Why did the human cross the road?

Humans at rest tend to stay at rest. Humans in motion tend to cross the road.

Panayiotis Charalambous
RISE Ltd
Nicosia, Cyprus
p.charalambous@rise.org.cy

Yiorgos Chrysanthou
RISE Ltd& University of Cyprus
Nicosia, Cyprus
y.chrysanthou@rise.org.cy

ABSTRACT

"Humans at rest tend to stay at rest. Humans in motion tend to cross the road – Isaac Newton." Even though this response is meant to be a joke to indicate the answer is quite obvious, this important feature of real world crowds is rarely considered in simulations. Answering this question involves several things such as how agents balance between reaching goals, avoid collisions with heterogeneous entities and how the environment is being modeled. As part of a preliminary study, we introduce a reinforcement learning framework to train pedestrians to cross streets with bidirectional traffic. Our initial results indicate that by using a very simple goal centric representation of agent state and a simple reward function, we can simulate interesting behaviors such as pedestrians crossing the road through crossings or waiting for cars to pass.

CCS CONCEPTS

• **Theory of computation** → *Multi-agent learning*; • **Computing methodologies** → **Physical simulation**; *Collision detection.*

KEYWORDS

animation, crowd simulation, traffic simulation, reinforcement learning, proximal policy optimization

ACM Reference Format:

Panayiotis Charalambous and Yiorgos Chrysanthou. 2019. Why did the human cross the road? Humans at rest tend to stay at rest. Humans in motion tend to cross the road. . In *Motion, Interaction and Games (MIG '19), October 28–30, 2019, Newcastle upon Tyne, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3359566.3364696

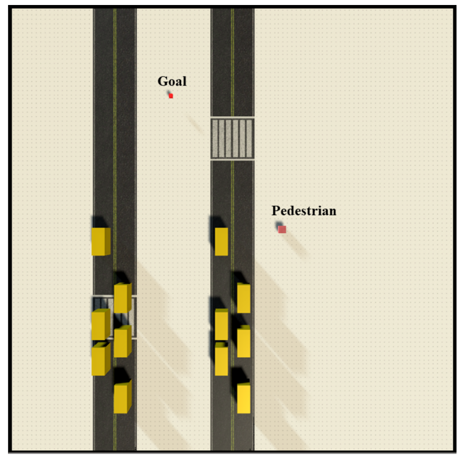

Figure 1: Training environment.

1 INTRODUCTION

The world around us is animated. We experience and interact with human crowds daily in many places such as streets, workplaces, shopping malls, football stadiums or concerts. Humans in crowds participate in various types of interactions with various entities such as other humans, cars and public transportation. The dynamics of crowd motion and the richness of these interactions can significantly impact the ambiance and believability of a scene, and are thus a crucial element of computer generated environments used in computer games, movies, urban studies, safety, traffic control and management and autonomous driving. Despite some really high quality results, most of these systems do not take into account important heterogeneous interactions between humans and traffic such as pedestrians crossing the streets. Moreover, research has shown that autonomous vehicles could potentially decrease road accidents that are caused by human error, by up to 80% by 2040[1]. The development of safe autonomous vehicles and road networks requires an in-depth analysis of the interaction of vehicles, pedestrians and the environment. In this work, we show initial results of a deep reinforcement learning based framework to train agents to cross streets with traffic. Our framework is able to simulate pedestrian street crossing behavior under various conditions without any explicit knowledge of the rules that govern this behavior.

2 RELATED WORK

Crowds There are many crowd simulation techniques that can broadly be categorized as macroscopic or microscopic. In the macroscopic approaches, crowds are modeled as a whole with no distinction of the individuals; these methods fail to simulate variety in motion and behaviours. Microscopic approaches on the other hand consider each individual separately allowing for more variety and aim to get emergent global behaviour. Interested readers can refer to [14] for a more comprehensive discussion on crowd simulation techniques. Of particular interest to this work are the microscopic data-driven models; the promise here is that agents will "learn" how

MIG '19, October 28–30, 2019, Newcastle upon Tyne, United Kingdom

ACM ISBN 978-1-4503-6994-7/19/10…$15.00
https://doi.org/10.1145/3359566.3364696

[1]KPMG, Marketplace of change: Automobile insurance in the era of autonomous vehicle

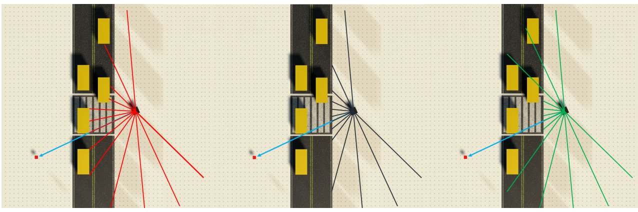

Figure 2: Overview. The agent state is defined relative to its current goal and consists of three consecutive agent observations. Agent observations are defined suing three distinct sets of rays that are cast towards a set of predetermined angles. These rays find nearest distances related to a) vehicles, b) roads and c) zebra crossings. State also includes the current velocity relative to the goal vector.

to behave from real-world examples, keeping the natural crowd ambiance with a wide range of complex individual behaviors [1, 5, 7–9, 21]. Some techniques use observations of real people to learn parameter values for simulators [12, 13, 15, 20]. Recently, several authors proposed Reinforcement Learning as approaches to learn crowd simulation policies by simulation [2, 4, 6, 10].

Pedestrian-car interactions Rasouli et al. [17] introduced large datasets of interactions between pedestrians and human-driven vehicles, to analyze the pedestrian crossing behavior. Most of the studies agree that pedestrian's crossing decision depends mostly on vehicle dynamics (vehicle's position, velocity, acceleration); these can be summarized using the time to collision (TTC) parameter [11]. However, [18] show that vehicle speed, is the most determinant factor of pedestrian's decision process. Apart from car direct indications such as its movement, non-verbal communication between the driver (or vehicle) and the pedestrian was also examined [16, 17]. For instance, pedestrians intending to cross a non-signalized road, seek to have an eye contact communication with the oncoming vehicle's driver in order to agree if the driver will yield. Other studies showed that crossing behavior also depends on agent properties as their age, gender or group size[3].

3 OVERVIEW

We propose a Reinforcement Learning (RL) environment to train agents to cross streets. We employed the Proximal Policy Optimization Algorithm [19] as the learning approach.

States/Actions. The state of an agent is defined in a goal centric 2D-local coordinate system; i.e., it is located at the current position of the agent and has a y-axis that is aligned towards the current goal of the agent Figure 2. We found that defining the state in a goal centric system is a) more stable than defining it using the agent's velocity, b) it converges to a good policy faster and c) this state representation generalizes better than a global representation of state. The agent perceives the environment in 220 degrees using three distinct batches of rays that record closest distances towards a) cars, b) streets and c) crosswalks at 13 predetermined angles. Additionally, we record two values that indicate if the agent is currently on a crosswalk or on a street. Three such consecutive observations define the agent centric state; this representation indirectly encodes the relative movement of the agent as compared to cars, streets and crosswalks. This gives a state $\mathbf{s} \in \mathbb{R}^{123}$. An *action* $\mathbf{a} \in \mathbb{R}^2$ in our framework is velocity that is relative to the agent's local coordinate system.

Learning. In an RL setting, an agent interacts with an environment over a sequence of episodes trying to maximize expected cumulative rewards. We employ a simple as possible environment that will allow us to test different ideas and allow to incrementally extend the learning system to more complex behaviors and environments. To improve training time, we train multiple agents concurrently. We initialize a $25m * 25m$ environment with two bidirectional roads (Figure 1). We initialize cars with random speeds $v \in [1, 10]m/s$ that move both ways; these cars decelerate a) when they approach slower moving cars and b) when they reach a crosswalk. When cars leave one side of the environment, they are translated to the opposite side with randomized speed to help in generalization. We concurrently train 24 agents in 24 similar environments using Proximal Policy Optimization (PPO) [19]. At each training episode, an agent and its goal are placed randomly in the environment. An episode finishes when an agent a)reaches its goal, b) hits a car, c) leaves the bounds of the environment or d) does 1000 simulation steps (20 seconds of simulation time) failing to reach its goal. Agents in our environments make 10 decisions per second.

Reward Function The reward function $R(s, a, s')$ of an agent transitioning between states s and s' by taking an action a defines the task. In the crossing scenario, agents need to a) successfully reach their goals, b) avoid collisions with cars, c) prefer to move through crossings if advantageous and d) prefer to move towards their goals. We defined the following reward function:

$$R(s, a, s') = \begin{cases} -1.0, & \text{car collision} \\ 0.5, & \text{reached goal} \\ R_l + R_r + R_c + R_{gm} + R_g, & \text{otherwise} \end{cases}$$

$R_l = -0.0001$ is a living penalty that *motivates* agents to move instead of standing still, $R_c = 0.001$ is a reward if the agent is on a crossing, $R_r = 0.002$ is a penalty if the agent is on a road and $R_g = 0.0001 * g_{pr}$ rewards or punishes how much the agent progressed towards the goal (g_{pr} is the difference in distance towards the goal between consecutive decisions of the agent.).

4 DISCUSSION

Initial results are very promising; our learned policies show agents exhibiting interesting behaviors such as the ones we described in the previous sections. We refer the interested reader to the short video that accompanies this poster. This is preliminary work and many things need to be considered such as more complex environments and interactions.

REFERENCES

[1] Panayiotis Charalambous and Yiorgos Chrysanthou. 2014. The PAG Crowd: A Graph Based Approach for Efficient Data-Driven Crowd Simulation. *Computer Graphics Forum* 33, 8 (2014), 95–108. https://doi.org/10.1111/cgf.12403

[2] Julio E Godoy, Ioannis Karamouzas, Stephen J Guy, and Maria Gini. 2015. Adaptive learning for multi-agent navigation. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1577–1585.

[3] Andrea Gorrini, Giuseppe Vizzari, and Stefania Bandini. 2016. Towards modelling pedestrian-vehicle interactions: Empirical study on urban unsignalized intersection. *arXiv preprint arXiv:1610.07892* (2016).

[4] Peter Henry, Christian Vollmer, Brian Ferris, and Dieter Fox. 2010. Learning to navigate through crowded environments. In *Robotics and Automation (ICRA),*

2010 IEEE International Conference on. IEEE, 981–986.
[5] Eunjung Ju, Myung Geol Choi, Minji Park, Jehee Lee, Kang Hoon Lee, and Shigeo Takahashi. 2010. Morphable Crowds. *ACM Trans. Graph.* 29, 6, Article 140 (Dec. 2010), 10 pages. https://doi.org/10.1145/1882261.1866162
[6] Jaedong Lee, Jungdam Won, and Jehee Lee. 2018. Crowd simulation by deep reinforcement learning. In *Proceedings of the 11th Annual International Conference on Motion, Interaction, and Games*. ACM, 2.
[7] Kang Hoon Lee, Myung Geol Choi, Qyoun Hong, and Jehee Lee. 2007. Group Behavior from Video: A Data-driven Approach to Crowd Simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '07)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 109–118. http://dl.acm.org/citation.cfm?id=1272690.1272706
[8] Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. 2007. Crowds by Example. *Computer Graphics Forum* 26, 3 (2007), 655–664. https://doi.org/10.1111/j.1467-8659.2007.01089.x
[9] Alon Lerner, Yiorgos Chrysanthou, Ariel Shamir, and Daniel Cohen-Or. 2010. Context-Dependent Crowd Evaluation. *Computer Graphics Forum* 29, 7 (2010), 2197–2206. https://doi.org/10.1111/j.1467-8659.2010.01808.x
[10] Pinxin Long, Tingxiang Fan, Xinyi Liao, Wenxi Liu, Hao Zhang, and Jia Pan. 2017. Towards Optimally Decentralized Multi-Robot Collision Avoidance via Deep Reinforcement Learning. *arXiv preprint arXiv:1709.10082* (2017).
[11] Gustav Markkula, Richard Romano, Ruth Madigan, Charles W Fox, Oscar T Giles, and Natasha Merat. 2018. Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation. *Transportation Research Record* 2672, 37 (2018), 153–163.
[12] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. 2010. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one* 5, 4 (2010), e10047.
[13] Sebastien Paris, Julien Pettre, and Stephane Donikian. 2007. Pedestrian Reactive Navigation for Crowd Simulation: a Predictive Approach. *Computer Graphics Forum* 26, 3 (2007), 665–674.
[14] Nuria Pelechano, Jan M Allbeck, Mubbasir Kapadia, and Norman I Badler. 2016. *Simulating heterogeneous crowds with interactive behaviors*. CRC Press.
[15] Julien Pettré, Jan Ondrej, Anne-Hélène Olivier, Armel Crétual, and Stéphane Donikian. 2009. Experiment-based Modeling, Simulation and Validation of Interactions between Virtual Walkers. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. 189–198.
[16] Anantha Pillai et al. 2017. Virtual reality based study to analyse pedestrian attitude towards autonomous vehicles. (2017).
[17] Amir Rasouli, Iuliia Kotseruba, and John K Tsotsos. 2017. Agreeing to cross: How drivers and pedestrians communicate. In *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 264–269.
[18] Friederike Schneemann and Irene Gohl. 2016. Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments. In *2016 IEEE intelligent vehicles symposium (IV)*. IEEE, 38–43.
[19] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
[20] B. van Basten, S. Jansen, and I. Karamouzas. 2009. Exploiting motion capture to enhance avoidance behaviour in games. *Motion in Games* (2009), 29–40.
[21] M Zhao, W Cai, and SJ Turner. 2017. CLUST: Simulating Realistic Crowd Behaviour by Mining Pattern from Crowd Videos. In *Computer Graphics Forum*. Wiley Online Library.