Guest Editorial Special Issue on Reinforcement Learning-Based Control: Data-Efficient and Resilient Methods

As an important branch of machine learning, reinforcement learning (RL) has proved its efficiency in many emerging applications in science and engineering. A remarkable advantage of RL is that it enables agents to maximize their cumulative rewards through online exploration and interactions with unknown (or partially unknown) and uncertain environments, which is regarded as a variant of data-driven adaptive optimal control methods. However, the successful implementation of RL-based control systems usually relies on a good quantity of online data due to its data-driven nature. Therefore, it is imperative to develop data-efficient RL methods for control systems to reduce the required number of interactions with the external environment. Moreover, network-aware issues, such as cyberattacks, dropout packet and communication latency, and actuator and sensor faults, are challenging conundrums that threaten the safety, security, stability, and reliability of network control systems. Consequently, it is significant to develop safe and resilient RL mechanisms.

This special issue focuses on promoting the development of cutting-edge RL and intelligent control systems, including theoretic studies, controller design, algorithm development, applications, and experimental validations, and encouraging the collaboration between researchers from diverse fields, such as computational intelligence, control systems, RL, and cybersecurity. We have received a large number of submissions to this special issue. After the peer-review process, 22 articles were accepted for publication that cover a wide variety of RL-based control topics from both theoretical and practical points of view. We separate these articles into four groups based on their research objectives.

The first group of articles aims at ameliorating the data efficiency of traditional RL methods based on the exploitation of historical data, prior knowledge, and replayed experience. Among them, Gao et al. [A6] built a transfer RL framework that enables the learning of controllers to employ the prior knowledge extracted from previously learned tasks and data, which improves the learning performance and data efficiency. In order to boost the sample efficiency of soft actor–critic, which is an off-policy RL strategy, Banerjee et al. [A3] applied a mixture of prioritized off-policy data with the latest on-policy data to train the policy and the value function networks. Yang et al. [A19] presented a composite critic learning mechanism, which combines the instantaneous data with the historical data, leading to an efficient, experience-replay-based RL approach. Li et al. [A5] have studied the adaptive optimal tracking control problem and developed a model-free Q-learning algorithm with a high efficiency of data utilization. Xu and Wu [A18] proposed a data-efficient off-policy RL approach for distributed output tracking control problems of heterogeneous multiagent systems with uncertain exosystem dynamics. Considering the grid emergency voltage control problem in electric power systems, a multiagent graph-attention-based deep RL algorithm is proposed to improve the decision accuracy in a data-efficient manner [A20].

The second group attempts to guarantee the resiliency, safety, and robustness of the closed-loop systems in terms of RL. Xie et al. [A17] investigated the problems of communication-efficient and resilient multiagent RL. They considered the scenario that some agents are adversarial captured by the Byzantine attack model and established the trade-off between optimality and resilience where Byzantine agents are present. Jiang et al. [A9] studied the adaptive optimal control problem of networked nonlinear systems and proposed an RL-based control approach so that the closed-loop system can handle stochastic sensor and actuator dropouts. Kokolakis and Vamvoudakis [A10] developed a safe pursuit-evasion game to enable finite-time capture, optimal performance, and adaption to an unknown cluttered environment in terms of Gaussian processes and RL. Modares et al. [A13] proposed a data-driven safe RL algorithm for discrete-time nonlinear systems to ensure both safety and stability of closed-loop systems. Yu et al. [A21] investigated the fault-tolerant formation control problem of networked fixed-wing unmanned aerial vehicles (UAVs) based on the integration of RL and fractional-order adaptive control.

The third group focuses on extensions of existing RL-based control methods, and combinations of RL and advanced control techniques, such as model predictive control (MPC) and distributed control. To be more specific, Jiang et al. [A22] introduced a λ-policy iteration method for optimal control problems of discrete-time linear systems in order to relax the reliance on an initial admissible control policy. A value iteration method is developed to solve the optimal control problem of discrete-time nonlinear systems with constrained cost [A16]. Wang et al. [A4] proposed a supplementary control approach for discrete-time nonlinear systems in the framework of goal representation heuristic dynamic programming. For the sake of realizing coordination of multiagent systems in an adaptive optimal sense, Wang et al. [A8] developed distributed controllers via RL to address semiglobal output regulation

Digital Object Identifier 10.1109/TNNLS.2024.3362092

problems in the presence of input saturation. Lan et al. [A12] proposed a distributed time-varying optimal formation control protocol for nonlinear uncertain multiagent systems. Lian et al. [A11] proposed a distributed minimax strategy for multiplayer games via RL, and the stability and robustness of the closed-loop systems have been discussed. The combination of RL and MPC appears in [A14] where the MPC serves as a policy generator and the RL is leveraged to evaluate the policy.

The fourth group is related to applications of RL-based control methods, including UAVs, industrial processes, and robotics. Bøhn et al. [A1] showed that deep RL can successfully learn to perform attitude control of fixed-wing UAVs with nonlinear dynamics. They have developed a data-efficient learning framework that yields flight-worthy attitude controllers within minutes of learning time. Li et al. [A2] modeled the temperature field control problem in roller skin via partial differential equations (PDEs) and proposed an event-triggered adaptive dynamic programming method for PDE systems. Liu et al. [A15] studied a dynamic operation optimization problem for a steelmaking process in terms of deep RL. Regarding the multirobot efficient search problem, Guo et al. [A7] proposed a distributional RL-based researcher with respect to moving targets.

We would like to express our sincere gratitude to all the authors and reviewers for their great dedication to this special issue. This special issue would not have been successful without their efforts. Finally, we highly appreciate the strong support from the Editor-in-Chief, Prof. Yongduan Song, the past Editor-in-Chief, Prof. Haibo He, and the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS's Editorial Office.

WEINAN GAO, *Guest Editor*
State Key Laboratory of Synthetical
Automation for Process Industries
Northeastern University
Shenyang, Liaoning 110819, China

NA LI, *Guest Editor*
School of Engineering and Applied Sciences
Harvard University
Allston, MA 02134 USA

KYRIAKOS G. VAMVOUDAKIS, *Guest Editor*
Daniel Guggenheim School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332 USA

FEI RICHARD YU, *Guest Editor*
Department of Systems and Computer Engineering
Carleton University
Ottawa, ON K1S 5B6, Canada

ZHONG-PING JIANG, *Guest Editor*
Tandon School of Engineering
New York University
Brookyn, NY 11201 USA

APPENDIX: RELATED ARTICLES

[A1] E. Bohn, E. M. Coates, D. Reinhardt, and T. A. Johansen, "Data-efficient deep reinforcement learning for attitude control of fixed-wing UAVs: Field experiments," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3168–3180, Mar. 2024.

[A2] B. Li, N. Chen, B. Luo, J. Chen, C. Yang, and W. Gui, "ADP-based event-triggered constrained optimal control on spatiotemporal process: Application to temperature field in roller kiln," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3229–3241, Mar. 2024.

[A3] C. Banerjee, Z. Chen, and N. Noman, "Improved soft actor-critic: Mixing prioritized off-policy samples with on-policy experiences," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3121–3129, Mar. 2024.

[A4] X. Wang, D. Ding, X. Ge, and Q.-L. Han, "Supplementary control for quantized discrete-time nonlinear systems under goal representation heuristic dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3202–3214, Mar. 2024.

[A5] C. Li, J. Ding, F. L. Lewis, and T. Chai, "Model-free Q-learning for the tracking problem of linear discrete-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3191–3201, Mar. 2024.

[A6] X. Gao, J. Si, and H. Huang, "Reinforcement learning control with knowledge shaping," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3156–3167, Mar. 2024.

[A7] H. Guo, Q. Peng, Z. Cao, and Y. Jin, "DRL-searcher: A unified approach to multirobot efficient search for a moving target," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3215–3228, Mar. 2024.

[A8] B. Wang, L. Xu, X. Yi, Y. Jia, and T. Yang, "Semiglobal suboptimal output regulation for heterogeneous multi-agent systems with input saturation via adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3242–3250, Mar. 2024.

[A9] Y. Jiang, L. Liu, and G. Feng, "Adaptive optimal control of networked nonlinear systems with stochastic sensor and actuator dropouts based on reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3107–3120, Mar. 2024.

[A10] N. T. Kokolakis and K. G. Vamvoudakis, "Safety-aware pursuit-evasion games in unknown environments using Gaussian processes and finite-time convergent reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3130–3143, Mar. 2024.

[A11] B. Lian, V. S. Donge, W. Xue, F. L. Lewis, and A. Davoudi, "Distributed minmax strategy for multiplayer games: Stability, robustness, and algorithms," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3265–3277, Mar. 2024.

[A12] J. Lan, Y.-J. Liu, D. Yu, G. Wen, S. Tong, and L. Liu, "Time-varying optimal formation control for second-order multiagent systems based on neural network observer and reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3144–3155, Mar. 2024.

[A13] A. Modares, N. Sadati, B. Esmaeili, F. A. Yaghmaie, and H. Modares, "Safe reinforcement learning via a model-free safety certifier," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3302–3311, Mar. 2024.

[A14] M. Lin, Z. Sun, Y. Xia, and J. Zhang, "Reinforcement learning-based model predictive control for discrete-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3312–3324, Mar. 2024.

[A15] C. Liu, L. Tang, and C. Zhao, "A novel dynamic operation optimization method based on multiobjective deep reinforcement learning for steelmaking process," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3325–3339, Mar. 2024.

[A16] Q. Wei and T. Li, "Constrained-cost adaptive dynamic programming for optimal control of discrete-time nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3251–3264, Mar. 2024.

[A17] Y. Xie, S. Mou, and S. Sundaram, "Communication-efficient and resilient distributed Q-learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3351–3364, Mar. 2024.

[A18] Y. Xu and Z.-G. Wu, "Data-efficient off-policy learning for distributed optimal tracking control of HMAS with unidentified exosystem dynamics," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3181–3190, Mar. 2024.

[A19] Y. Yang, Y. Pan, C.-Z. Xu, and D. C. Wunsch, "Hamiltonian-driven adaptive dynamic programming with efficient experience replay," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3278–3290, Mar. 2024.

[A20] Y. Zhang, M. Yue, J. Wang, and S. Yoo, "Multi-agent graph-attention deep reinforcement learning for post-contingency grid emergency voltage control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3340–3350, Mar. 2024.

[A21] Z. Yu, J. Li, Y. Xu, Y. Zhang, B. Jiang, and C.-Y. Su, "Reinforcement learning-based fractional-order adaptive fault-tolerant formation control of networked fixed-wing UAVs with prescribed performance," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3365–3379, Mar. 2024.

[A22] H. Jiang, B. Zhou, and G.-R. Duan, "Modified λ-policy iteration based adaptive dynamic programming for unknown discrete-time linear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3291–3301, Mar. 2024.

Weinan Gao (Senior Member, IEEE) received the B.Sc. degree in automation and the M.Sc. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2011 and 2013, respectively, and the Ph.D. degree in electrical engineering from New York University, Brooklyn, NY, USA, in 2017.

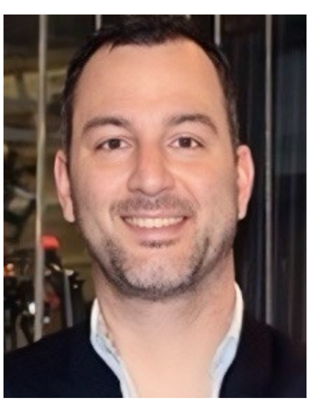
He is currently a Professor with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. Previously, he was an Assistant Professor of mechanical and civil engineering at the Florida Institute of Technology, Melbourne, FL, USA; an Assistant Professor of electrical and computer engineering at Georgia Southern University, Statesboro, GA, USA; and a Visiting Professor with the Mitsubishi Electric Research Laboratory (MERL), Cambridge, MA, USA. His research interests include reinforcement learning, adaptive dynamic programming (ADP), optimal control, cooperative adaptive cruise control (CACC), intelligent transportation systems, sampled-data control systems, and output regulation theory.

Prof. Gao was a recipient of the Best Paper Award at the IEEE Data Driven Control and Learning Systems (DDCLS) Conference in 2023, the International Conference on Control, Automation and Information Sciences (ICCAIS) in 2023, and the IEEE International Conference on Real-time Computing and Robotics (RCAR) in 2018, and the David Goodman Research Award at New York University in 2019. He is an Associate Editor of IEEE Transactions on Neural Networks and Learning Systems, IEEE/Chinese Association of Automation (CAA) Journal of Automatica Sinica, *Control Engineering Practice*, and IEEE Transactions on Circuits and Systems—II: Express Briefs, a member of Editorial Board of *Neural Computing and Applications*, and a Technical Committee Member of IEEE Control Systems Society on Nonlinear Systems and Control, IEEE Control Systems Society on Smart Cities, IFAC TC 1.2 Adaptive and Learning Systems, IFAC TC 2.3 Non-Linear Control Systems, and Chinese Association for Artificial Intelligence (CAAI) Industrial Artificial Intelligence.

Na Li (Senior Member, IEEE) received the bachelor's degree in mathematics from Zhejiang University, Hangzhou, China, in 2007, and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2013.

From 2013 to 2014, she was a Post-Doctoral Associate with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. She is currently the Winokur Family Professor of Electrical Engineering and Applied Mathematics at Harvard University, Cambridge. She has held a variety of short-term visiting appointments, including the Simons Institute for the Theory of Computing, Berkeley, CA, USA, MIT, and Google Brain, Mountain View, CA, USA. Her research interests include the control, learning, and optimization of networked systems, including theory development, algorithm design, and applications to real-world cyber-physical societal systems.

Dr. Li was a recipient of the NSF Career Award in 2016, the AFSOR Young Investigator Award in 2017, the ONR Young Investigator Award in 2019, the Donald P. Eckman Award in 2019, the McDonald Mentoring Award in 2020, and the International Federation of Automatic Control (IFAC) Manfred Thoma Medal in 2023, along with some other awards.



Kyriakos G. Vamvoudakis (Senior Member, IEEE) was born in Athens, Greece. He received the Diploma (a five-year degree, equivalent to a Master of Science) degree (Hons.) in electronic and computer engineering from the Technical University of Crete, Kounoupidiana, Greece, in 2006, and the M.S. and Ph.D. degrees in electrical engineering The University of Texas at Arlington, Arlington, TX, USA, in 2008 and 2011, respectively, with Frank L. Lewis as his advisor.

From May 2011 to January 2012, he worked as an Adjunct Professor and a Faculty Research Associate at The University of Texas at Arlington and the Automation and Robotics Research Institute, Fort Worth, TX, USA. From 2012 to 2016, he was a Project Research Scientist at the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA, USA. He was an Assistant Professor with the Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA, USA, until

2018. He is currently the Dutton-Ducoffe Endowed Professor at the Daniel Guggenheim School of Aerospace Engineering, Georgia Tech, Atlanta, GA, USA. He holds a secondary appointment at the School of Electrical and Computer Engineering. His expertise is in reinforcement learning, control theory, game theory, cyber-physical security, bounded rationality, and safe/assured autonomy.

Dr. Vamvoudakis was a recipient of the 2019 ARO YIP Award, the 2018 NSF CAREER Award, the 2018 DoD Minerva Research Initiative Award, and the 2021 GT Chapter Sigma Xi Young Faculty Award. His work has been recognized with best paper nominations and several international awards, including the 2016 International Neural Network Society Young Investigator (INNS) Award, the Best Paper Award for Autonomous/Unmanned Vehicles at the 27th Army Science Conference in 2010, the Best Presentation Award at the World Congress of Computational Intelligence in 2010, and the Best Researcher Award from the Automation and Robotics Research Institute in 2011. He had served as a Guest Editor for IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING (Special Issue on Learning from Imperfect Data for Industrial Automation), IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (Special Issue on Reinforcement Learning Based Control: Data-Efficient and Resilient Methods), IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS (Special Issue on Industrial Artificial Intelligence for Smart Manufacturing), and IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (Special Issue on Unmanned Aircraft System Traffic Management). He is an Associate Editor of *Automatica*, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE COMPUTATIONAL INTELLIGENCE *Magazine*, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, *Neurocomputing*, *Journal of Optimization Theory and Applications*, and *Frontiers in Control Engineering—Adaptive*, *Robust and Fault Tolerant Control*.

Fei Richard Yu (Fellow, IEEE) received the Ph.D. degree in electrical engineering from The University of British Columbia (UBC), Vancouver, BC, Canada, in 2003.

He is currently with the Guangdong Laboratory of Artificial Intelligence and Digital Economy, Shenzhen, China.

Mr. Yu is a Distinguished Lecturer, the Vice President-Membership, and an elected member of the Board of Governors (BoG), IEEE Vehicular Technology Society. He serves on the Editorial Board for several journals, including the Co-Editor-in-Chief for *Ad Hoc and Sensor Wireless Networks* and the Lead Series Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and IEEE COMMUNICATIONS SURVEYS AND TUTORIALS. He has served as the technical program committee (TPC) co-chair for numerous conferences. He is a registered Professional Engineer in the Province of Ontario, Canada.

Zhong-Ping Jiang (Fellow, IEEE) received the M.Sc. degree in statistics from the University of Paris XI, Paris France, in 1989, and the Ph.D. degree in automatic control and mathematics from ParisTech-Mines, Paris, in 1993, under the direction of Laurent Praly.

He is currently a Professor of electrical and computer engineering with the Tandon School of Engineering, New York University, New York, NY, USA. His main research interests include stability theory, robust/adaptive/distributed nonlinear control, robust adaptive dynamic programming, reinforcement learning, and their applications to information, mechanical, and biological systems. In these fields, he has written six books and is the author/coauthor of more than 500 peer-reviewed journal articles and conference papers.

Prof. Jiang is a fellow of International Federation of Automatic Control (IFAC), Chinese Association of Automation (CAA), and Asia-Pacific Artificial Intelligence Association (AAIA), and a foreign member of the Academia Europaea (Academy of Europe) and the European Academy of Sciences and Arts. He is among the Clarivate Analytics Highly Cited Researchers and Stanford's top 2% most highly cited scientists. In 2022, he received the Excellence in Research Award from the NYU Tandon School of Engineering. He was a recipient of the prestigious Queen Elizabeth II Fellowship Award from the Australian Research Council, the CAREER Award from the U.S. National Science Foundation, the JSPS Invitation Fellowship from the Japan Society for the Promotion of Science, the Distinguished Overseas Chinese Scholar Award from the National Natural Science Foundation of China, and several best paper awards. He has served as the Deputy Editor-in-Chief, a senior editor, and an associate editor for numerous journals.