Behaviour Mining for Collision Avoidance in Multi-robot Systems

(Extended Abstract)

Jeffery Raphael, Eric Schneider, Simon Parsons and Elizabeth I. Sklar
Department of Computer Science
University of Liverpool, UK
{jeffery.raphael,eric.schneider,s.d.parsons,e.i.sklar}@liverpool.ac.uk

ABSTRACT

Humans are very good at abstract spatial reasoning, both in the physical world and in virtual settings. For example, *gamers* controlling mobile agents in virtual environments can quickly identify potential collisions and effect evasive action. In a multi-robot system, avoiding collisions is a necessary and fundamental behaviour. Recent research has recognised challenges in scaling two-robot collision avoidance methods to larger populations. The work presented here takes advantage of humans' spatial reasoning skills, by collecting data from a simple 2D game where multiple robots move around in an enclosed arena and a human player prevents the robots from colliding with each other. Behaviours for avoiding collisions are learned by mining data logged during games. These mined behaviours are then deployed on robots in a non-interactive environment and compared to a fixed-threshold technique. Results show that mined behaviours provide an effective and flexible alternative.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Learning, Robotics

General Terms

Human Factors, Experimentation

Keywords

Multi-robot systems, collision avoidance, machine learning, behaviour modelling

1. INTRODUCTION

Behaviour mining combines *data mining* and *behaviour modelling*. Here, we describe an approach to behaviour mining that involves building agent-based models of human activity by mining data collected from interactive applications, such as decision-making tools or games [5]. The aim is for the acquired model to be used as a controller for an agent. This could be an agent that helps a novice human make decisions in a complex environment by suggesting actions that might be taken by more experienced human users; or it could be an agent that challenges a human in a competitive game environment by emulating moves of more creative or advanced human players. Here, we explore the following: can we apply data mining and agent-based modelling techniques to a human's clickstream data in order to learn an effective policy for collision avoidance?

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

2. APPROACH

Our approach involves first collecting data from human subjects interacting with physical robots through a point-and-click interface. This graphical interface provides the human "player" with a bird's-eye view of an environment in which a team of robots is deployed for performing exploration tasks. The robots' environment consists of an arena with six "rooms" interconnected via a corridor.

At the start of a "game", the human player and robot team is given a mission: the robot team must explore a set of *interest points* in the arena. The team uses a simple auction-based mechanism to distribute the interest points amongst the robots [4, 3]. The robots begin each mission in a designated "home" room within the arena. Once the interest points are distributed, the robots start moving around the arena, exploring their assigned interest points.

During exploration, it is the human player's responsibility to ensure that the robots do not collide with each other. The player can click on a robot to pause its movement, if she is worried that the robot might crash into a teammate. The player can then click on the same robot again, to resume its motion once the danger of colliding with a teammate has passed. The human can also pause/resume all the robots with the press of a single button. Note that the robots have an internal map of their environment, so they know where the walls are and how to avoid them (i.e., the collision avoidance behaviour investigated here only focuses on robots avoiding each other).

During a game, the system logs data continuously. The log file for any run can be used to reconstruct the state of the environment, at each timestep recorded. We mine these logs to learn the human player's policy for pausing and resuming robot motion. The training set used to develop the model described here consists of a collection of pair-wise configurations of robots and corresponding labels, extracted from the log data. Each exemplar in the training set is represented by a seven-dimensional state vector:

$$\langle range, \theta, V_x, V_y, H_x, H_y, \alpha \rangle$$

where $range$ is the Euclidean distance (in cm); θ is the angle (in radians) to a teammate in the environment; V_x and V_y are the x and y velocities of the robot; H_x and H_y are the heading of the robot to its next waypoint; and α is a flag indicating whether the human clicked on the robot in this state or not.

Our overarching aim is to learn a small-footprint policy that can be placed on-board a low-cost robot platform that has limited computing capacity and memory. So it is important that the learned policy can be represented and executed with a conservative amount of memory and computation time. This means that many machine learning techniques would be unsuitable for this task. Neural networks, which are very good at generalising and identifying non-linear relationships, typically require storing a large number of weights and performing many computations. Similar barriers impede the use of *Gaussian Mixture Models (GMM)*, which have been applied to learning from human interaction [1]. GMMs are computationally greedy, relying on an iterative algorithm (*Expectation-Maximization*, *EM*) to maximise the log of a likelihood function [2].

We used WEKA [6] (version 3.6.7), an open source data mining tool, to develop our model. Our learned policy is implemented using WEKA's J48 decision tree (Quinlan's *C4.5* algorithm [7]). The decision tree is transcribed into a sequence of nested `if-then` statements and embedded into our multi-robot system. We then deployed one robot using this policy into a non-interactive version of the multi-robot game described earlier.

3. RESULTS

This section describes our experiments and results. First, we collected data from human subjects to seed the behaviour mining process. We asked each human subject to play 5 games. In each game, three robots were deployed to explore a total of 8 interest points, configured differently for each game (the robots always started in the same positions). Next, we ran J48 on the data logged during the games, resulting (on average) in 81% correctly classified examples during training and 72% correct classifications during stratified 10-fold cross validation. Then, to test the ability of the mined policy to guide a robot to avoid collisions, we deployed this policy on one robot in a non-interactive version of the system.

An analysis of paths taken by the mined-policy robot during these test games demonstrated successful collision avoidance behaviour. Table 1 illustrates two metrics computed with respect to one representative human player. The first metric, number of pauses, counts the number of times that the human player or the robot following the mined policy paused to avoid a collision. The second metric, total pause time, sums the amount of time that the paused robot in each condition remained motionless until resuming movement, either because the human clicked on it a second time or because the mined policy sensed the environment was safe for the robot to move again. On average, the human initiated pausing less frequently than the mined policy. However, the amount of time that the robot remained motionless in the human-controlled game was significantly longer than in the non-interactive simulation. The biggest indicator of this is the average pause time, which is only 2.27 seconds for the mined policy as compared to 12.70 seconds for the human-controlled robot.

	human	*mined policy*
number of pauses	5.00 (1.22)	16.87 (6.84)
total pause time (sec)	63.52 (27.68)	38.36 (25.41)
average pause time	12.70	**2.27**

Table 1: Metrics comparing performance of mined policy with that of human trainer (mean and standard deviation, over 5 games by human and 15 games by mined policy).

4. SUMMARY

We have presented a method for behaviour mining, applied to collision avoidance in multi-robot systems. This methodology has several advantages. First, it is a general technique for constructing behaviour models from data and can be applied in a variety of domains. Second, it does not require a human teacher to explain their actions, engineer training examples or spend a long time demonstrating behaviours for a learner. Third, the resulting policy does not require a lot of memory to store or computational power to execute, which makes it appropriate for deployment on low-cost, small-compute-footprint robot platforms. Follow-on work will extend this method to a wider user study, gathering data from many human subjects and mining policies from a broader population.

Acknowledgments

This work was partially funded by the National Science Foundation (NSF) under grant #IIS-1116843, by a University of Liverpool Research Fellowship and by a Fulbright-King's College London Scholar Award.

5. REFERENCES

[1] B. Argall, S. Chernova, B. Browning, and M. Veloso. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5), 2009.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Secaucus, NJ, USA, 2006.

[3] A. T. Özgelen, E. Schneider, E. I. Sklar, M. Costantino, S. L. Epstein, and S. Parsons. A first step toward testing multiagent coordination mechanisms on multi-robot teams. In *AAMAS Workshop on Autonomous Robots and Multirobot Systems (ARMS)*, 2013.

[4] E. Schneider, O. Balas, A. T. Ozgelen, E. I. Sklar, and S. Parsons. An Empirical Evaluation of Auction-based Task Allocation in Multi-robot Teams (Extended Abstract). In *Proc of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2014.

[5] E. I. Sklar, A. D. Blair, and J. B. Pollack. Training Intelligent Agents Using Human Data Collected on the Internet (chapter 8). In *Agent Engineering*. World Scientific, Singapore, 2001.

[6] WEKA. `http://www.cs.waikato.ac.nz/~ml/weka`, 2012.

[7] I. H. Witten, E. Frank, and M. A. Hall. *Data Mining, Practical Machine Learning Tools and Techniques*. Elsevier, 3rd edition, 2011.