Simulation Modelling Practice and Theory 74 (2017) 117–133

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Emergent behaviors and scalability for multi-agent reinforcement learning-based pedestrian models

Francisco Martinez-Gil[a,*], Miguel Lozano[a], Fernando Fernández[b]

[a] Departament d'Informàtica, ETSE-UV, Universitat de València, Valencia, Spain
[b] Computer Science Department, Universidad Carlos III, Madrid, Spain

ARTICLE INFO

Article history:
Received 11 May 2016
Revised 24 December 2016
Accepted 7 March 2017

Keywords:
Pedestrian simulation and modeling
Multi-Agent Reinforcement Learning (MARL)
Behavioural simulation
Emergent behaviours

ABSTRACT

This paper analyzes the emergent behaviors of pedestrian groups that learn through the multiagent reinforcement learning model developed in our group. Five scenarios studied in the pedestrian model literature, and with different levels of complexity, were simulated in order to analyze the robustness and the scalability of the model. Firstly, a reduced group of agents must learn by interaction with the environment in each scenario. In this phase, each agent learns its own kinematic controller, that will drive it at a simulation time. Secondly, the number of simulated agents is increased, in each scenario where agents have previously learnt, to test the appearance of emergent macroscopic behaviors without additional learning. This strategy allows us to evaluate the robustness and the consistency and quality of the learned behaviors. For this purpose several tools from pedestrian dynamics, such as fundamental diagrams and density maps, are used. The results reveal that the developed model is capable of simulating human-like micro and macro pedestrian behaviors for the simulation scenarios studied, including those where the number of pedestrians has been scaled by one order of magnitude with respect to the situation learned.

1. Introduction

Interactive simulations with artificial groups or crowds offer a difficult control problem because the simulated people must exhibit very complex behaviors to be realistic. This complexity mainly depends on the simulated scenario and the size of the group being simulated. For instance, in many crowded scenarios, such as buildings, cities, etc., artificial pedestrians must reflect intelligent path planning in stochastic environments, as humans are constantly adjusting their speed to reflect congestion and other dynamic factors. Moreover, when the size of the simulated group increases in many structured scenarios, the problem of providing realistic path planning also increases, as a result some emergent behaviours are expected from a macroscopic perspective (lane formation, clogging effects, etc). A paradigmatic situation is the shortest *vs* quickest scenario (see Section 4.1). It offers an illustrative example of a simple pedestrian facility (2 rooms connected through two doors) which many kind of agents can easily solve. However, when the number of people involved increases a congestion appears very soon, and then a more complex problem has to be faced. In these scenarios, intelligent agents must show macroscopic self organized patterns, normally without considering strategic considerations or coordination techniques, that is, emergent collective behaviors. This type of behaviours emerge from the combination of local interactions between

* Corresponding author. .
E-mail addresses: Francisco.Martinez-Gil@uv.es (F. Martinez-Gil), Miguel.Lozano@uv.es (M. Lozano), ffernand@inf.uc3m.es (F. Fernández).

http://dx.doi.org/10.1016/j.simpat.2017.03.003
1569-190X/

individuals or agent models and they have been studied for many decades ([1,2]). Emergent behaviours have been also incorporated to crowd based simulations, normally under social force models assumptions although different kinds of crowd models currently offer this feature (see Section 2). Nevertheless, one of the challenges in crowd simulations nowadays is to automatically generate macroscopic level behaviors and emergent phenomena from these local rules [3].

Crowd simulation typically requires complex mathematical models to drive the agents in their environments. Multi-agent reinforcement learning (RL) models propose an interesting approach for several reasons. RL agents are efficient because during simulations they are continually performing two main tasks per cycle. Thus, they classify the feature vector provided by the sensors (state recognition), and then, they find the best action to carry out according to the current state. The classification involved in the recognition of the current state has a linear computational cost with the number of generalized states used (see Section 3.2 for a formal state definition). The navigational decision making (find the maximum likelihood action in a given state) is also linear with the number of actions defined. Furthermore, stochastic models also offer interesting possibilities for controlling the variability of the simulated behaviours when cloning them to increase the size of the group, which is an important issue in crowd simulation. During the learning phase, the agents involved are considered as prototypes and once the learning process has been completed, they can simply be cloned or combined. Moreover, the Multi-agent paradigm allows to define independent learning processes for each autonomous agent generating variability in the behaviors learned. By contrast, the learned behaviors normally suffer from poor controllability during the learning process and they are difficult to edit.

The key contributions of this paper are:

- A scalability and performance evaluation analysis of the Multi-agent RL model in different scenarios studied in pedestrian modeling literature.
- An evaluation of the capability of the learned behaviors to create emergent collective behaviors while scaling up the number of simulated agents without additional learning (i.e. generalization).

The rest of the paper has been organized as follows. The next section summarizes the related work on pedestrian simulation and it also includes a specific review of machine learning-based models for pedestrian simulation. Some foundations of RL and a description of the Multi-agent RL framework used is explained in Section 3. Section 4 describes the scenarios used in the experiments and relate them to the literature of the field, then, the configuration of the learning processes is described. In Section 5, the results of these experiments are displayed and the limitations of the approach are discussed. Section 6 presents the conclusions and describes the future work.

2. Related work

Navigational behaviors from individual agents to virtual crowds have been studied in different research areas including social sciences, computer graphics, robotics, engineering (traffic), etc. In our case, pedestrian dynamics must be considered as a research area that inspires many of the navigational models presented in this section.

2.1. Pedestrian models

Efficient path-planning algorithms have been developed for Multi-agent navigation in virtual environments [4–6]. There is a considerable work on local dynamics models able to produce emergent crowd behaviors. Reynolds, in his seminal works [7,8] demonstrated that simple local rules can generate emergent flocking and other behaviors. Among these local methods, the social forces model [9] has been actively studied and many extensions have also been proposed [10–12]. In this context, issues such as sociological factors [13], psychological effects [14], situation-guided control [15] and cognitive and behavioral models [16,17], have also been integrated into the social force model. From a pedestrian dynamics perspective, emergent behaviours have been analyzed in the most popular microscopic models: social forces [1,18,19], cellular automata [20] and agent-based models [21]. Recently, in the work [22] the authors studied the stop-and-go waves that emerge from unidirectional pedestrian traffic. This work presents a numerical model of a following behavior inspired by the analogy with car traffic with an evaluation at a microscopic scale with real examples. The aim of that work is similar to ours, as both present an experimental study to evaluate different emergent behaviours.

In this paper, we use a bottom-up methodology to create crowds. First, a group of agents learn the navigational problem individually. Then, these learned behaviors are reproduced in a crowd to analyze their robustness. Therefore, crowds are simulated using a pure microscopic approach, where each individual is autonomous. In the work [23], the authors also propose a method to clone crowd motion data, to animate crowded scenes. Multi-scale approaches, on the contrary, combine micro and macroscopic characteristics in the same simulation. As a result, the system is capable of reproducing adequate values of the characteristics that define the scale of observation (local collisions, overtaking, etc. in the case of the microscopic scale and mean velocities, flows and densities in the case of the macroscopic scale). The works [24–26] describe formally ways of coupling microscopic and macroscopic scales. A popular approach to the multi-scale problem consists of developing a hybrid framework with different layers where each one assumes the control of a specific scale (a microscopic layer to assume the operational level of control and a macroscopic layer that assumes tactical and strategic levels). Works inspired by this approach are [27,28,29].