Real-Time Imitation-Based Adaptation of Gaming Behaviour in Modern Computer Games

Steffen Priesterjahn
University of Paderborn
Dept. of Computer Science
33098 Paderborn, Germany
spriesterjahn@upb.de

Alexander Weimer
Prosis GmBH
Dieselstraße 22
85080 Gaimersheim,Germany
aweimer@upb.de

Markus Eberling
University of Paderborn
Dept. of Computer Science
33098 Paderborn, Germany
markus.eberling@upb.de

ABSTRACT

In the course of the recent complexification and sophistication of commercial computer games, the creation of competitive artificial players that are able to behave intelligently and successfully in the featured highly dynamic and complex virtual worlds has become a considerable challenge. This paper describes an evolutionary real-time adaptation approach to produce competitive artificial players in an action game. The proposed method is inspired by the idea of social learning or cultural evolution. Thus, the agents try to adapt to the level of their opponents by the exchange of information about advantageous behaviours within the population. In addition, the behaviour of the opponents and other players is recorded and used to create more sophisticated and human-like agents.

Categories and subject descriptors:
I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—Intelligent agents, Multiagent systems; I.2.6 [Artificial Intelligence] Learning—Knowledge Acquisition

General Terms: Algorithms

Keywords: Artificial Intelligence, Games, Machine Learning, Multiagent Systems, Online Algorithms

1. INTRODUCTION

The implementation of artificial characters or game agents in computer games raises several challenges. First of all, the agents have to be as competitive as the human players. In addition, as games are often played in real time, the agents also have to be able to adapt and learn online in an ongoing game and to reach a competitive level as fast as possible. Finally, the game characters should not only be competitive, but also believable. They should show human like behaviours and ultimately be indistinguishable from humanly controlled game characters.

In our previous work [5, 6, 4] we have proposed the usage of imitation to generate more believable behaviours by imitating other players. Though, as imitation is inherently done with errors, we proposed the addition of a learning method on top of the raw copying of recorded behaviours to reach a competitive performance. In particular we successfully used an evolutionary algorithm to breed well performing game agents that were initialised from a recording of a role model. However, because of its explorative nature, this approach can not be used in an ongoing game. The problem is that in each generation defective and malfunctioning agents are produced by the variation operators. Yet, when we tried to apply Q-learning [4], a typical online learning method, we found out, that the usage of a population of agents and their combined experiences as the basis for learning greatly enhances the quality of the results and the stability of the learning process in the highly uncertain game environment.

Therefore, this paper presents a learning method that combines the advantages of population-based learning with the ability to learn online by incorporating ideas from reinforcement learning. It is inspired by the concepts of social learning [3] and memetics [1] and follows the idea that in a group of individuals the low performing ones try to improve themselves by imitating the good performing individuals. Therefore, we chose to call it *imitation learning*.

We have already published parts of this method in 2007 [6]. However, this paper presents the method in a much improved and clearer form. In particular we will apply imitation learning to learn combat in the game QUAKE III (©1999, id software) - a very popular three-dimensional action game.

2. BASIC IDEA

Imitation learning can in general be applied to all reinforcement learning problems or problems in which an agent has to learn to become competitive in an uncertain environment. In the proposed form it requires that the encoding of the behaviour of an agent is done as a set of rules - very similar to learning classifier systems. The adaptation process can be divided into four steps: evaluation, elite identification, rule replacement and individual adaptation / mutation.

Algorithm 1 Imitation Learning

```
inputs: $\mu, \sigma, n \in \mathbb{N}, \mu \leq n$
initialise $n$ agents
loop
    evaluate agents
    elite identification: determine the $\mu$ elite agents
    for all non-elite agents do
        choose a random role model from the elite agents
        select $\sigma$ rules from the role model
        replace rules
        mutate $\sigma$ worst rules
    end for
end loop
```

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-130-9/08/07.

Initialisation

As we stated above, the agents are initialised with behaviour rules from a recorded player. These rules simply state what movement the player made in a corresponding game situation. For details on the encoding and modelling of these rules we refer to our previous work [4]. In our experiments we used the built-in QUAKE III agent as the source for the recording because it has a constant performance and can be used to reliably measure, if the generated agents are competitive.

Evaluation

After initialisation the agents and their rule sets are evaluated for a certain timespan, whereas the performance of an agent is determined by the accumulated rewards over the evaluation phase and the rule values are determined by using a policy evaluation-based method.

In our case the agents are evaluated for one minute of combat gameplay against the built-in QUAKE III agent. It is important to notice that - unlike in a learning classifier system - in our implementation there is only one rule that is executed in each time frame. Rewards are determined by the damage that the agent applied to its opponents minus the damage that it received when the respective rule was applied. At the end of the evaluation a policy evaluation-derived method is used to determine the real rule values based on the garnered rewards and the transition probabilities between the rules.

Elite Identification

After the evaluation, the $\mu \in \mathbb{N}$ best performing agents, as indicated by their accumulated rewards, are identified as the *elite agents*. The others become the *imitators*.

Rule Replacement

The rule replacement is the crucial part of the adaptation mechanism. As the single rules are not independent from each other, many behaviours are encoded by a sequence of rule applications. A simple replacement of rules with low values by elite rules with high values will often lead to defective agents. Therefore, we devise a more careful mechanism - as illustrated by figure 1 - that is inspired by results from memetics and viral marketing research about the acceptance and spreading of new ideas [2].

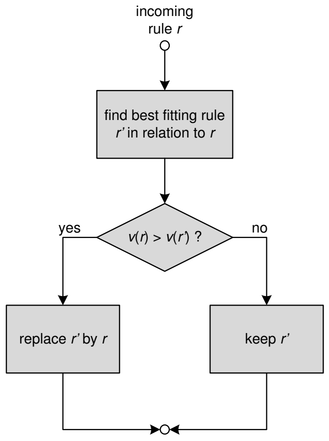

Figure 1: Rule Replacement Mechanism

First of all, each imitator randomly selects just one elite agent as its role model, from which it receives new rules. Therefore, the incoming rule package is known to have successfully worked together. In our application the transmitted rules are simply the σ highest valued rules of the chosen elite agent.

To maintain the coherence of the rule sets a new rule can only replace the rule that is most similar to itself. So, for each incoming rule, the imitator identifies what rule it would apply in the most similar situation. These two rules then compete by their value. The new rule will only replaces the old one, if it proposes a higher utility.

Individual Adaptation / Mutation

For individual adaptation we simply devise a mutation mechanism, though other adaptation techniques are possible. With respect to the uncertainty and high dynamics of our considered environments, the elite agents do not change their rules. This keeps the learning process more stable. They therefore can be seen as taking the role of the parents in an evolutionary algorithm.

The imitators mutate the σ rules with the lowest utility. In our implementation, the agents only mutate a rule, if it has a negative utility, which indicates that the rule has led to a situation in which the agent received damage. Therefore, it is reasonable to do something else in the respective situation.

3. RESULTS

Concerning the results, the limited space gives us only the possibility to briefly state the overall results. A much more detailled discussion of the method and its results can be found in [4]. We have conducted an extensive series of experiments to compare imitation learning with our previous approaches and to find out, how sensitive it reacts to parameter changes.

First of all, imitation learning was able to deliver the same quality of agent behaviour as the previous evolutionary approach. Yet, to accomplish this, it only needed half the population size. Because of its more careful design and more online learning and exploitation-oriented adaptation technique, it was possible to reach this without the generation of defective agents.

In addition, imitation learning has proven to be very robust and stable. It needs a certain minimum amount of agents (in our case 16) to work well, but then reacts very robustly to parameter changes. The reason for this mainly lies in the adaptive manner of the rule replacement mechanism. We had to employ very extreme setups, like using no discounting, or having only one possible elite agent to choose from, to diminish its performance.

4. REFERENCES

[1] S. J. Blackmore. *The Meme Machine*. Oxford University Press, 2000.

[2] K. Chielens and F. Heylighen. Operationalization of Meme Selection Criteria: Methodologies to Empirically Test Memetic Predictions. In *Proceedings of the Joint Symposium on Socially Inspired Computing (AISB'05)*, pages 14–20, 2005.

[3] R. Conte and M. Paolucci. Intelligent Social Learning. *Journal of Artificial Societies and Social Simulation*, 4(1):U61–U82, 2001.

[4] S. Priesterjahn. *Online Adaptation and Imitation in Modern Computer Games*. PhD thesis, University of Paderborn, 2008.

[5] S. Priesterjahn, O. Kramer, A. Weimer, and A. Goebels. Evolution of Reactive Rules in Multi-Player Computer Games Based on Imitation. In *Proceedings of the International Conference on Natural Computation (ICNC'06)*, volume 2, pages 744–755. Springer, 2005.

[6] S. Priesterjahn and A. Weimer. An Evolutionary Online Adaptation Method for Modern Computer Games. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'07)*, 2007.