MODELLING MOTIVATION FOR EXPERIENCE-BASED ATTENTION FOCUS IN REINFORCEMENT LEARNING

A thesis submitted in fulfilment of the requirements for the
degree of Doctor of Philosophy in the School of Information Technologies at
The University of Sydney

Kathryn Elizabeth Merrick

August, 2007

I hereby declare that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other university or institution.

Kathryn E. Merrick

Sydney
August, 2007

Acknowledgements

A PhD thesis is rarely the work of just one person. This thesis has benefited from the input and advice of members of the Key Centre for Design Computing and Cognition at the University of Sydney. My supervisor, Mary Lou Maher, has been both enthusiastic and supportive throughout. Mary introduced me to John Gero who, in turn, showed me the work of Rob Saunders, which has influenced the models developed in this thesis. Rob has provided support and guidance during the writing process. I would also like to thank Professor Peter Eades from the National ICT Australia for his help and advice over the past three years. This research was supported by an Australian Postgraduate Award and a National ICT Australia PhD scholarship.

Abstract

Computational models of motivation are software reasoning processes designed to direct, activate or organise the behaviour of artificial agents. Models of motivation inspired by psychological motivation theories permit the design of agents with a key reasoning characteristic of natural systems: experience-based attention focus. The ability to focus attention is critical for agent behaviour in complex or dynamic environments where only small amounts of available information is relevant at a particular time. Furthermore, experience-based attention focus enables adaptive behaviour that focuses on different tasks at different times in response to an agent's experiences in its environment. This thesis is concerned with the synthesis of motivation and reinforcement learning in artificial agents. This extends reinforcement learning to adaptive, multi-task learning in complex, dynamic environments.

Reinforcement learning algorithms are computational approaches to learning characterised by the use of reward or punishment to direct learning. The focus of much existing reinforcement learning research has been on the design of the learning component. In contrast, the focus of this thesis is on the design of computational models of motivation as approaches to the reinforcement component that generates reward or punishment. The primary aim of this thesis is to develop computational models of motivation that extend reinforcement learning with three key aspects of attention focus: rhythmic behavioural cycles, adaptive behaviour and multi-task learning in complex, dynamic environments. This is achieved by representing such environments using context-free grammars, modelling maintenance tasks as observations of these environments and modelling achievement tasks as events in these environments. Motivation is modelled by processes for task selection, the computation of experience-based reward signals for different tasks and arbitration between reward signals to produce a motivation signal. Two specific models of motivation based on the experience-oriented psychological concepts of interest and competence are designed within this framework. The first models motivation as a function of environmental experiences while the second models motivation as an introspective process.

This thesis synthesises motivation and reinforcement learning as motivated reinforcement learning agents. Three models of motivated reinforcement learning are presented to explore the combination of motivation with three existing reinforcement learning components. The first model combines motivation with flat reinforcement learning for highly adaptive learning of behaviours for performing multiple tasks. The second model facilitates the recall of learned behaviours by combining motivation with multi-option reinforcement learning. In the third model, motivation is combined with an hierarchical reinforcement learning

component to allow both the recall of learned behaviours and the reuse of these behaviours as abstract actions for future learning.

Because motivated reinforcement learning agents have capabilities beyond those of existing reinforcement learning approaches, new techniques are required to measure their performance. The secondary aim of this thesis is to develop metrics for measuring the performance of different computational models of motivation with respect to the adaptive, multi-task learning they motivate. This is achieved by analysing the behaviour of motivated reinforcement learning agents incorporating different motivation functions with different learning components. Two new metrics are introduced that evaluate the behaviour learned by motivated reinforcement learning agents in terms of the variety of tasks learned and the complexity of those tasks.

Persistent, multi-player computer game worlds are used as the primary example of complex, dynamic environments in this thesis. Motivated reinforcement learning agents are applied to control the non-player characters in games. Simulated game environments are used for evaluating and comparing motivated reinforcement learning agents using different motivation and learning components. The performance and scalability of these agents are analysed in a series of empirical studies in dynamic environments and environments of progressively increasing complexity. Game environments simulating two types of complexity increase are studied: environments with increasing numbers of potential learning tasks and environments with learning tasks that require behavioural cycles comprising more actions.

A number of key conclusions can be drawn from the empirical studies, concerning both different computational models of motivation and their combination with different reinforcement learning components. Experimental results confirm that rhythmic behavioural cycles, adaptive behaviour and multi-task learning can be achieved using computational models of motivation as an experience-based reward signal for reinforcement learning. In dynamic environments, motivated reinforcement learning agents incorporating introspective competence motivation adapt more rapidly to change than agents motivated by interest alone. Agents incorporating competence motivation also scale to environments of greater complexity than agents motivated by interest alone. Motivated reinforcement learning agents combining motivation with flat reinforcement learning are the most adaptive in dynamic environments and exhibit scalable behavioural variety and complexity as the number of potential learning tasks is increased. However, when tasks require behavioural cycles comprising more actions, motivated reinforcement learning agents using a multi-option learning component exhibit greater scalability. Motivated multi-option reinforcement learning also provides a more scalable approach to recall than motivated hierarchical reinforcement learning.

In summary, this thesis makes contributions in two key areas. Computational models of motivation and motivated reinforcement learning extend reinforcement learning to adaptive, multi-task learning in complex, dynamic environments. Motivated reinforcement learning agents allow the design of non-player characters for computer games that can progressively adapt their behaviour in response to changes in their environment.

Table of Contents

Index of Figures

Index of Tables

List of Abbreviations

CFG	Context-Free Grammar
GPI	Generalised Policy Interaction
HRL	Hierarchical Reinforcement Learning
HSOM	Habituated Self-Organising Map
IHDR	Incremental Hierarchical Discriminant Regression
MDP	Markov Decision Process
MEU	Maximum Expected Utility
MFRL	Motivated Flat Reinforcement Learning
MHRL	Motivated Hierarchical Reinforcement Learning
MMORL	Motivated Multi-Option Reinforcement Learning
MMORPG	Massively Multiplayer, Online Role-Playing Game
MORL	Multi-Option Reinforcement Learning
MRL	Motivated Reinforcement Learning
MSL	Motivated Supervised Learning
MUL	Motivated Unsupervised Learning
NPC	Non-Player Character
POMDP	Partially Observable Markov Decision Process
RL	Reinforcement Learning
RPG	Role-Playing Game
SL	Supervised Learning
SMDP	Semi-Markov Decision Process
SOM	Self-Organising Map
TD	Temporal Difference
UL	Unsupervised Learning