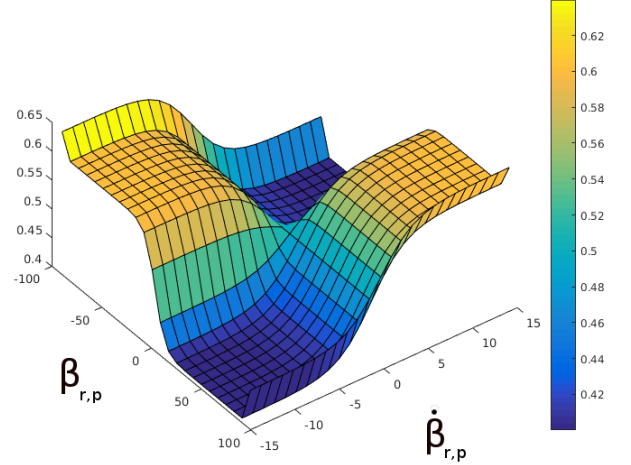


Fig. 2: Shared effort space defines $\alpha_{p,r}$ (both axis in degrees). Its value indicates the relative cost of the robot's deviation from its baseline motion in relation a person's deviation.

head-on) collision scenarios. Additionally, a generalized logistic function represents the boundary between the head-on collision avoidance case and the perpendicular case (where there may be an unequal distribution of effort).

The function f , showcased in Fig. 2, is not applied in cases where there is no chance of collision, as there is no need to change its motion plan, or in cases where the person does not see the robot. In the latter case, for example, if a robot is trying to pass a person from behind it is not appropriate to expect the person to share effort with the robot as the robot is outside its field of view. Thus, in both cases the robot is responsible for the total motion effort.

B. Human-like collision avoidance

To correctly share effort between a person and a robot the optimization problem defined in Sec. 2 is presented in this section in a way can be solved using Reinforcement Learning [15]. The most usual way to represent reinforcement learning problems is as a Markov Decision Process (MDP) which defines a tuple containing $\langle Z, A, R, P \rangle$ that are, respectively, the set of possible states Z , the set of possible actions A , the reward function $R : Z \times A \times Z \rightarrow \mathbb{R}$ and also a transition function $P : Z \times A \rightarrow Z$. At each discrete time step the MDP observes the current state $z_0 \in Z$ and selects an action $a_0 \in A$, as a result, it reaches a new state z_1 and receives a reward r_1 . Given this formulation, the goal of the MDP is to reach a given terminal state s_f with the best expected reward or maximize the expected reward within a certain time frame.

A particular robot behavior, that is, a relation between every state and action is defined as $\psi : Z \rightarrow A$ and called policy. The goal of a reinforcement learning is thus to learn a policy ψ^* that provides better reward than any other policy. Among available methods of Reinforcement Learning, TEXPLORE [16] was selected as our choice as it is robust to noise and able to handle continuous state features.

In order for ψ to make a decision about the future

robot motion, the robot represents its own internal state and the state of nearby persons into a form that can be used in RL. As such, its RL state, defined as z_t , is a tuple $\langle \beta_{r,g}, d_{r,g}, \zeta_{r,p}, \text{ttc}, \beta_{r,p}, \dot{\beta}_{r,p}, d_{r,p} \rangle$ that is used a person where its current motion has risk of collision with the robot, where ttc represents the number of time steps to collision (up to n steps ahead) given linear projection of current velocities, and $\dot{\beta}$ is the rate of change of the bearing angle (see Fig. 1). One limitation of this state space formulation is that it only allows for shared effort in the one person and one robot scenario, given that adding more persons would require an unbounded number of new states to the state space, according to the number of people in the environment.

Using the relative angle and distance to the goal allows the robot to learn what actions better leads him to the goal. For instance, in the absence of collision risk, maintaining the bearing angle of the robot to the goal, $\beta_{r,g}$, at near zero guarantees the reward is maximum. In a similar sense $\dot{\beta}$ is used to allow the agent to measure the risk of collision, the direction of the collision is given by $\beta_{r,p}$ and $\zeta_{r,p}$. When collision is detected within the visible range the ttc is set to the predicted amount of time steps, its value is an arbitrary maximum distance of collision detection otherwise.

The possible actions are a discretization of the control space, represented as forward motion and also left and right motions in 45° angles. The discretization was chosen in such way to reduce learning times. To avoid sharp turns as a result of this discretization, the generated trajectories are smoothed using a B-spline [25]. Given this control space, each action a_t in our model can be represented by a control $u(t)$. Furthermore, the motion $u(t)$ can be seen a trajectory of two points and one time step, where its cost can be expressed in terms of Γ , thus for each action a_t in state z_t its reward is given by

$$r_{t+1} = -|(1 - \alpha_{r,p}) \cdot \Gamma(u_p(t)) - \alpha_{r,p} \cdot \Gamma(u_r(t))| \quad (7)$$

The reward in Eq. 7 is used in case the robot did not reach its goal and there was no collision, in case otherwise, the reward is set to, respectively twenty and minus twenty.

V. Results

In this section we evaluate our approach to shared effort in HRM. The tests were executed inside the ROS framework and its packages but most trajectory planning is done inside ORCA space. The persons are simulated as holonomic agents using ORCA and are able to change their speed, conversely, the robot has a discretized control space that is always at maximum speed. The robot is set to have maximum speed equal to a person's maximum speed, however our model is compatible with any particular proportion between these two speeds. Moreover, in our tests both simulated person and robot have a circular shape with diameter of 34 cm (similar to TurtleBot 2). The robot motion model used to find the trajectory is point mass but restricted to three acceptable actions, see section IV-B for details.

In these tests, the time step between t and $t + 1$ of our prediction is equal 0.25 seconds.

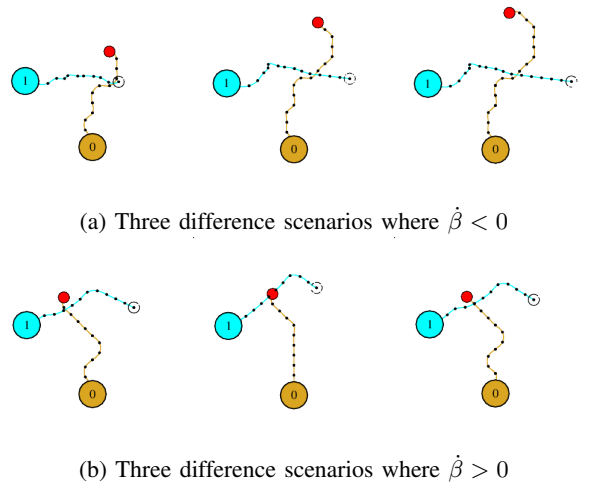


Fig. 3: Crossing angle of 90° , where zero indicates the effort-aware robot and one the human-like planner

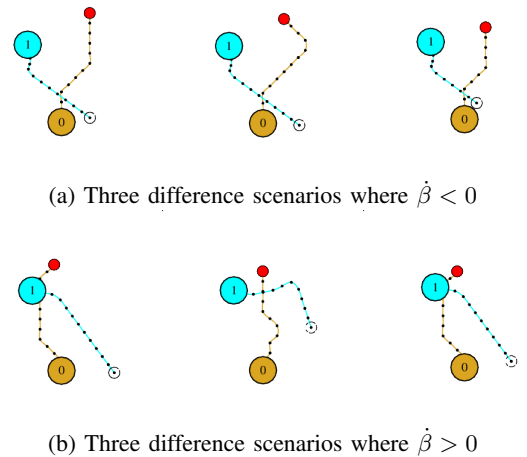


Fig. 4: Crossing angle of 45° , where zero indicates the effort-aware robot and one the human-like planner

A. Trajectories based on crossing order

The trajectories presented were made accounting for different crossing angles and crossing order expectations in order to evaluate their feasibility. The goal of the person is a point with a fixed distance away while the goal of the robot is a random position away and an angle near the direction of their heading, this allows one to randomize the crossing order without altering relative velocities. This is so as the person and the robot are set to have near equal speeds.

It is important to note that there is no perceived order in crossing scenarios with angles of 0° , depicted in Fig. 5. Whereas in the case of crossing angle 45° (Fig. 4) and 90° (Fig. 3), we showcase different available trajectories in the cases where a robot has crossing order priority ($\dot{\beta}_{r,p} < 0$) and in cases where a person has the priority ($\dot{\beta}_{r,p} > 0$).

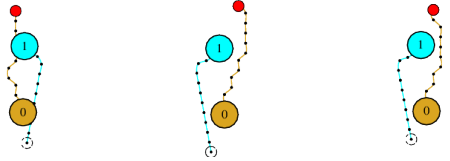


Fig. 5: Case with no crossing order, where zero indicates the effort-aware robot and one the human-like planner

B. Runtime performance

The runtime performance analysis of our approach is presented in Table I to showcase that, after the policy is trained, it can be used to provide response times that are compatible with real-world requirements.

Policy type	Avg. 1-step (s)	Avg. n -steps (s)
Online	0.100489s	3.701310s
Offline	0.000479s	0.016419s

TABLE I: Runtime performance comparison with an online policy, that updates its model while taking actions, and also an offline policy, which only applies learned behavior.

VI. Discussion and Conclusion

This work presented an approach to allow a robot to share the effort required to avoid collision with a person by learning a policy that encodes stereotypical behaviors from persons during collision avoidance. The results observed during experimental evaluation show that the robot is capable of sharing effort with angles 0° , 45° , 90° without simply yielding to the person.

To our knowledge, this is the first work that approximates the human asymmetrical effort sharing during collision avoidance in 90° crossing scenarios in different crossing orders. This can allow a robot to better represent human-like behavior, this is important as following stereotypical motions were shown in recent works to reduce planning effort for persons in the environment.

For the short term, our plan is improve our model as to allow effort sharing with multiple persons, as our current approach is limited to the one person scenario. This would allow observation of cases where avoiding collisions with someone could by consequence cause additional effort to somebody else. Our long term goal is to apply this model into a real robot that has to avoid collision with multiple persons in a real environment.

ACKNOWLEDGMENT

This work is supported by the Brazilian National Counsel of Technological and Scientific Development (CNPq).

References

[1] T. Fraichard and H. Asama, "Inevitable collision states—a step towards safer robots?" *Advanced Robotics*, vol. 18, no. 10, 2004.
[2] F. Lindner and C. Eschenbach, "Towards a formalization of social spaces for socially aware robots," *Spatial Information Theory*, 2011.

[3] C. Lichtenthaler, T. Lorenzy, and A. Kirsch, "Influence of legibility on perceived safety in a virtual human-robot path crossing task," in *IEEE Int. Work. Robot Hum. Interact. Commun.*, Paris (FR), Sep. 2012.
[4] P. Henry, C. Vollmer, B. Ferris, and D. Fox, "Learning to navigate through crowded environments," in *IEEE Int. Conf. Robot. Autom.*, Anchorage (UM), May 2010.
[5] M. Shiomi, F. Zanlungo, K. Hayashi, and T. Kanda, "Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model," *Int. J. Soc. Robot.*, vol. 6, no. 3, 2014.
[6] P. Basili, M. Saglam, T. Kruse, M. Huber, A. Kirsch, and S. Glasauer, "Strategies of locomotor collision avoidance," *Gait and Posture*, vol. 37, no. 3, 2013.
[7] D. Carton, W. Olszowy, and D. Wollherr, "Measuring the effectiveness of readability for mobile robot locomotion," *International Journal of Social Robotics*, 2016.
[8] G. Ferrer and A. Sanfeliu, "Proactive kinodynamic planning using the extended social force model and human motion prediction in urban environments," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, Chicago (UM), Sep. 2014.
[9] B. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. Bagnell, M. Hebert, A. Dey, and S. Srinivasa, "Planning-based prediction for pedestrians," in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, St. Louis (UM), Oct. 2009.
[10] A. Olivier, A. Marin, A. Crétual, A. Berthoz, and J. Pettré, "Collision avoidance between two walkers: Role-dependent strategies," *Gait and Posture*, vol. 38, no. 4, 2013.
[11] M. Huber, Y. Su, M. Kruger, K. Faschian, S. Glasauer, and J. Hermsdorfer, "Adjustments of speed and path when avoiding collisions with another pedestrian," *PLoS ONE*, vol. 9, no. 2, 2014.
[12] S. Jansen, A. Toet, and P. Werkhoven, "Human locomotion through a multiple obstacle environment: Strategy changes as a result of visual field limthrough," *Experimental Brain Research*, vol. 212, no. 3, 2011.
[13] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Physical Review E*, vol. 51, no. 5, 1995.
[14] J. van den Berg, S. Guy, M. Lin, and D. Manocha, *Reciprocal n-Body Collision Avoidance*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
[15] R. Sutton and A. Barto, *Reinforcement Learning: An introduction*, M. Press, Ed., 1998.
[16] T. Hester and P. Stone, "Texplora: real-time sample-efficient reinforcement learning for robots," *Machine Learning*, vol. 90, no. 3, 2013.
[17] Y. Morales, A. Watanabe, F. Ferreri, T. Even, J. Ikeda, K. Shinozawa, T. Miyashita, and N. Hagita, "Including human factors for planning comfortable paths," in *IEEE Int. Conf. Robot. Autom.*, Seattle (UM), May 2015.
[18] R. Narain, A. Golas, S. Curtis, and M. Lin, "Aggregate dynamics for dense crowd simulation," in *ACM Transactions on Graphics (TOG)*, vol. 28, no. 5. ACM, 2009.
[19] A. Bera and D. Manocha, "Realtime multilevel crowd tracking using reciprocal velocity obstacles," in *Int. Conf. on Pattern Recognition (ICPR)*. IEEE, 2014.
[20] S. M. LaValle, *Planning algorithms*. Cambridge University Press, 2006.
[21] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Autonomous Robots*, vol. 28, no. 3, 2009.
[22] J. Cutting, P. Vishton, and P. Braren, "How we avoid collisions with stationary and moving obstacles," *American Psychological Association*, vol. 102, no. 4, 1995.
[23] D. Marigold, V. Weerdesteijn, A. Patla, and J. Duysens, "Keep looking ahead? re-direction of visual fixation does not always occur during an unpredictable obstacle avoidance task," *Experimental Brain Research*, vol. 176, no. 1, 2007.
[24] J. Sardegna, S. Shelly, and S. Steidl, *The encyclopedia of blindness and vision impairment*. Infobase Publishing, 2002.
[25] P. Dierckx, *Curve and surface fitting with splines*. Oxford University Press, 1995.