

Preface to the special issue: adaptive and learning agents

DAAN BLOEMBERGEN¹, TIM BRYNS² and LOGAN YLINIEMI³

¹*Department of Computer Science, University of Liverpool, Ashton Building, Ashton Street, Liverpool L69 3BX, UK;*
e-mail: d.bloembergen@liverpool.ac.uk;

²*Artificial Intelligence Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium;*
e-mail: timbryns@vub.ac.be;

³*Amazon Robotics, 300 Riverpark Drive, North Reading, MA 01864, USA;*
e-mail: loganyli@amazon.com

1 Introduction

Adaptive and learning agents (ALA) are able to optimize their behaviour in unknown and potentially changing environments, while using previous experience to improve their performance with respect to some evaluation measure. The community of ALA studies systems that are capable of acting autonomously and adapting to their surroundings.

While the development of a single learning agent may already present a serious challenge, current research frontiers also have a large focus on systems where multiple agents interact in a shared environment. Often, these systems are inherently decentralized, rendering a centralized single-agent learning approach infeasible. Examples of such systems are, for example, multi-robot set-ups, decentralized network routing, distributed load-balancing, electronic auctions, traffic control, and many others.

In multi-agent settings, agents not only have to deal with a dynamic environment, but also with other agents that act, learn, and change over time. When agent objectives are aligned and all agents try to achieve a common goal, coordination among the agents is still required to reach optimal results. When agents have conflicting goals, a clear optimal solution may no longer exist and an equilibrium between agent behaviours is generally sought. These issues have given rise to an important research track studying coordination mechanisms in multi-agent learning.

In addition, current research within the ALA community focusses on how agents can share experience with other agents, or how human operators can guide the learning process. Work in this direction falls under the scope of transfer learning, human–agent interaction, teaching, reward shaping, and advice.

This special issue contains selected papers from the 2016 ALA workshop, held as a satellite workshop at the Autonomous Agents and MultiAgent Systems conference in Singapore. The goal of the ALA workshop is to increase awareness and interest in adaptive agent research, encourage collaboration, and provide a representative overview of current research in the area of ALA. It aims at bringing together not only different areas of computer science (e.g. agent architectures, reinforcement learning, and evolutionary algorithms), but also different fields studying similar concepts (e.g. game theory, bio-inspired control, and mechanism design). The workshop serves as an interdisciplinary forum for the discussion of ongoing or completed work in ALA and multi-agent systems.

2 Contents of the special issue

This special issue contains four papers, carefully selected out of 21 initial workshop submission. All papers were presented at the ALA 2016 workshop and have been thoroughly reviewed and revised over two separate review rounds. The result provides an excellent overview of the current research directions and state of the art within the ALA community.

The first paper, *Multi-Agent Credit Assignment in Stochastic Resource Management Games* by Patrick Mannion, Sam Devlin, Jim Duggan, and Enda Howley, discusses the problem of credit assignment in multi-agent systems. Especially in cooperative multi-agent systems, it is not always straightforward to design a reward function for the individual agents that produces the desired result for the system as a whole. In this paper, different credit assignment structures are compared empirically in two stochastic resource management problems. The authors show that multi-agent systems that use credit assignment structures based on potential-based reward shaping (PBRS) can achieve near optimal performance, outperforming systems that use unshaped local or global rewards. In addition, the authors note interesting differences between PBRS with either state-based or action-based potential functions, and urge researchers to always test both approaches when implementing PBRS in their specific problem domain.

The second paper, *Autonomous UAV Landing in Windy Conditions with MAP-Elites* by Sierra Adibi, Scott Forer, Jeremy Fries, and Logan Yliniemi, investigates how a neural network controller can be trained to safely land a fixed-wing unmanned aerial vehicle in high-wind conditions. The authors use the MAP-Elites algorithm to evolve the weights of the neural network, and find that learning two separate controllers for higher and lower altitude yields better performance than learning a single controller. The resulting controllers are able to land the aircraft safely in adverse weather conditions.

The third paper, *Environmental Effects on Simulated Emotional and Moody Agents* by Joe Collenette, Katie Atkinson, Daan Bloembergen, and Karl Tuyls, studies the effect that simulated emotions and mood have on decision making within societies of self-interested mobile agents. Emotions are short-term directed feelings towards specific opponents whereas mood is long-term and undirected. The authors distinguish different emotional characters and investigate which character is most successful in different environments. Additionally, the authors find that the addition of mood increases cooperation in the Prisoner's Dilemma, with high moods leading to more rapid changes in behaviour, and the strength of this effect depending on the structure of the environment.

Finally, the paper *Limits and Limitations of No-Regret Learning in Games* by Barnabé Monnot and Georgios Piliouras studies the limits of no-regret dynamics in congestion games. No-regret learning has been subject to considerable interest in the multi-agent learning community in the past due to its simplicity in implementation, but has not been completely understood theoretically. In this paper, the authors contribute an algorithm to steer agents towards a Nash Equilibrium of a one-shot game whilst maintaining the no-regret property, and introduce the Price of Learning and Value of Learning metrics for analyzing how the class of coarse correlated equilibria relate to no-regret dynamics.