

Editorial: Special Section on Reinforcement Learning and Approximate Dynamic Programming

Approximate dynamic programming (ADP) is to compute near-optimal solutions to Markov decision problems (MDPs) with large or continuous spaces. In recent years, the research works on ADP have been brought together with the reinforcement learning (RL) community [1-4]. RL is a machine learning framework for solving sequential decision making problems that can also be modeled as the MDP formalism. The common objective of RL and ADP is to develop efficient algorithms for sequential decision making under uncertain complex conditions. Therefore, there are many potential applications of RL and ADP in real-world problems such as autonomous robots, intelligent control, resource allocation, network routing, etc.

This special section of JILSA focuses on key research problems emerging at the junction of RL and ADP. After a rigorous reviewing process, three papers were accepted for publication in this special section.

The first paper by M. A. Wiering [5] focuses on the applications of reinforcement learning with value function approximation in game playing. In the paper, three different schemes were studied for learning to play Backgammon with temporal difference learning. The three training schemes include: 1) self-play, 2) playing against an expert program, and 3) viewing experts play against each other. Extensive experimental results using temporal difference methods with neural networks were provided to compare the three learning schemes. It was illustrated that the drawback of learning from experts is that the learning program has few chances for exploration. The results also indicate that observing an expert play is the worst method and learning by playing against an expert seems to be the best strategy.

The second paper by J. H. Zaragoza, and E. F. Morales [6] proposed a relational reinforcement learning approach with continuous actions, called TS-RRLCA, which is based on the combination of behavioral cloning and locally weighted regression. The TS-RRLCA approach includes two main stages to learn continuous action policy for robots in partially known environments. The first stage is to develop a relational representation of robot states and actions and the rQ-learning algorithm is applied with behavioral cloning so that optimized control policies with discrete actions can be obtained efficiently. In the second stage, the learned policy is transformed into a relational policy with continuous actions through a

Locally Weighted Regression (LWR) process. The proposed method was successfully applied to a simulated and a real service robot for navigation and following tasks with different conditions.

The combination of reinforcement learning or approximate dynamic programming with learning from demonstration is studied in the third paper [7]. A learning strategy was proposed to generate a control field for a mobile robot in an unknown and uncertain environment, which integrates learning, generalization, and exploration into a unified architecture. Some Simulation results were provided to evaluate the performance of the proposed method.

Although RL and ADP provide efficient ways for developing machine intelligence in a trial-and-error manner, the incorporation of human intelligence is important for the successful applications of RL and ADP. In this special section on RL and ADP, all the three papers studied the relationships between machine intelligence and human intelligence in different aspects. The results of the first paper demonstrate that an expert program for game playing will be very helpful to develop computer programs using RL [5]. The usage of relational RL to incorporate human examples was investigated in the second paper [6]. In the third paper [7], the method of learning from human demonstration was employed to generate initial control field for an autonomous mobile robots. Therefore, the results in this special section will be good references for future research in related topics.

At last, I would like to thank all of the authors and reviewers who have made contributions to this special section.

Xin Xu
Editor-in-Chief,
JILSA

REFERENCES

- [1] F. Y. Wang, H. G. Zhang and D. R. Liu, "Adaptive Dynamic Programming: An Introduction," *IEEE Computational Intelligence Magazine*, May 2009, pp. 39-47.
- [2] W. B. Powell, "Approximate Dynamic Programming: Solving the Curses of Dimensionality," Wiley, Princeton,

- NJ, 2007.
- [3] R. S. Sutton and A. G. Barto, "Reinforcement Learning: an Introduction," MIT Press, Cambridge, MA, 1998.
- [4] D. P. Bertsekas and J. N. Tsitsiklis, "Neuro-Dynamic Programming. Belmont," Athena Scientific, MA, 1996.
- [5] M. A. Wiering, "Self-play and Using an Expert to Learn to Play Backgammon with Temporal Difference Learning," *Journal of Intelligent Learning Systems and Applications*, Vol. 2, 2010, pp. 55-66.
- [6] J. H. Zaragoza and E. F. Morales, "Relational Reinforcement Learning with Continuous Actions by Combining Behavioural Cloning and Locally Weighted Regression," *Journal of Intelligent Learning Systems and Applications*, Vol. 2, 2010, pp. 67-77.
- [7] D. Goswami and P. Jiang, "Experience Based Learning Controller," *Journal of Intelligent Learning Systems and Applications*, Vol. 2, 2010, pp. 78-83.