



**ADAPTIVE DYNAMIC PROGRAMMING
AND REINFORCEMENT LEARNING
DEVELOPMENT TRENDS AND APPLICATIONS**

**自适应动态规划与强化学习
发展现状与应用进展**

中国自动化学会
Chinese Association of Automation
自适应动态规划与强化学习专业委员会
TC on Adaptive Dynamic Programming and Reinforcement Learning
2026年7月

目录

主编：刘德荣

Editor-in-Chief: Derong Liu

副主编：张化光、徐昕、杨秦敏、魏庆来

Deputy Editors-in-Chief: Huaguang Zhang, Xin Xu, Qinmin Yang, Qinglai Wei

编委会

白伟伟、董璐、姜钟平、李宏亮、梁明明、林铭铎、刘艳军、罗彪、罗艳红、吕永峰、穆朝絮、那靖、宋睿卓、孙景亮、王鼎、夏宏兵、许斌、徐勇、薛珊、杨雄、杨永亮、杨湛宇、张坤、张顺超、张勇威、赵博、赵冬斌

Editorial Board

Weiwei Bai, Lu Dong, Zhong-Ping Jiang, Hongliang Li, Mingming Liang, Mingduo Lin, Yanjun Liu, Biao Luo, Yanhong Luo, Yongfeng Lv, Chaoxu Mu, Jing Na, Ruizhuo Song, Jingliang Sun, Ding Wang, Hongbing Xia, Bin Xu, Yong Xu, Shan Xue, Xiong Yang, Yongliang Yang, Zhanyu Yang, Kun Zhang, Shunchao Zhang, Yongwei Zhang, Bo Zhao, Dongbin Zhao



中国自动化学会

Chinese Association of Automation

自适应动态规划与强化学习专业委员会

TC on Adaptive Dynamic Programming and Reinforcement Learning

2026年7月

1

溯源拾光·ADPRL发展简史

1

2

铸魂基石·ADPRL基础理论、核心要素与应用

7

2.1 动态规划方法及其近似解法

7

2.2 强化学习

9

2.3 自适应动态规划

9

2.4 自适应动态规划应用概述

12

3

硕果华章·璀璨星河·ADPRL专委会学术成果

16

3.1 著作

16

3.2 期刊论文

16

3.3 授权发明专利

37

3.4 学术奖励

42

4

聚力同行·携手芳华·ADPRL专委会组织构架、职能与活动回顾

44

4.1 中国自动化学会自适应动态规划与强化学习专业委员会简介

44

4.2 中国自动化学会第二届自适应动态规划与强化学习专业委员会组织构架

45

4.3 中国自动化学会自适应动态规划与强化学习专业委员会活动回顾

46

4.3.1 大事件：2016自适应动态规划与强化学习专业委员会成立大会

46

4.3.2 前沿论坛：2017第三期自动化前沿论坛暨自适应动态规划与强化学习研讨会

49

4.3.3 前沿论坛：2018自适应动态规划与强化学习研讨会暨第十八期自动化前沿论坛

52

4.3.4 学术会议：2019自适应动态规划与强化学习研讨会

54

4.3.5 前沿论坛：2020自适应动态规划与强化学习研讨会暨第二十期自动化前沿论坛

56

4.3.6 大事件：2021自适应动态规划与强化学习专业委员会换届大会暨学术研讨会

59

4.3.7 专题论坛：2022自适应动态规划与强化学习专业委员会青年论坛

61

4.3.8 专题论坛：2021中国自动化大会“自适应动态规划与强化学习”专题论坛

62

4.3.9 专题论坛：2022 ADPRL研讨会及中国自动化大会“元控制”专题论坛

63

4.3.10 学术会议：2023自适应动态规划与强化学习研讨会

65

4.3.11 学术会议：2023强化学习与自适应动态规划前沿进展研讨会

67

4.3.12 学术会议：2024自适应动态规划与强化学习前沿研讨会

68

4.3.13 学术会议：2024自适应动态规划与强化学习研讨会

70

4.3.14 走进系列：2024 自适应动态规划与强化学习系列研讨会“走进衢州”

72

4.3.15 走进系列：2025自适应动态规划与强化学习专委会“走进安徽大学”学术活动

73

4.3.16 学术会议：2025自适应动态规划与强化学习研讨会

77

4.3.17 走进系列：2026 自适应动态规划与强化学习专委会“走进北航”学术活动

79

4.3.18 走进系列：2026 自适应动态规划与强化学习专委会“走进海南大学”学术活动

81

ADPRL 发展简史

1952 年 Bellman 在 PNAS 发表第一篇动态规划期刊论文^[1]
1957 年 Bellman 出版动态规划首部专著^[2]
1958 年 Bellman 在 PNAS 发表论文提出动态规划近似求解方法^[3]
1960–1962 年 Howard、Bellman、Dreyfus 提出策略迭代动态规划近似求解方法^[4,5,6]
1967 年 Larson 在 TAC 发表近似动态规划综述论文^[7]，并于 1968 年出版相关专著一部^[8]
1977 年 Werbos 在 GSY 提出启发式动态规划^[9]，之后命名为自适应评判设计^[10,11]
1983 年 Barto、Sutton、Anderson 在 SMC 发表论文，用强化学习方法成功控制倒立摆^[12]
1987 年 Werbos 发表 SMC 论文再次阐述 10 年前 GSY 论文中提出的启发式动态规划^[13]
1988 年 Sutton 提出时间差分（TD）强化学习算法^[14]
1992 年 Watkins 和 Dayan 提出 Q-学习算法^[15]
1992 年王飞跃和 Saridis 提出近似动态规划迫敛解法^[16,17]
1996 年 Bertsekas 和 Tsitsiklis 出版《神经动态规划》专著^[18]
1997 年 Prokhorov 和 Wunsch 在 TNN 发表自适应评判设计综述论文^[19]
1998 年 Sutton 和 Barto 出版《强化学习》专著^[20]
2002 年 Murray 等人在 SMCC 发表自适应动态规划论文^[21]，次年发表完整理论证明^[22]
2002 年美国国家基金会 NSF 在墨西哥召开“学习与近似动态规划”研讨会^[23]
2004 年出版 NSF 墨西哥研讨会论文集《学习与近似动态规划手册》^[24]
2006 年美国国家基金会 NSF 在墨西哥再一次召开近似动态规划研讨会^[25]
2006 年 Lincoln 和 Rantzer 在 TAC 论文中给出 RDP 算法并证明其收敛性和最优性^[26]
2007 年 Powell 出版《近似动态规划：解决维数灾难》专著^[27]
2007 年 IEEE CIS 创办每两年一届的 ADPRL Symposium，2014 年开始改为每年一届^[28]
2008 年 IEEE CIS 创办 ADPRL 技术委员会/Technical Committee，刘德荣任创始主任^[29]
2008 年 SMCB 出版 ADPRL 专刊，Lewis 等人论文证明迭代 ADP 算法的收敛性和最优性^[30,31]
2009 年王飞跃、张化光、刘德荣委员在 IEEE CIS Magazine 发表 ADP 综述论文^[32]
2009 年 Lewis 和 Vrabie 在 IEEE CAS Magazine 发表 ADPRL 综述论文^[33]
2010 年徐昕委员出版《增强学习与近似动态规划》专著^[34]
2011 年 JCTA 出版 ADPRL 专刊^[35]
2013 年张化光、刘德荣委员出版英文专著《自适应动态规划控制：算法和稳定性》^[36]
2013 年 Lewis 和刘德荣委员出版英文编著《近似动态规划与强化学习反馈控制》^[37]
2014 年 Silver 等人在第 31 届国际机器学习大会上提出确定性策略梯度算法^[38]
2015 年 Google’s AlphaGo 击败欧洲围棋冠军樊麾^[39]
2016 年 Google’s AlphaGo 击败世界围棋冠军李世石^[40]
2016 年中国自动化学会创建 ADPRL 专业委员会，刘德荣任创始主任^[41]
2017 年 Google’s DeepMind 发布 AlphaGo Zero^[42]
2017 年 Google’s AlphaGo 击败排名第一的世界围棋冠军柯洁^[43]
2018 年 TNNLS 出版 Deep RL 和 ADP 专刊^[44]
2020 年 Google’s DeepMind 发布 AlphaFold^[45]
2021 年 Google’s DeepMind 发布 AlphaFold 升级版^[46]
2022 年 OpenAI 发布 ChatGPT，引发全球大语言模型热潮^[47]
2024 年 Hopfield 和 Hinton 获得诺贝尔物理学奖^[48]
2024 年 Baker、Hassabis 和 Jumper 获得诺贝尔化学奖^[49]
2024 年 Barto 和 Sutton 获图灵奖^[50]

参考文献

[1] R. Bellman, “On the theory of dynamic programming,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 8, pp. 716–719, Aug. 1952.
[2] R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.
[3] R. Bellman, “Dynamic programming, successive approximations, and monotone convergence,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 44, no. 6, pp. 578–580, Jun. 1958.
[4] R. A. Howard, *Dynamic Programming and Markov Processes*. New York: Wiley, 1960.
[5] R. Bellman, *Adaptive Control Processes*. Princeton, NJ: Princeton University Press, 1961.
[6] R. Bellman and S. Dreyfus, *Applied Dynamic Programming*. Princeton, NJ: Princeton University Press, 1962.
[7] R. E. Larson, “A survey of dynamic programming computational procedures,” *IEEE Transactions on Automatic Control*, vol. 12, no. 6, pp. 767–774, Dec. 1967.
[8] R. E. Larson, *State Increment Dynamic Programming*. New York: Elsevier, 1968.
[9] P. J. Werbos, “Advanced forecasting methods for global crisis warning and models of intelligence,” *General Systems Yearbook*, vol. 22, pp. 25–38, 1977.
[10] P. J. Werbos, “A menu of designs for reinforcement learning over time,” in *Neural Networks for Control*, W. T. Miller, R. S. Sutton, and P. J. Werbos, Eds., Cambridge, MA: The MIT Press, 1990, Chapter 3.
[11] P. J. Werbos, “Approximate dynamic programming for real-time control and neural modeling,” in *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, D. A. White and D. A. Sofge, Eds., New York, NY: Van Nostrand Reinhold, 1992, Chapter 13.
[12] A. G. Barto, R. S. Sutton, and C. W. Anderson, “Neuronlike adaptive elements that can solve difficult learning control problems,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 5, pp. 834–846, Sep./Oct. 1983.
[13] P. J. Werbos, “Building and understanding adaptive systems: A statistical/numerical approach to factory automation and brain research,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-17, no. 1, pp. 7–20, Jan. 1987.
[14] R. S. Sutton, “Learning to predict by the methods of temporal differences,” *Machine Learning*, vol. 3, pp. 9–44, Aug. 1988.
[15] C. J. C. H. Watkins and P. Dayan, “Q-learning,” *Machine Learning*, vol. 8, pp. 279–292, May 1992.
[16] F.-Y. Wang and G. N. Saridis, “Suboptimal control for nonlinear stochastic systems,” in *Proc. 31st IEEE Conference on Decision and Control*, Tucson, AZ, Dec. 1992, pp. 1856–1861.
[17] G. N. Saridis and F.-Y. Wang, “Suboptimal control of nonlinear stochastic systems,” *Control-Theory and Advanced Technology*, vol. 10, no. 4, pp. 847–871, Dec. 1994.
[18] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Belmont, MA: Athena Scientific, 1996.
[19] D. V. Prokhorov and D. C. Wunsch, “Adaptive critic designs,” *IEEE Transactions on Neural Networks*, vol. 8, pp. 997–1007, Sep. 1997.
[20] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press, 1998.
[21] J. J. Murray, C. J. Cox, G. G. Lendaris, and R. Saeks, “Adaptive dynamic programming,” *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 32, no. 2, pp. 140–153, May 2002.
[22] J. J. Murray, C. J. Cox, and R. E. Saeks, “The adaptive dynamic programming theorem,” in

Stability and Control of Dynamical Systems with Applications, D. Liu and P. J. Antsaklis, Eds., Boston, MA: Birkhauser, 2003, Chapter 19.

- [23] <https://grantome.com/grant/NSF/ECCS-0223696>
- [24] J. Si, A. G. Barto, W. B. Powell, and D. Wunsch, Eds., *Handbook of Learning and Approximate Dynamic Programming*. Piscataway, NJ: IEEE Press-Wiley, 2004.
- [25] <https://grantome.com/grant/NSF/ECCS-0541949>
- [26] B. Lincoln and A. Rantzer, “Relaxing dynamic programming,” *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1249–1260, Aug. 2006.
- [27] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken, NJ: Wiley, 2007.
- [28] <https://ewh.ieee.org/conf/ssci/2007/ssci-2007.pdf>
- [29] <https://cis.ieee.org/activities/technical-activities/>
- [30] F. L. Lewis, D. Liu, and G. G. Lendaris, Guest Editorial, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, special issue on Adaptive Dynamic Programming and Reinforcement Learning in Feedback Control, vol. 38, no. 4, pp. 896–897, Aug. 2008.
- [31] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, “Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 4, pp. 943–949, Aug. 2008.
- [32] F.-Y. Wang, H. Zhang, and D. Liu, “Adaptive dynamic programming: An introduction,” *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 39–47, May 2009.
- [33] F. L. Lewis and D. Vrabie, “Reinforcement learning and adaptive dynamic programming for feedback control,” *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, Aug. 2009.
- [34] 徐昕, *增强学习与近似动态规划*. 北京: 科学出版社, 2010.
- [35] S. Ferrari, J. Sarangapani, and F. L. Lewis, Editorial, *Journal of Control Theory and Applications*, special issue on Approximate Dynamic Programming and Reinforcement Learning, vol. 9, no. 3, p. 309, Aug. 2011.
- [36] H. Zhang, D. Liu, Y. Luo, and D. Wang, *Adaptive Dynamic Programming for Control: Algorithms and Stability*. London: Springer-Verlag, 2013.
- [37] F. L. Lewis and D. Liu, Eds., *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. Hoboken, NJ: Wiley, 2013.
- [38] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, “Deterministic policy gradient algorithms,” in *Proc. 31st International Conference on Machine Learning*, Beijing, China, Jun. 2014, PMLR, vol. 32, no. 1, pp. 387–395.
- [39] D. Silver, A. Huang, C. Maddison, et al., “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.
- [40] <https://www.bbc.co.uk/news/technology-35785875>
- [41] <https://www.caa.org.cn/article/205/1092.html>
- [42] D. Silver, J. Schrittwieser, K. Simonyan, et al., “Mastering the game of Go without human knowledge,” *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.
- [43] <https://www.bbc.co.uk/news/technology-40042581>
- [44] D. Zhao, D. Liu, F. L. Lewis, J. C. Principe, and S. Squartini, Guest Editorial, *IEEE Transactions on Neural Networks and Learning Systems*, special issue on Deep Reinforcement Learning and Adaptive Dynamic Programming, vol. 29, no. 6, pp. 2038–2041, Jun. 2018.
- [45] A. W. Senior, R. Evans, J. Jumper, et al., “Improved protein structure prediction using potentials from deep learning,” *Nature*, vol. 577, no. 7792, pp. 706–710, Jan. 2020.
- [46] J. Jumper, R. Evans, A. Pritzel, et al., “Highly accurate protein structure prediction with AlphaFold,” *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021.
- [47] <https://www.bbc.com/news/technology-63861322>
- [48] <https://www.nobelprize.org/prizes/physics/2024/summary/>
- [49] <https://www.nobelprize.org/prizes/chemistry/2024/summary/>
- [50] <https://awards.acm.org/about/2024-turing/>