Why did the human 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?. In *Motion, Interaction and Games (MIG '19), October 28–30, 2019, Newcastle upon Tyne, United Kingdom.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3359566.3364696

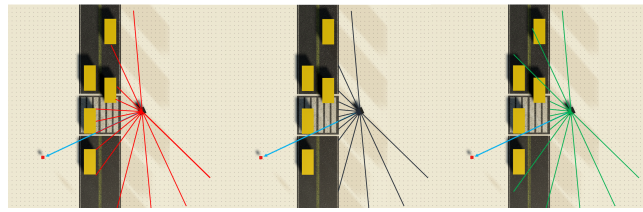

Figure 1: Agent state is defined relative to its current goal and consists of consecutive agent observations; these are found using three distinct sets of rays (vehicles, roads and crossings).

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

Crowd simulation techniques 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 [Pelechano et al. 2016] 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 to behave from real-world examples [Charalambous and Chrysanthou 2014; Lee et al. 2007; Lerner et al. 2007]. Some techniques use data to learn parameter values for simulators [Moussaïd et al. 2010; Paris et al. 2007; Pettré et al. 2009]. Recently, several authors proposed Reinforcement Learning approaches to learn crowd simulation policies by simulation [Lee et al. 2018; Long et al. 2017].

Rasouli et al. [2017] introduced datasets of interactions between pedestrians and human-driven vehicles. Most studies agree that pedestrian's crossing decision depends mostly on vehicle dynamics

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

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

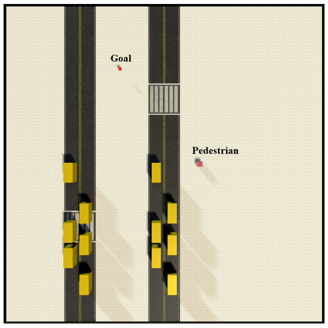

Figure 2: Training environment.

which can be summarized using the time to collision (TTC) parameter [Markkula et al. 2018]. [Schneemann and Gohl 2016] show that vehicle speed, is the most determinant factor of pedestrian's decision process. In addition to car dynamics, non-verbal communication between drivers and pedestrians was examined [Rasouli et al. 2017]. For instance, pedestrians intending to cross a road, seek to have eye contact communication with the oncoming vehicle's driver in order to agree if the driver will yield. Others showed that crossing behavior also depends on features such as age, gender or group size[Gorrini et al. 2016].

3 OVERVIEW

We propose a Reinforcement Learning (RL) framework to train agents to cross streets.

The *state* of an agent is defined in a goal centric 2D-local coordinate system; i.e., located at the current position of the agent with a y-axis that is aligned towards the current goal of the agent Figure 1. 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 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 13 rays that record closest distances towards a) cars, b) streets and c) crosswalks. Additionally, we record if the agent is currently on a crosswalk or on a street. Three such consecutive observations define the agent state $\mathbf{s} \in \mathbb{R}^{123}$; this representation indirectly encodes the relative movement of the agent as compared to cars, streets and crosswalks. An *action* $\mathbf{a} \in \mathbb{R}^2$ in our framework is velocity that is relative to the agent's local coordinate system.

To *learn* in the 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. We initialize a $25m * 25m$ environment with two bidirectional roads (Figure 2). We initialize cars with random speeds $v \in [1, 10]m/s$; 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 similar environments using Proximal Policy Optimization (PPO) [Schulman et al. 2017]. At each episode, agents and goals are randomly placed 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 make 10 decisions per second.

The *reward function* $R(s, a, s')$ of 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 and d) prefer to move towards their goals. We give a reward of -1 if the agent collided with a car and 0.5 if the agent reached its goals. In any other case we define $R(s, a, s') = R_l + R_r + R_c + R_{gm} + R_g$. $R_l = -0.0001$ 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; we demonstrate agents with interesting behaviors such as the ones we described in the previous sections. We refer the interested reader to the accompanying video. This is preliminary work and many things need to be considered such as more complex environments and interactions.

REFERENCES

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.

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).

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.

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, 109–118.

Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. 2007. Crowds by Example. *Computer Graphics Forum* 26, 3 (2007), 655–664.

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).

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.

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.

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.

Nuria Pelechano, Jan M Allbeck, Mubbasir Kapadia, and Norman I Badler. 2016. *Simulating heterogeneous crowds with interactive behaviors*. CRC Press.

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.

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.

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.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).