

Chen, Xin; Chen, Gang; Cao, Weihua; Wu, Min

Cooperative learning with joint state value approximation for multi-agent systems. (English)

Zbl 1299.93001

J. Control Theory Appl. 11, No. 2, 149-155 (2013).

Summary: This paper relieves the ‘curse of dimensionality’ problem, which becomes intractable when scaling reinforcement learning to multi-agent systems. This problem is aggravated exponentially as the number of agents increases, resulting in large memory requirement and slowness in learning speed. For cooperative systems which widely exist in multi-agent systems, this paper proposes a new multi-agent Q-learning algorithm based on decomposing the joint state and joint action learning into two learning processes, which are learning individual action and the maximum value of the joint state approximately. The latter process considers others’ actions to insure that the joint action is optimal and supports the updating of the former one. The simulation results illustrate that the proposed algorithm can learn the optimal joint behavior with smaller memory and faster learning speed compared with friend-Q learning and independent learning.

MSC:

- 93A14 Decentralized systems
- 93C85 Automated systems (robots, etc.) in control theory
- 68T05 Learning and adaptive systems in artificial intelligence
- 68T42 Agent technology and artificial intelligence

Keywords:

multi-agent system; Q-learning; cooperative system; curse of dimensionality; decomposition

Software:

R-MAX

Full Text: [DOI](#)

References:

- [1] G. Weiss. Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. Cambridge: MIT Press, 1999.
- [2] N. Vlassis. A concise introduction to multiagent systems and distributed artificial intelligence. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2007, 1(1): 1–71. · Zbl 1198.68265 · doi:10.2200/S00091ED1V01Y200705AIM002
- [3] M. Wu, W. Cao, J. Peng, et al. Balanced reactive-deliberative architecture for multi-agent system for simulation league of RoboCup. International Journal of Control, Automation and Systems, 2009, 7(6): 945–955. · doi:10.1007/s12555-009-0611-z
- [4] K. Tumer, A. Agogino. Improving air traffic management with a learning multiagent system. IEEE Intelligent Systems, 2009, 24(1):18–21. · Zbl 1343.68208 · doi:10.1109/MIS.2009.10
- [5] S. Proper, P. Tadepalli. Solving multiagent assignment Markov decision processes. Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems. Richland: IFAAMAS, 2009: 681–688.
- [6] J. R. Kok, M. T. J. Spaan, N. Vlassis. Non-communicative multi-robot coordination in dynamics environments. Robotics and Autonomous Systems, 2005, 50(2/3): 99–114. · Zbl 02223497 · doi:10.1016/j.robot.2004.08.003
- [7] M. L. Littman. Friend-or-Foe Q-learning in general-sum games. Proceedings of the 18th International Conference on Machine Learning. Williamstown: Morgan Kaufmann Press, 2001: 322–328.
- [8] X. Wang, T. Sandholm. Reinforcement learning to play an optimal Nash equilibrium in team Markov games. Proceedings of the Advances Neural Information Processing Systems. Cambridge: MIT Press, 2002: 1571–1578.
- [9] R. I. Brafman, M. Tennenholtz. R-Max-a general polynomial time algorithm for near-optimal reinforcement learning. Journal of Machine Learning Research, 2002, 3(2): 213–231. · Zbl 1088.68694
- [10] L. Busoniu, R. Babuska, B. De Schutter. A comprehensive survey of multi-agent reinforcement learning. IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, 2008, 38(2): 156–172. · doi:10.1109/TSMCC.2007.913919
- [11] N. Mehta, S. Natarajan, P. Tadepalli, et al. Transfer in variable-reward hierarchical reinforcement learning. Machine Learning, 2008, 73(3):289–312. · Zbl 05537392 · doi:10.1007/s10994-008-5061-y

- [12] J. R. Kok, N. Vlassis. Collaborative multiagent reinforcement learning by payoff propagation. *Journal of Machine Learning Research*, 2006, 7: 1789–1828. · [Zbl 1222.68235](#)
- [13] S. Kapetanakis, D. Kudenko. Reinforcement learning of coordination in cooperative multi-agent systems. *Proceedings of the 18th National Conference on Artificial Intelligence*. Washington: IEEE Computer Society, 2002: 326–331. · [Zbl 1032.68692](#)
- [14] C. Claus, C. Boutilier. The dynamics of reinforcement learning in cooperative multiagent systems. *Proceedings of the 15th National Conference on Artificial Intelligence*. Madison: AAAI Press, 1998:746–752. 746–752.
- [15] C. J. C. H. Watkins, P. Dayan. Q-learning. *Machine Learning*, 1992, 8(3/4): 279–292. · [Zbl 0773.68062](#)
- [16] C. S. Szepesvari, M. L. Littman. A unified analysis of value-function-based reinforcement-learning algorithms. *Neural Computation*, 1999, 11(8): 2017–2059. · [doi:10.1162/089976699300016070](#)
- [17] R. S. Sutton. Learning to predict by the method of temporal differences. *Machine Learning*, 1988, 3(1): 9–44.
- [18] A. Bab, R. I. Brafman. Multi-agent reinforcement learning in common interest and fixed sum stochastic games: an experimental study. *Journal of Machine Learning Research*, 2008, 9: 2635–2675. · [Zbl 1225.68145](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.