Developing a Statistical Baseline for Robot Pursuit and Evasion using a Real World Control Architecture

Mark Mansfield, J.J. Collins, Shane O'Sullivan, Malachy Eaton, David Haskett and Thomas Collins
Dept.of Computer Science and Information Systems
University of Limerick,
Limerick. Ireland.
Email: mark.mansfield@ul.ie

Abstract

The canonical robotic pursuit and evasion (PE) contest is one where two agents, an evader and a pursuer, aim to maximise or minimise contact and distance of separation between each other. This paper highlights the need for a greater understanding of such contests and details a method of developing a statistical baseline against which the utility of various pursuer and evader agents can be measured. Also described is an extensible real world control architecture that allows pursuit and evasion capabilities to be integrated with other requirements for a robotic agent operating in the real world. Initial results and future research directions are also discussed.

1 Introduction

PE scenarios receive a lot of attention in robotics because they are highly dynamic contests in an ever-changing environment, that require a tightly coupled sensory-motor loop to facilitate the necessary real-time control. As PE experimentation is relatively new to the field of embodied autonomous robotics, many of the results reported so far have been independent of other results, as there is no common benchmarking system in place. This makes it difficult to assess the overall relevance of much work. For this reason there exists a requirement for some standard characterisation of agent behaviour within PE contests. Also, it is noticable that the majority of results are based on how agents behave. It may be significant to classify agents according to their informational content or internal workings. To this end some work has been carried out by Ficici and Pollack [1, 2] using information theoretic tools to characterise internal dynamics. Results reported in this paper are based on their experimental framework.

Another noticeable trait of some robotic PE research is that it is not grounded in robotics. Simulations take place in unbounded two-dimensional planes meaning that the problems inherent in embodied robotics such as sensor and odometry error are sometimes ignored. It is hoped that by including PE capabilities in an overall global robot control architecture, this work will have a greater and more relevant impact than research that is distinct from any robotic platform.

The next section outlines the motivations for studying PE with section 3 describing the real world control architecture that has been developed. Section 4 details the experimental framework and section 5 describes some preliminary results of these experiments before the penultimate section mentions possible future research directions before the concluding remarks.

2 Pursuit and Evasion Motivation

There are several reasons why there is currently much interest and research into PE scenarios not only within the field of robotics but in other scientific areas also. In a series of papers by Cliff and Miller [3, 4, 5] some of these reasons are outlined. They suggest that PE is the perfect domain in which to study co-evolution where two hostile populations compete in a co-evolutionary arms race to complexity because it is the simplest situation in which co-evolution seems to occur. In addition PE contests are a common setting for protean behaviour. Protean behaviour, also termed adaptive unpredictability, is often observed in predator-prey situations in the animal world where fleeing animals are frequently observed to zig-zag unpredictably when predators are within extremely close range. Proteanism serves to confuse competition as it is the opposite to predictability, and predictability is a weakness that can be exploited by opponents.

The modelling of PE also has implications for a range of sciences including Ethology. Studies of animal behaviour such as Holley's study of PE habits of brown hares and foxes [6] would be complemented by a PE simulation environment where various theories could be tested and perhaps outcomes explained. Endowing robots with the characteristics of specific prey and predators would allow many more situations to be observed and manipulated.

From a robotics point of view PE contests are an interesting case because they require a tight sensory-motor loop, as the introduction of another agent makes the environment highly dynamic meaning that a change of environmental state may require an immediate change of direction or action. While this falls under the umbrella of reactive control there is also a requirement for some deliberation as the robot moves. The robot may have some overall pursuit or evasion strategy that is too complex to compute in real-time. This need for both reactive and deliberative control makes PE an interesting case for the robotics community.

The experimental setup described in this paper is based on research undertaken by Ficici and Pollack [2]. In this work the authors realise the lack of a rigorous metric for PE agent behaviour and use methods from information theory to characterise this. They develop an artificial recurrent neural network to control an evolving evader. There are no inputs to the network with the evader effectively being blind. The only output indicates the angle that the robot is to turn to for the next move. The evader can be seen to be evolving as it learns over generations, based on fitness returns. Using tools from information theory, hand built pursuers develop a statistical model of an evader based on a period of observation of ten thousand time steps. The model is built using n^{th} order statistics. The order of a pursuer describes the number of movement observations made when computing the probability of sequences of moves occuring. A 0 order pursuer makes a single observation before recalculating probabilities while a 10^{th} order pursuer updates probabilities based on 11 observations.

Simulation takes place in a 2-D plane where both agents move at a rate of one bodylength, in a particular direction, per time-step. The pursuer perceives the evader to be moving in one of eight directions and develops a statistical model approximating the evader's movement patterns. The pursuer uses its model to predict the movement of the evader for a specified number of time-steps. If the pursuer can get to any point on the predicted path before the evader, then it will move directly to that point in order to intercept the evader. Otherwise the pursuer moves to the first point on the path.

Fitness for the evader is determined by the average distance it can maintain between itself and the pursuer over the course of the game while the pursuer aims to maximise the number of times it comes within one body length of the evader.

Results are based on observation of simulated PE contests and show that higher order pursuers are more successful, while evaders evolved against higher order pursuers develop more unpredictable movement patterns. Results based on the accuracy of pursuer modeling capabilities are reported in a later section of this paper.

3 A Real World Control Architecture

In order to deploy an embodied agent with the abilities to pursue and evade it has been necessary to develop a real world control architecture that takes into account both the needs of real-time control as well as time-consuming deliberation. The extensible architecture diagrammed in figure 1 shows the significant classes for an architecture that allows robots to build maps as well as pursue and evade. Client controllers take readings from the physical robot and pass these to the *serviceControl* module. The *serviceController* is where the inherent extensibility of the system is realised. It is endowed with different queues for distinct types of data. This is necessary because map building and PE contests require the manipulation of different types of data. Map building work by O' Sullivan [7] has required the use of sonar data while PE uses data concerning the position and orientation of robots. The *serviceControl* module queues data, which is popped from its queues by the robots client services that are the classes *Mapping* and *Referee* in this case. The clients can then manipulate the data as they see fit. The *mappingModules* build maps useful for navigation while the *Pursuer* and *Evader* modules decide on long-term policies for robot motion. Effectively the *serviceController* is a store for data with the client controllers and mapping and PE services being clients that register with it.

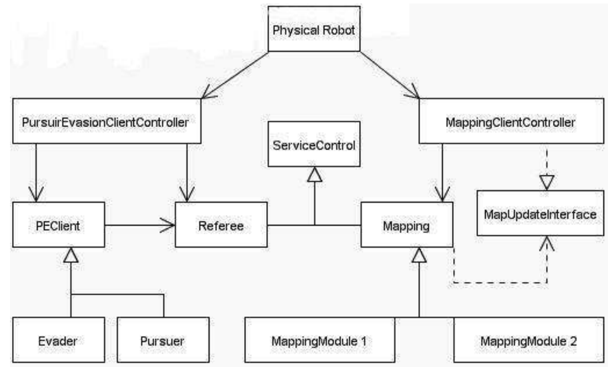

Figure 1: Extensible Control Architecture

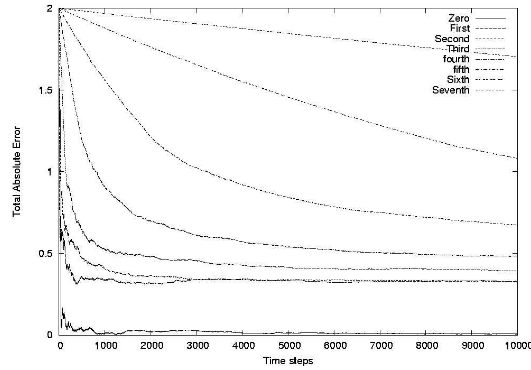

Figure 2: Total Absolute Error

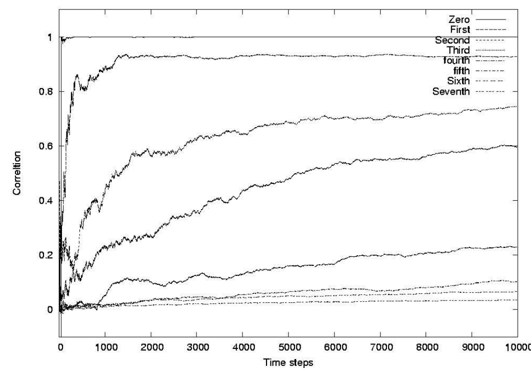

Figure 3: Baron's Correlation Coefficient

4 Experimentation

4.1 Setup

For the work described here, the evolving evader of Ficici and Pollack has been replaced by an evader that bases its moves on a predefined probability distribution, which is created using n^{th} order statistics. A 0-order evader effectively behaves in a random fashion with actions having an impact for only one time step into the future, whereas a 7^{th} order evader commits to movements for eight time steps into the future. The pursuers however, are replicas of those described above, computing their model of the evader using a sliding-window mechanism. From the pursuers point of view it is irrelevant that the neural network generator has been replaced by a probability distribution because the evader is only able to observe the movement of the evader and not its decision process.

4.2 Simulation Suite

As already stated these experiments are being deployed in a real world robotic environment. Testing is taking place using the Saphira/Aria software robot simulation suite. Testing in an environment such as this is considered necessary, as much previous simulation in this domain has taken place in unbounded two-dimensional planes where real world error is not considered. Also a software simulation suite is suited to the code-debug-test cycle and there is no chance of damage to physical robots. One Saphira time step is 100ms in duration. This has lead to some difficulty because a robot takes more that 100ms to turn and move in a real world environment as is necessary for these experiments.

5 Preliminary Results

Preliminary results regarding the accuracy of pursuer prediction capabilities have been recorded. Experiments by Ficici and Pollack have used agents up to an order of 12 that can move in any one of 8 directions. It has become apparent that an observational period of 10,000 time steps is not sufficient to give a clear indication of the probability model of the evader, as a 12^{th} order evader would have 8^{13} potential move sequences. For evaders of higher order it is suggested that an extended observational period is required. Evaders up to an order of 7 that can move in 4 directions have been modelled. Figure 2 shows the total absolute error for evaders of order 0 to 7. With a maximum of 2, it is clear that for a 0 order evader, the error rapidly drops but for evaders of higher order there is a slow gradual decrease in error. For a 7^{th} order evader the error drops approximately 12.5%.

Figure 3 shows the Baron cross correlation coefficient for the eight evaders calculated by:

$$C_N(y) = \frac{< I_T T > - < I_T >< T >}{\sigma(I_T)\sigma(T)} \quad (1)$$

where C_N(y) is Baron's cross correlation coefficient, $I_T T$ is the evaders probability distribution, T is the pursuers estimation, σ is the standard deviation over the distributions and <> is the average operator. This results in a value between -1 and 1, with values close to 1 meaning there is a high correlation between the two distributions. Figure 3 shows that zero order pursuers are able to make a good approximation of evader characteristics quite rapidly, while 10000 time-steps is not long enough for higher order pursuers to develop an accurate model of an evader. Both of these diagrams relate to experiments using agents of the same order.

6 Future Work

Once the information theoretic experiments described above have been completed, there appears to be two promising avenues for future work. The first is the study of co-evolution using PE as mentioned earlier. The second is the application of Reinforcement Learning (RL) to PE. Smart and Kaebling [8] suggest that RL may be well suited to robotics because a RL robot is given a task specification that it must learn to accomplish itself rather than instructions on how to complete the task. This makes sense because autonomous embodied robots are often deployed in unpredictable environments that can be difficult to model and specify in advance. Work by Asada et al [9] has shown that with a suitable descretisation of the state space, simple robotic tasks are achievable, while Ono et al [10] show that effective PE is a real possibility using RL. Future work involves replacing the neural network controlled evolving evader with a RL agent and observing how these compare when inserted into the co-evolutionary framework.

7 Conclusion

Currently there exist few metrics to characterise the behaviour of pursuit and evasion agents. This paper has described previous and current investigation into this area, reasons for simulating PE in robotics and a real world control architecture that has been developed for use in PE contests. Clearly there are useful insights to be gained into co-evolutionary dynamics among other things from the study and simulation of PE contests. The introduction of RL into this setting may provide agents that can learn the best ways to pursue and evade. This is the subject of future work.

References

[1] Ficici, S.G. and Pollack J.B.(1998) Coevolving Communicative Behavior in a Linear Pursuer-Evader Game. From Animals to Animats 5, pp 505-514. MIT Press.

[2] Ficici, S.G. and Pollack, J.B. (1999). Statistical Reasoning Strategies in the Pursuit and Evasion Domain. Proc. of the Fifth European Conference on Artificial Life, pp 70-88. Springer.

[3] Miller G.F. and Cliff D. (1994) Protean Behaviour in Dynamic Games: Arguments for the co-evolution of pursuit-evasion tactics. From Animals to Animats 3, pp 411-420. MIT Press.

[4] Miller G.F. and Cliff D.T. (1994) Co-Evolution of Pursuit and Evasion I: Biological and Game-Theoretic Foundations. Technical Report CSRP311. University if Sussex (1994).

[5] Cliff D. and Miller G.F. (1996) Co-Evolution of Pursuit and Evasion II: Simulation Methods and Results. In From Animals to Animats 4, pp 506-515. MIT Press.

[6] Holley A.J.F. (1993) Do Brown Hares Signal to Foxes? Ethology 94, pp 21-30.

[7] O'Sullivan (2003) An Empirical Evaluation of Map Building Methologies in Mobile Robotics using The Feature Prediction Sonar Noise Filter and Metric Grid Map Benchmarking Suite. (2003) Masters Thesis, University of Limerick.

[8] Smart, W.D. and Kaelbling L.P. (2002) Effective Reinforcement Learning for Mobile Robots. International Conference on Robotics and Automation.

[9] Asada, M., Noda, S. et al (1994) Coordination of Multiple Behaviors Acquired By A Vision-Based Reinforcement Learning. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, pp917-924.

[10] Ono, N., Fukomoto K. et al (1996) Collective Behavior by Modular Reinforcement Learning Animats. In From Animals to Animats 4, pp 618-624. MIT Press.