圆桌论道|ICML2021强化学习相关189篇论文分类整理

作者：RLCN公众号 二维码：

Theory: non-stationary
Deep Reinforcement Learning amidst Continual Structured Non-Stationarity Annie Xie (Stanford University) · James Harrison (Stanford University) · Chelsea Finn (Stanford)
Theory: optimality
Joint Online Learning and Decision-making via Dual Mirror Descent Alfonso Lobos Ruiz (Microsoft) · Paul Grigas (UC Berkeley) · Zheng Wen (DeepMind)
UCB Momentum Q-learning: Correcting the bias without forgetting Pierre MENARD (Inria) · Omar Darwiche Domingues (Inria) · Xuedong Shang (Inria) · Michal Valko (DeepMind / Inria / ENS Paris-Saclay)
Kernel-Based Reinforcement Learning: Finite-Time Analysis for a Practical Algorithm Omar Darwiche Domingues (Inria) · Pierre Menard (Inria) · Matteo Pirotta (Facebook AI Research) · Emilie Kaufmann (CNRS, Univ. Lille) · Michal Valko (DeepMind / Inria / ENS Paris-Saclay)
Shortest-Path Constrained Reinforcement Learning for Sparse Reward Tasks Sungryull Sohn (University of Michigan) · Sungtae Lee (Yonsei University) · Jongwook Choi (University of Michigan) · Harm van Seijen (Microsoft Research) · Mehdi Fatemi (Microsoft Research) · Honglak Lee (Google / U. Michigan)
Modularity in Reinforcement Learning via Algorithmic Independence in Credit Assignment Michael Chang (UC Berkeley) · Sid Kaushik (UCB) · Sergey Levine (UC Berkeley) · Thomas Griffiths (Princeton University)
Doubly Robust off-policy Actor-Critic: Convergence and Optimality Tengyu Xu (The Ohio State University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U) · Yingbin LIANG (The Ohio State University)
Solving Challenging Dexterous Manipulation Tasks With Trajectory Optimisation and Reinforcement Learning Henry Charlesworth (University of Warwick) · Giovanni Montana (University of Warwick)
Density Constrained Reinforcement Learning Zengyi Qin (MIT) · Yuxiao Chen (California Institute of Technology) · Chuchu Fan (MIT)
Self-Paced Context Evaluation for Contextual Reinforcement Learning Theresa Eimer (Leibniz Universität Hannover) · André Biedenkapp (University of Freiburg) · Frank Hutter (University of Freiburg and Bosch Center for Artificial Intelligence) · Marius Lindauer (Leibniz University Hannover)
Policy Optimization via Differentiable Simulation Miguel Angel Zamora Mora (ETH Zurich) · Momchil Peychev (ETH Zurich) · Sehoon Ha (Georgia Institute of Technology) · Martin Vechev (ETH Zurich) · Stelian Coros (ETH Zurich)
Revisiting Rainbow: Promoting more insightful and inclusive deep Reinforcement Learning research Johan Obando Ceron (UAO) · Pablo Samuel Castro (Google Brain)
Logarithmic Regret for Reinforcement Learning with Linear Function Approximation Jiafan He (University of California, Los Angeles) · Dongruo Zhou (UCLA) · Quanquan Gu (University of California, Los Angeles)
Detecting Rewards Deterioration in Episodic Reinforcement Learning Ido Greenberg (Technion) · Shie Mannor (Technion)
Revisiting Peng's Q(λ) for Modern Reinforcement Learning Tadashi Kozuno (University of Alberta) · Yunhao Tang (Columbia University) · Mark Rowland (DeepMind) · Remi Munos (DeepMind) · Steven Kapturowski (Deepmind) · Will Dabney (DeepMind) · Michal Valko (DeepMind / Inria / ENS Paris-Saclay) · David Abel (DeepMind)
Ensemble Bootstrapping for Q-Learning Oren Peer (Technion) · Chen Tessler (Technion) · Nadav Merlis (Technion) · Ron Meir (Technion Israeli Institute of Technology)
Phasic Policy Gradient Karl Cobbe (OpenAI) · Jacob Hilton (OpenAI) · Oleg Klimov (OpenAI) · John Schulman (OpenAI)
SUNRISE: A Simple Unified Framework for Ensemble Learning in Deep Reinforcement Learning Kimin Lee (UC Berkeley) · Michael Laskin (UC Berkeley) · Aravind Srinivas (UC Berkeley) · Pieter Abbeel (UC Berkeley & Covariant)
Generative Adversarial Networks for Markovian Temporal Dynamics: Stochastic Continuous Data Generation Sung Woo Park (Chung-Ang Univ., Korea) · Dong Wook Shu (Chung-Ang Univ., Korea) · Junsoek Kwon (Chun-Ang University)
Dynamic Planning and Learning under Recovering Rewards David Simchi-Levi (MIT) · Zeyu Zheng (University of California, Berkeley) · Feng Zhu (Massachusetts Institute of Technology)
Objective Bound Conditional Gaussian Process for Bayesian Optimization Taewon Jeong (KAIST) · Heeyoung Kim (KAIST)
Guided Exploration with Proximal Policy Optimization using a Single Demonstration Gabriele Libardi (Pompeu Fabra University) · Gianni De Fabritiis (Universitat Pompeu Fabra) · Sebastian Dittert (Universitat Pompeu Fabra)
Improved Regret Bounds of Bilinear Bandits using Action Space Dimension Analysis Jang Kyoungseok (KAIST) · Kwang-Sung Jun (University of Arizona) · Se-Young Yun (KAIST) · Wanmo Kang (KAIST)
Learning and Planning in Average-Reward Markov Decision Processes Yi Wan (University of Alberta) · Abhishek Naik (University of Alberta) · Richard Sutton (DeepMind / Univ Alberta)
Decision-Making Under Selective Labels: Optimal Finite-Domain Policies and Beyond Dennis Wei (IBM Research)
Discovering symbolic policies with deep Reinforcement Learning Sookyung Kim (Lawrence Livermore National Laboratory) · Mikel Landajuela (Lawrence Livermore National Laboroatory) · Brenden Petersen (Lawrence Livermore National Laboratory) · Claudio Santiago (LLNL) · Ruben Glatt (LLNL) · Nathan Mundhenk (Lawrence Livermore National Labs) · Jacob Pettit (Lawrence Livermore National Laboratory) · Daniel Faissol (Lawrence Livermore National Laboratory)

Spectral Normalisation for Deep Reinforcement Learning: An Optimisation Perspective
Florin Gogianu (Bitdefender) · Tudor Berariu (Imperial College London) · Mihaela Rosca (DeepMind) · Claudia Clopath (Imperial College London) · Lucian Busoniu (Technical University of Cluj-Napoca) · Razvan Pascanu (DeepMind)

Posterior value functions: Hindsight Baselines for Policy Gradient Methods
Chris Nota (University of Massachusetts Amherst) · Philip Thomas (University of Massachusetts Amherst) · Bruno C. da Silva (University of Massachusetts)

Structured World Belief for Reinforcement Learning in POMDP
Gautam Singh (Rutgers University) · Skand Peri (Rutgers University, New Jersey) · Junghyun Kim (Rutgers University) · Hyunseok Kim (Electronics and Telecommunications Research Institute (ETRI), Korea) · Sungjin Ahn (Rutgers University)

EMaQ: Expected-Max Q-Learning Operator for Simple Yet Effective Offline and Online RL
Seyed Kamyar Seyed Ghasemipour (University of Toronto) · Dale Schuurmans (Google / University of Alberta) · Shixiang Gu (Google)

Deep Coherent Exploration For Continuous Control
Yijie Zhang (University of Copenhagen) · Herke van Hoof (University of Amsterdam)

Randomized Exploration in Reinforcement Learning with General value function Approximation
Haque Ishfaq (MILA / McGill University) · Qiwen Cui (Peking University) · Alex Ayoub (University of Alberta) · Viet Nguyen (McGill, Mila) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U) · Doina Precup (McGill University / DeepMind) · Lin Yang (UCLA)

Recomposing the Reinforcement Learning Building Blocks with Hypernetworks
shai keynan (Bar Ilan University) · Elad Sarafian (Bar-Ilan University) · Sarit Kraus (Bar-Ilan University)

Quantum algorithms for Reinforcement Learning with a generative model
Ashish Kapoor (Microsoft Research) · Robin Kothari (Microsoft) · Martin Roetteler (Microsoft) · Aarthi Sundaram (Microsoft) · Daochen Wang (University of Maryland)

Emphatic Algorithms for Deep Reinforcement Learning
Ray Jiang (DeepMind) · Tom Zahavy (DeepMind) · Zhongwen Xu (DeepMind) · Adam White (Deepmind, University of Alberta) · Matteo Hessel (DeepMind) · Charles Blundell (DeepMind) · Hado van Hasselt (DeepMind)

Beyond Variance Reduction: Understanding the True Impact of Baselines on Policy Optimization
Wesley Chung (Mila / McGill University) · Valentin Thomas (MILA) · Marlos C. Machado (Google Brain) · Nicolas Le Roux (Google)

Improved Regret Bound and Experience Replay in Regularized Policy Iteration
Nevena Lazic (DeepMind) · Dong Yin (DeepMind) · Yasin Abbasi-Yadkori (Adobe Research) · Csaba Szepesvari (DeepMind/University of Alberta)

Learning and Planning in Complex Action Spaces
Thomas Hubert (DeepMind) · Julian Schrittwieser (DeepMind) · Ioannis Antonoglou (Deepmind) · Mohammadamin Barekatain (DeepMind) · Simon Schmitt (DeepMind) · David Silver (Google DeepMind)

Low-Precision Reinforcement Learning: Running Soft Actor-Critic in Half Precision
Johan Björck (Cornell) · Xiangyu Chen (Cornell University) · Christopher De Sa (Cornell) · Carla Gomes (Cornell University) · Kilian Weinberger (Cornell University)

Mixed Nash Equilibria in the Adversarial Examples Game
Laurent Meunier (Facebook/Dauphine) · Meyer Scetbon (CREST, ENSAE) · Rafael Pinot (Dauphine University - CEA LIST) · Jamal Atif (Université Paris-Dauphine) · Yann Chevaleyre (Univ. Paris Dauphine)

Muesli: Combining Improvements in Policy Optimization
Matteo Hessel (DeepMind) · Ivo Danihelka (DeepMind) · Fabio Viola (DeepMind) · Arthur Guez (Google DeepMind) · Simon Schmitt (DeepMind) · Laurent Sifre (DeepMind) · Theophane Weber (DeepMind) · David Silver (Google DeepMind) · Hado van Hasselt (DeepMind)

Optimal Thompson Sampling strategies for support-aware CVaR bandits
Dorian Baudry (CNRS/INRIA) · Romain Gautron (CIRAD - CGIAR) · Emilie Kaufmann (CNRS, Univ. Lille) · Odalric-Ambrym Maillard (Inria Lille - Nord Europe)

PODS: Policy Optimization via Differentiable Simulation
Miguel Angel Zamora Mora (ETH Zurich) · Momchil Peychev (ETH Zurich) · Sehoon Ha (Georgia Institute of Technology) · Martin Vechev (ETH Zurich) · Stelian Coros (ETH Zurich)

Principled Exploration via Optimistic Bootstrapping and Backward Induction
Chenjia Bai (Harbin Institute of Technology) · Lingxiao Wang (Northwestern University) · Lei Han (Tencent AI Lab) · Jianye Hao (Tianjin University) · Animesh Garg (University of Toronto, Vector Institute, Nvidia) · Peng Liu (Harbin Institute of Technology) · Zhaoran Wang (Northwestern U)

Preferential Temporal Difference Learning
Nishanth Anand (Mila / McGill University) · Doina Precup (McGill University / DeepMind)

Provably Efficient Reinforcement Learning for Discounted MDPs with Feature Mapping
Dongruo Zhou (UCLA) · Jiafan He (University of California, Los Angeles) · Quanquan Gu (University of California, Los Angeles)

The Symmetry between Arms and Knapsacks: A Primal-Dual Approach for Bandits with Knapsacks
Xiaocheng Li (Imperial College London) · Chunlin Sun (Stanford University) · Yinyu Ye (Standord)

Reinforcement Learning for Cost-Aware Markov Decision Processes
Wesley Suttle (Stony Brook University) · Kaiqing Zhang (University of Illinois at Urbana-Champaign/MIT) · Zhuoran Yang (Princeton University) · Ji Liu (Stony Brook University) · David N Kraemer (Stony Brook University)

TeachMyAgent: a Benchmark for Automatic Curriculum Learning in Deep RL
Clément Romac (Inria) · Rémy Portelas (Inria Bordeaux - Sud-Ouest) · Katja Hofmann (Microsoft) · Pierre-Yves Oudeyer (Inria)

Towards Open Ad Hoc Teamwork Using Graph-based Policy Learning
Muhammad Arrasy Rahman (The University of Edinburgh) · Niklas Hopner (University of Amsterdam) · Filippos Christianos (University of Edinburgh) · Stefano Albrecht (University of Edinburgh)

When is Pessimism Warranted in Batch Policy Optimization?
Chenjun Xiao (Google / University of Alberta) · Yifan Wu (Carnegie Mellon University) · Jincheng Mei (University of Alberta / Google Brain) · Bo Dai (Google Brain) · Tor Lattimore (DeepMind) · Lihong Li (Google Research) · Csaba Szepesvari (DeepMind/University of Alberta) · Dale Schuurmans (Google / University of Alberta)

Theory: sample complexity

Towards Tight Bounds on the Sample Complexity of Average-reward MDPs
Yujia Jin (Stanford University) · Aaron Sidford (Stanford)

Tightening the Dependence on Horizon in the Sample Complexity of Q-Learning
Gen Li (Tsinghua University, China) · Changxiao Cai (Princeton University) · Yuxin Chen (Princeton University) · Yuantao Gu (Tsinghua University) · Yuting Wei (Carnegie Mellon University) · Yuejie Chi (CMU)

A Lower Bound for the Sample Complexity of Inverse Reinforcement Learning
Abi Komanduru (Purdue University) · Jean Honorio (Purdue University)

Model-Free Reinforcement Learning: from Clipped Pseudo-Regret to Sample Complexity
Zhang Zihan (Tsinghua University) · Yuan Zhou (UIUC) · Xiangyang Ji (Tsinghua University)

Policy Information Capacity: Information-Theoretic Measure for Task Complexity in Deep Reinforcement Learning
Hiroki Furuta (The University of Tokyo) · Tatsuya Matsushima (The University of Tokyo) · Tadashi Kozuno (University of Alberta) · Yutaka Matsuo (University of Tokyo) · Sergey Levine (UC Berkeley) · Ofir Nachum (Google Brain) · Shixiang Gu (Google)

Adaptive Sampling for Best Policy Identification in Markov Decision Processes
Aymen Al Marjani (ENS Lyon) · Alexandre Proutiere (KTH Royal Institute of Technology)

Finite-Sample Analysis of off-policy Natural Actor-Critic Algorithm
sajad khodadadian (georgia institute of technology) · Zaiwei Chen (Georgia Institute of Technology) · Siva Maguluri (Georgia Tech)

Sample Efficient Reinforcement Learning In Continuous State Spaces: A Perspective Beyond Linearity
Dhruv Malik (Carnegie Mellon University) · Aldo Pacchiano (UC Berkeley) · Vishwak Srinivasan (Carnegie Mellon University) · Yuanzhi Li (CMU)

Diversity Actor-Critic: Sample-Aware Entropy Regularization for Sample-Efficient Exploration
Seungyul Han (KAIST) · Youngchul Sung (KAIST)

Sparse Feature Selection Makes Batch Reinforcement Learning More Sample Efficient
Botao Hao (Princeton University) · Yaqi Duan (Princeton University) · Tor Lattimore (DeepMind) · Csaba Szepesvari (DeepMind/University of Alberta) · Mengdi Wang (Princeton University)

Self-play

DouZero: Mastering DouDizhu with Self-Play Deep Reinforcement Learning
Daochen Zha (Texas A&M University) · Jingru Xie (Kwai Inc.) · Wenye Ma (Kuaishou) · Sheng Zhang (Georgia Institute of Technology) · Xiangru Lian (Kwai Inc.) · Xia Hu (Texas A&M University) · Ji Liu (Kwai Seattle AI lab, University of Rochester)

potential

learning Nonzero-Sum Stochastic Games with Potentials
David Mguni (Noah's Ark Laboratory, Huawei) · Yutong Wu (Institute of Automation, Chinese Academy of Sciences) · Yali Du (University College London) · Yaodong Yang (Huawei) · Ziyi Wang (Peking University) · Minne Li (University College London) · Ying Wen (Shanghai Jiao Tong University) · Joel Jennings (Huawei) · Jun Wang (Huawei)

Application

On Reinforcement Learning with Adversarial Corruption and Its Application to Block MDP
Tianhao Wu (Peking University) · Yunchang Yang (Center for Data Science, Peking University) · Simon Du (University of Washington) · Liwei Wang (Peking University)

Controlling Graph Dynamics with Reinforcement Learning and Graph Neural Networks
Eli Meirom (NVIDIA Research) · Haggai Maron (NVIDIA Research) · Shie Mannor (Technion) · Gal Chechik (NVIDIA / Bar-Ilan University)

A Deep Reinforcement Learning Approach to Marginalized Importance Sampling with the Successor Representation
Scott Fujimoto (McGill University) · David Meger (McGill University) · Doina Precup (McGill University / DeepMind)

SCC: an efficient deep Reinforcement Learning agent mastering the game of StarCraft II
Xiangjun Wang (inspir.ai) · Junxiao SONG (inspir.ai) · Penghui Qi (InspirAI) · Peng Peng (inspir.ai) · Zhenkun Tang (inspir.ai) · Wei Zhang (inspir.ai) · Weimin Li (inspir.ai) · Xiongjun Pi (inspir.ai) · Jujie He (inspir.ai) · Chao Gao (inspir.ai) · Haitao Long (inspir.ai) · Quan Yuan (inspir.ai)

Online Algorithm Selection: A Rested Bandit Formulation
Leonardo Cella (Italian Institute of Technology) · Claudio Gentile (Google Research) · Massimiliano Pontil (Istituto Italiano di Tecnologia and University College London)

Efficient Generative Modelling of Protein Structure Fragments using a Deep Markov Model
Christian Thygesen (University of Copenhagen) · Christian Skjødt Steenmans (Evaxion Biotech) · Ahmad Salim Al-Sibahi (University of Copenhagen) · Lys Sanz Moreta (University of Copenhagen) · Anders Bundgård Sørensen (Evaxion Biotech) · Thomas Hamelryck (University of Copenhagen)

Exploration

Fast active learning for pure exploration in Reinforcement Learning
Pierre MENARD (Inria) · Omar Darwiche Domingues (Inria) · Anders Jonsson (Universitat Pompeu Fabra) · Emilie Kaufmann (CNRS, Univ. Lille) · Edouard Leurent () · Michal Valko (DeepMind / Inria / ENS Paris-Saclay)

Locally Persistent Exploration in Continuous Control Tasks with Sparse Rewards
Susan Amin (McGill University) · Maziar Gomrokchi (McGill University) · Hossein Aboutalebi (University of Waterloo) · Harsh Satija (McGill University) · Doina Precup (McGill University / DeepMind)

Offline/batch RL

Offline Reinforcement Learning with Fisher Divergence Critic Regularization
Ilya Kostrikov (Google/New York University) · Rob Fergus (DeepMind) · Jonathan Tompson (Google Brain) · Ofir Nachum (Google Brain)

Uncertainty Weighted Actor-Critic for Offline Reinforcement Learning
Yue Wu (Carnegie Mellon University) · Shuangfei Zhai (Apple) · Nitish Srivastava (Apple) · Joshua Susskind (Apple, Inc.) · Jian Zhang (Apple Inc.) · Ruslan Salakhutdinov (Carnegie Mellen University) · Hanlin Goh (Apple)

Batch value-function Approximation with Only Realizability
Tengyang Xie (University of Illinois at Urbana-Champaign) · Nan Jiang (University of Illinois at Urbana-Champaign)

Exponential Lower Bounds for Batch Reinforcement Learning: Batch RL can be Exponentially Harder than Online RL
Andrea Zanette (Stanford University)

Risk Bounds and Rademacher Complexity in Batch Reinforcement Learning
Yaqi Duan (Princeton University) · Chi Jin (Princeton University) · Zhiyuan Li (Princeton University)

OptiDICE: Offline Policy Optimization via Stationary Distribution Correction Estimation
Jongmin Lee (KAIST) · Wonseok Jeon (MILA, McGill University) · Byung-Jun Lee (Gauss Labs Inc.) · Joelle Pineau (McGill University / Facebook) · Kee-Eung Kim (KAIST)

How Sensitive is Offline RL to Distribution Shift?
Ruosong Wang (Carnegie Mellon University) · Yifan Wu (Carnegie Mellon University) · Sham Kakade (University of Washington) · Ruslan Salakhutdinov (Carnegie Mellen University)

Actionable Models: Unsupervised Offline Reinforcement Learning of Robotic Skills
Yevgen Chebotar (Google) · Karol Hausman (Google Brain) · Yao Lu (Google Research) · Ted Xiao (Google) · Dmitry Kalashnikov (Google Inc.) · Jacob Varley (Google) · Alexander Irpan (Google) · Benjamin Eysenbach (CMU, Google Brain) · Ryan Julian (Google) · Chelsea Finn (Google Brain) · Sergey Levine (Google)

Is Pessimism Provably Efficient for Offline RL?
Ying Jin (Stanford University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U)

Meta learning/transfer/multi-task/generalization

Offline Meta-Reinforcement Learning with Advantage Weighting
Eric Mitchell (Stanford) · Rafael Rafailov (Stanford University) · Xue Bin Peng (UC Berkeley) · Sergey Levine (University of California, Berkeley) · Chelsea Finn (Stanford)

MetaCURE: Meta Reinforcement Learning with Empowerment-Driven Exploration
Jin Zhang (Tsinghua University) · Jianhao Wang (Tsinghua University) · Hao Hu (Tsinghua University) · Tong Chen (Tsinghua University) · Yingfeng Chen (NetEase Fuxi AI Lab) · Changjie Fan (NetEase Fuxi AI Lab) · Chongjie Zhang (Tsinghua University)

Decoupling Exploration and Exploitation for Meta-Reinforcement Learning without Sacrifices
Evan Liu (Stanford University) · Aditi Raghunathan (Stanford) · Percy Liang (Stanford University) · Chelsea Finn (Stanford)

How Important is the Train-Validation Split in Meta-Learning?
Yu Bai (Salesforce Research) · Minshuo Chen (Georgia Tech) · Pan Zhou (Salesforce) · Tuo Zhao (Georgia Tech) · Jason Lee (Princeton) · Sham Kakade (University of Washington) · Huan Wang (Salesforce Research) · Caiming Xiong (Salesforce)

MURAL: Meta-Learning Uncertainty-Aware Rewards for Outcome-Driven Reinforcement Learning
Kevin Li (UC Berkeley) · Abhishek Gupta (UC Berkeley) · Ashwin D Reddy (UC Berkeley) · Vitchyr Pong (UC Berkeley) · Aurick Zhou (UC Berkeley) · Justin Yu (Berkeley) · Sergey Levine (UC Berkeley)

Improving Generalization in Meta-learning via Task Augmentation
Huaxiu Yao (Stanford University) · Long-Kai Huang (Tencent AI Lab) · Linjun Zhang (Rutgers University) · Ying WEI (City University of Hong Kong) · Li Tian (Tencent) · James Zou (Stanford University) · Junzhou Huang (University of Texas at Arlington / Tencent AI Lab) · Zhenhui (Jessie) Li (Penn State University)

Bayesian Online Meta-Learning
Pauching Yap (University College London) · Hippolyt Ritter (University College London) · David Barber (University College London)

REPAINT: Knowledge Transfer in Deep Reinforcement Learning
Yunzhe Tao (ByteDance) · Sahika Genc (Amazon AI) · Jonathan Chung (AWS) · TAO SUN (Amazon.com) · Sunil Mallya (Amazon AWS)

A Distribution-dependent Analysis of Meta Learning
Mikhail Konobeev (University of Alberta) · Ilja Kuzborskij (University of Milan) · Csaba Szepesvari (DeepMind/University of Alberta)

MetaCURE: Meta Reinforcement Learning with Empowerment-Driven Exploration
Jin Zhang (Tsinghua University) · Jianhao Wang (Tsinghua University) · Hao Hu (Tsinghua University) · Tong Chen (Tsinghua University) · Yingfeng Chen (NetEase Fuxi AI Lab) · Changjie Fan (NetEase Fuxi AI Lab) · Chongjie Zhang (Tsinghua University)

Memory Efficient Online Meta Learning
Durmus Alp Emre Acar (Boston University) · Ruizhao Zhu (Boston University) · Venkatesh Saligrama (Boston University)

Reinforcement Learning of Implicit and Explicit Control Flow Instructions
Ethan Brooks (University of Michigan) · Janarthanan Rajendran (University of Michigan) · Richard Lewis (University of Michigan) · Satinder Singh (University of Michigan)

LTL2Action: Generalizing LTL Instructions for Multi-Task RL
Pashootan Vaezipoor (University of Toronto) · Andrew Li (University of Toronto and Vector Institute) · Rodrigo A Toro Icarte (University of Toronto) · Sheila McIlraith (University of Toronto)

Decoupling Value and Policy for Generalization in Reinforcement Learning
Roberta Raileanu (NYU) · Rob Fergus (Facebook / NYU)

Generalizable Episodic Memory for Deep Reinforcement Learning
Hao Hu (Tsinghua University) · Jianing Ye (Peking University) · Guangxiang Zhu (Tsinghua University) · Zhizhou Ren (University of Illinois at Urbana-Champaign) · Chongjie Zhang (Tsinghua University)

Grounding Language to Entities and Dynamics for Generalization in Reinforcement Learning
Austin W. Hanjie (Princeton University) · Victor Zhong (University of Washington) · Karthik Narasimhan (Princeton)

Bridging Multi-Task Learning and Meta-Learning: Towards Efficient Training and Effective Adaptation
Haoxiang Wang (University of Illinois at Urbana-Champaign) · Han Zhao (University of Illinois at Urbana-Champaign) · Bo Li (UIUC)

Exploration in Approximate Hyper-State Space for Meta Reinforcement Learning
Luisa Zintgraf (University of Oxford) · Leo Feng (Mila) · Cong Lu (University of Oxford) · Maximilian Igl (University of Oxford) · Kristian Hartikainen (UC Berkeley) · Katja Hofmann (Microsoft) · Shimon Whiteson (University of Oxford)

Goal-Conditioned Reinforcement Learning with Imagined Subgoals
Elliot Chane-Sane (INRIA Paris) · Cordelia Schmid (Inria/Google) · Ivan Laptev (INRIA Paris)

What Structural Conditions Permit Generalization in Reinforcement Learning?
Simon Du (University of Washington) · Sham Kakade (University of Washington) · Jason Lee (Princeton) · Shachar Lovett (University of California San Diego) · Gaurav Mahajan (UCSD) · Wen Sun (Cornell University) · Ruosong Wang (Carnegie Mellon University)

Representation

PEBBLE: Feedback-Efficient Interactive Reinforcement Learning via Relabeling Experience and Unsupervised Pre-training
Kimin Lee (UC Berkeley) · Laura Smith (UC Berkeley) · Pieter Abbeel (UC Berkeley & Covariant)

Towards Better Laplacian Representation in Reinforcement Learning with Generalized Graph Drawing
Kaixin Wang (National University of Singapore) · Kuangqi Zhou (National University of Singapore) · Qixin Zhang (city university of hong kong) · Jie Shao (Fudan University) · Bryan Hooi (National University of Singapore) · Jiashi Feng (National University of Singapore)

Learning Representations by Humans, for Humans
Anna Hilgard (Harvard University) · Nir Rosenfeld (Harvard) · Mahzarin Banaji (Harvard University) · Jack Cao (Harvard) · David Parkes (Harvard University)

Scaling Up Visual and Vision-Language Representation Learning With Noisy Text Supervision
Chao Jia (Google) · Yinfei Yang (Google Research) · Ye Xia (Google) · Yi-Ting Chen (Google) · Zarana Parekh (Google) · Hieu Pham (Google) · Zhen Li (Google) · Tom Duerig (Google) · Yun-Hsuan Sung (Google Research) · Quoc Le (Google Brain)

Self-supervised Graph-level Representation Learning with Local and Global Structure
Minghao Xu (Shanghai Jiao Tong University) · Hang Wang (Shanghai Jiao Tong University) · Bingbing Ni (Shanghai Jiao Tong University) · Hongyu Guo (National Research Council Canada) · Jian Tang (HEC Montreal & MILA)

Reinforcement Learning with Prototypical Representations
Denis Yarats (New York University) · Rob Fergus (Facebook / NYU) · Alessandro Lazaric (Facebook AI Research) · Lerrel Pinto (NYU/Berkeley)

UniSpeech: Unified Speech Representation Learning with Labeled and Unlabeled Data
Chengyi Wang (Nankai University) · Yu Wu (Microsoft Research) · Yao Qian (Microsoft) · Kenichi Kumatani (Microsoft) · Shujie Liu (Microsoft Research Asia) · Furu Wei (Microsoft Research Asia) · Michael Zeng (Microsoft) · Xuedong Huang (Microsoft)

Causal Curiosity: RL Agents Discovering Self-supervised Experiments for Causal Representation Learning
Sumedh Sontakke (University of Southern California) · Arash Mehrjou (Max Planck Institute for Intelligent Systems) · Laurent Itti (University of Southern California) · Bernhard Schölkopf (MPI for Intelligent Systems Tübingen, Germany)

Multi-Task Reinforcement Learning with Context-based Representations
Shagun Sodhani (Facebook AI Research) · Amy Zhang (FAIR / UC Berkeley) · Joelle Pineau (McGill, Facebook)

RRL: Resnet as representation for Reinforcement Learning
Rutav M Shah (Indian Institute of Technology, Kharagpur) · Vikash Kumar (Univ. Of Washington)

Near-Optimal Representation Learning for Linear Bandits and Linear RL
Xiaoyu Chen (Peking University) · Jiachen Hu (Peking University) · Chi Jin (Princeton University) · Lihong Li (Google Research) · Liwei Wang (Peking University)

Decoupling Representation Learning from Reinforcement Learning
Adam Stooke (UC Berkeley) · Kimin Lee (UC Berkeley) · Pieter Abbeel (UC Berkeley & Covariant) · Michael Laskin (UC Berkeley)

Learning node representations using stationary flow prediction on large payment and cash transaction networks
Ciwan Ceylan (KTH Royal Institute of Technology & SEB) · Salla Franzén (SEB AB) · Florian T. Pokorny (KTH Royal Institute of Technology)

Variational Empowerment as Representation Learning for Goal-Conditioned Reinforcement Learning
Jongwook Choi (University of Michigan) · Archit Sharma () · Honglak Lee (Google / U. Michigan) · Sergey Levine (Google) · Shixiang Gu (Google)

Contrastive Learning

Towards Domain-Agnostic Contrastive Learning
Vikas Verma (Aalto University) · Thang Luong (Google Brain) · Kenji Kawaguchi (Harvard University) · Hieu Pham (Google) · Quoc Le (Google Brain)

Large-Margin Contrastive Learning with Distance Polarization Regularizer
Shuo Chen (RIKEN) · Gang Niu (RIKEN) · Chen Gong (Nanjing University of Science and Technology) · Jun Li (Nanjing University of Science and Technology) · Jian Yang (Nanjing University of Science and Technology) · Masashi Sugiyama (RIKEN / The University of Tokyo)

Hierarchical RL/option/skill

TempoRL: Learning When to Act
André Biedenkapp (University of Freiburg) · Raghu Rajan (University of Freiburg) · Frank Hutter (University of Freiburg and Bosch Center for Artificial Intelligence) · Marius Lindauer (Leibniz University Hannover)

Near-Optimal Representation Learning for Linear Bandits and Linear RL
Xiaoyu Chen (Peking University) · Jiachen Hu (Peking University) · Chi Jin (Princeton University) · Lihong Li (Google Research) · Liwei Wang (Peking University)

Imitation Learning/reward-free/goal/inverse

Cross-domain Imitation from Observation
Dripta S. Raychaudhuri (University of California, Riverside) · Sujoy Paul (Google Research) · Jeroen Vanbaar (MERL) · Amit K. Roy-Chowdhury (University of California, Riverside)

Inverse Constrained Reinforcement Learning
Shehryar Malik (Information Technology University) · Usman Anwar (Information Technlogy University, Lahore.) · Alireza Aghasi (Georgia State University) · Ali Ahmed (Information Technology University)

Adversarial Option-Aware Hierarchical Imitation Learning
Mingxuan Jing (Tsinghua University) · Wenbing Huang (Tsinghua University) · Fuchun Sun (Tsinghua) · Xiaojian Ma (University of California, Los Angeles) · Tao Kong (Bytedance) · Chuang Gan (MIT-IBM Watson AI Lab) · Lei Li (ByteDance AI Lab)

Near Optimal Reward-Free Reinforcement Learning
Zhang Zihan (Tsinghua University) · Simon Du (University of Washington) · Xiangyang Ji (Tsinghua University)

Keyframe-Focused Visual Imitation Learning
Chuan Wen (Tsinghua University) · Jierui Lin (UT Austin) · Jianing Qian (University of Pennsylvania) · Yang Gao (Tsinghua University) · Dinesh Jayaraman (University of Pennsylvania)

LogME: Practical Assessment of Pre-trained Models for Transfer Learning
Kaichao You (Tsinghua University) · Yong Liu (Tsinghua University) · Jianmin Wang (Tsinghua University) · Mingsheng Long (Tsinghua University)

Demonstration-Conditioned Reinforcement Learning for Few-Shot Imitation
Christopher Dance (NAVER LABS Europe) · Perez Julien (Naver Labs Europe) · Théo Cachet (Naver Labs Europe)

Inverse Decision Modeling: Learning Interpretable Representations of Behavior
Daniel Jarrett (University of Cambridge) · Alihan Hüyük (University of Cambridge) · Mihaela van der Schaar (University of Cambridge and UCLA)

Policy Gradient Bayesian Robust Optimization for Imitation Learning
Daniel Brown (University of Texas at Austin) · Ashwin Balakrishna (University of California, Berkeley) · Zaynah Javed (UC Berkeley) · Satvik Sharma (UC Berkeley) · Jerry Zhu (UC Berkeley) · Marek Petrik (University of New Hampshire) · Anca Dragan (University of California, Berkeley) · Ken Goldberg (UC Berkeley)

On Reward-Free RL with Kernel and Neural Function Approximations: Single-Agent MDP and Markov Game
Shuang Qiu (University of Michigan) · Zhuoran Yang (Princeton University) · Jieping Ye (University of Michigan) · Zhaoran Wang (Northwestern U)

Reward Identification in Inverse Reinforcement Learning
Kuno Kim (Stanford University) · Shivam Garg (Stanford University) · Kirankumar Shiragur (Stanford University) · Stefano Ermon (Stanford University)

PsiPhi-Learning: Reinforcement Learning with Demonstrations using Successor Features and Inverse Temporal Difference Learning
Angelos Filos (University of Oxford) · Clare Lyle (University of Oxford) · Yarin Gal (University of Oxford) · Sergey Levine (UC Berkeley) · Natasha Jaques (Google Brain, UC Berkeley) · Gregory Farquhar (University of Oxford)

Model-based

A Sharp Analysis of Model-based Reinforcement Learning with Self-Play
Qinghua Liu (Princeton University) · Tiancheng Yu (MIT) · Yu Bai (Salesforce Research) · Chi Jin (Princeton University)

Model-based Reinforcement Learning for Continuous Control with Posterior Sampling
Ying Fan (University of Wisconsin-Madison) · Yifei Ming (University of Wisconsin-Madison)

Model-Based Reinforcement Learning via Latent-Space Collocation
Oleh Rybkin (University of Pennsylvania) · Chuning Zhu (University of Pennsylvania) · Anusha Nagabandi (UC Berkeley) · Kostas Daniilidis (University of Pennsylvania) · Igor Mordatch (Google Brain) · SergeyLevine (UC Berkeley)

Dynamic Balancing for Model Selection in Bandits and RL
Ashok Cutkosky (Boston University) · Christoph Dann (Google) · Abhimanyu Das (Google) · Claudio Gentile (Google Research) · Aldo Pacchiano (UC Berkeley) · Manish Purohit (Google Research)

PC-MLP: Model-based Reinforcement Learning with Policy Cover Guided Exploration
Yuda Song (University of California, San Diego) · Wen Sun (Cornell University)

Conservative Objective Models for Effective Offline Model-Based Optimization
Brandon L Trabucco (UC Berkeley) · Aviral Kumar (UC Berkeley) · Xinyang Geng (UC Berkeley) · Sergey Levine (UC Berkeley)

Temporal Predictive Coding For Model-Based Planning In Latent Space
Tung Nguyen (VinAI Research, Vietnam) · Rui Shu (Stanford University) · Tuan Pham (VinAI Research) · Hung Bui (VinAI Research) · Stefano Ermon (Stanford University)

Automatic RNN Repair via Model-based Analysis
Xiaofei Xie (Nanyang Technological University) · Wenbo Guo (Pennsylvania State University) · Lei Ma (University of Alberta) · Wei Le (Iowa State University) · Jian Wang (Nanyang Technological University) · Lingjun Zhou (College of Intelligence and Computing,Tianjin University) · Yang Liu (Nanyang Technology University, Singapore) · Xinyu Xing (The Pennsylvania State University)

Continuous-time Model-based Reinforcement Learning
Cagatay Yildiz (Aalto University) · Markus Heinonen (Aalto University) · Harri Lähdesmäki (Aalto University)

Multi-agent/game theory

Cooperative Exploration for Multi-Agent Deep Reinforcement Learning
Iou-Jen Liu (University of Illinois at Urbana-Champaign) · Unnat Jain (UIUC) · Raymond Yeh (University of Illinois at Urbana–Champaign) · Alexander Schwing (UIUC)

Global Convergence of Policy Gradient for Linear-Quadratic Mean-Field Control/Game in Continuous Time
Weichen Wang (Two Sigma Investments, LP) · Jiequn Han (Princeton University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern)

Tesseract: Tensorised Actors for Multi-Agent Reinforcement Learning
Anuj Mahajan (Dept. of Computer Science, University of Oxford) · Mikayel Samvelyan (University College London) · Lei Mao (NVIDIA) · Viktor Makoviychuk (NVIDIA) · Animesh Garg (University of Toronto, Vector Institute, Nvidia) · Jean Kossaifi (NVIDIA) · Shimon Whiteson (University of Oxford) · Yuke Zhu (University of Texas - Austin) · Anima Anandkumar (Caltech and NVIDIA)

Randomized Entity-wise Factorization for Multi-Agent Reinforcement Learning
Shariq Iqbal (University of Southern California) · Christian Schroeder (University of Oxford) · Bei Peng (University of Oxford) · Wendelin Boehmer (Delft University of Technology) · Shimon Whiteson (University of Oxford) · Fei Sha (Google Research)

UneVEn: Universal Value Exploration for Multi-Agent Reinforcement Learning
Tarun Gupta (University of Oxford) · Anuj Mahajan (Dept. of Computer Science, University of Oxford) · Bei Peng (University of Oxford) · Wendelin Boehmer (Delft University of Technology) · Shimon Whiteson (University of Oxford)

DFAC Framework: Factorizing the value function via Quantile Mixture for Multi-Agent Distributional Q-Learning
Wei-Fang Sun (National Tsing Hua University) · Cheng-Kuang Lee (NVIDIA Corporation) · Chun-Yi Lee (National Tsing Hua University)

A Policy Gradient Algorithm for Learning to Learn in Multiagent Reinforcement Learning
Dong Ki Kim (MIT) · Miao Liu (IBM) · Matthew Riemer (IBM Research) · Chuangchuang Sun (MIT) · Marwa Abdulhai (MIT) · Golnaz Habibi (MIT) · Sebastian Lopez-Cot (MIT) · Gerald Tesauro (IBM Research) · Jonathan How (MIT)

Learning Interaction Kernels for Agent Systems on Riemannian Manifolds
Mauro Maggioni (Johns Hopkins University) · Jason Miller (Johns Hopkins University) · Hongda Qiu (Johns Hopkins University) · Ming Zhong (Johns Hopkins University)

Scalable Evaluation of Multi-Agent Reinforcement Learning with Melting Pot
Joel Z Leibo (DeepMind) · Edgar Duenez-Guzman (DeepMind) · Alexander Vezhnevets (DeepMind) · John Agapiou (DeepMind) · Peter Sunehag () · Raphael Koster (DeepMind) · Jayd Matyas (DeepMind) · Charles Beattie (DeepMind Technologies Limited) · Igor Mordatch (Google Brain) · Thore Graepel (DeepMind)

Multi-Agent Training beyond Zero-Sum with Correlated Equilibrium Meta-Solvers
Luke Marris (DeepMind) · Paul Muller (DeepMind) · Marc Lanctot (DeepMind) · Karl Tuyls (DeepMind) · Thore Graepel (DeepMind)

Coach-Player Multi-agent Reinforcement Learning for Dynamic Team Composition
Bo Liu (University of Texas, Austin) · Qiang Liu (UT Austin) · Peter Stone (University of Texas at Austin) · Animesh Garg (University of Toronto, Vector Institute, Nvidia) · Yuke Zhu (University of Texas - Austin) · Anima Anandkumar (California Institute of Technology)

Learning While Playing in Mean-Field Games: Convergence and Optimality
Qiaomin Xie (Cornell University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U) · Andreea Minca (Cornell University)

Learning Fair Policies in Decentralized Cooperative Multi-Agent Reinforcement Learning
Matthieu Zimmer (Shanghai Jiao Tong University) · Claire Glanois (Shanghai Jiao Tong University) · Umer Siddique (Shanghai Jiao Tong University) · Paul Weng (Shanghai Jiao Tong University)

FOP: Factorizing Optimal Joint Policy of Maximum-Entropy Multi-Agent Reinforcement Learning
Tianhao Zhang (Peking University) · yueheng li (Peking university) · Chen Wang (Peking University) · Zongqing Lu (Peking University) · Guangming Xie (1. State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University; 2. Institute of Ocean Research, Peking University)

Provably Efficient Fictitious Play Policy Optimization for Zero-Sum Markov Games with Structured Transitions
Shuang Qiu (University of Michigan) · Zhuoran Yang (Princeton University) · Xiaohan Wei (Facebook) · Jieping Ye (University of Michigan) · Zhaoran Wang (Northwestern U)

Scaling Multi-Agent Reinforcement Learning with Selective Parameter Sharing
Filippos Christianos (University of Edinburgh) · Georgios Papoudakis (The University of Edinburgh) · Muhammad Arrasy Rahman (The University of Edinburgh) · Stefano Albrecht (University of Edinburgh)

The Emergence of Individuality in Multi-Agent Reinforcement Learning
Jiechuan Jiang (Peking University) · Zongqing Lu (Peking University)

Emergent Social Learning via Multi-agent Reinforcement Learning
Kamal Ndousse (OpenAI) · Douglas Eck (Google Brain) · Sergey Levine (UC Berkeley) · Natasha Jaques (Google Brain, UC Berkeley)

Parallel Droplet Control in MEDA Biochips using Multi-Agent Reinforcement Learning
Tung-Che Liang (Duke University) · Jin Zhou (Duke University) · Yun-Sheng Chan (National Chiao Tung University) · Tsung-Yi Ho (National Tsing Hua University) · Krishnendu Chakrabarty (Duke University) · Cy Lee (National Chiao Tung University)

Robust/safe/constrained/uncertainty

Safe Reinforcement Learning with Linear Function Approximation
Sanae Amani (University of California, Los Angeles) · Christos Thrampoulidis (University of British Columbia) · Lin Yang (UCLA)

State Entropy Maximization with Random Encoders for Efficient Exploration
Younggyo Seo (KAIST) · Lili Chen (UC Berkeley) · Jinwoo Shin (KAIST) · Honglak Lee (Google / U. Michigan) · Pieter Abbeel (UC Berkeley & Covariant) · Kimin Lee (UC Berkeley)

High Confidence Generalization for Reinforcement Learning
James Kostas (University of Massachusetts Amherst) · Yash Chandak (University of Massachusetts Amherst) · Scott M Jordan (University of Massachusetts) · Georgios Theocharous (Adobe Research) · Philip Thomas (University of Massachusetts Amherst)

Robust Asymmetric Learning in POMDPs
Andrew Warrington (University of Oxford) · Jonathan Lavington (University of British Columbia) · Adam Scibior (University of British Columbia) · Mark Schmidt (University of British Columbia) · Frank Wood (University of British Columbia)

Risk-Sensitive Reinforcement Learning with Function Approximation: A Debiasing Approach
Yingjie Fei (Cornell University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U)

Model-Free and Model-Based Policy Evaluation when Causality is Uncertain
David Bruns-Smith (UC Berkeley)

Improved Corruption Robust Algorithms for Episodic Reinforcement Learning
Yifang Chen (University of Washington) · Simon Du (University of Washington) · Kevin Jamieson (University of Washington)

CRPO: A New Approach for Safe Reinforcement Learning with Convergence Guarantee
Tengyu Xu (The Ohio State University) · Yingbin LIANG (The Ohio State University) · Guanghui Lan (Georgia Institute of Technology)

SAINT-ACC: Safety-Aware Intelligent Adaptive Cruise Control for Autonomous Vehicles Using Deep Reinforcement Learning
Lokesh Chandra Das (The University of Memphis) · Myounggyu Won (University of Memphis)

Robust Reinforcement Learning using Least Squares Policy Iteration with Provable Performance Guarantees
Kishan Panaganti (TAMU) · Dileep Kalathil (TAMU)

Combining Pessimism with Optimism for Robust and Efficient Model-Based Deep Reinforcement Learning
Sebastian Curi (ETH) · Ilija Bogunovic (ETH Zurich) · Andreas Krause (ETH Zurich)

Monotonic Robust Policy Optimization with Model Discrepancy
yuankun jiang (Shanghai Jiao Tong University) · Chenglin Li (Shanghai Jiao Tong University) · Wenrui Dai (Shanghai Jiao Tong University) · Junni Zou (Shanghai Jiao Tong University) · Hongkai Xiong (Shanghai Jiao Tong University)

Distributionally Robust Optimization with Markovian Data
Mengmeng Li (EPFL) · Tobias Sutter (EPFL Lausanne) · Daniel Kuhn (EPFL)

Connecting Interpretability and Robustness in Decision Trees through Separation
Michal Moshkovitz (UC San Diego) · Kamalika Chaudhuri (University of California at San Diego) · Yao-Yuan Yang (UCSD)

1-bit Adam: Communication Efficient Large-Scale Training with Adam's Convergence Speed
Hanlin Tang (University of Rochester) · Shaoduo Gan (ETH Zurich) · Ammar Ahmad Awan (Microsoft) · Samyam Rajbhandari (Microsoft) · Conglong Li (Microsoft) · Xiangru Lian (University of Rochester) · Ji Liu (Kwai Seattle AI lab, University of Rochester) · Ce Zhang (ETH Zurich) · Yuxiong He (Microsoft)

Robust Policy Gradient against Strong Data Corruption
Xuezhou Zhang (UW-Madison) · Yiding Chen (University of Wisconsin-Madison) · Jerry Zhu (University of Wisconsin-Madison) · Wen Sun (Cornell University)

Reinforcement Learning Under Moral Uncertainty
Adrien Ecoffet (OpenAI) · Joel Lehman ()

Doubly Robust off-policy Actor-Critic: Convergence and Optimality
Tengyu Xu (The Ohio State University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U) · Yingbin LIANG (The Ohio State University)

Policy Gradient Bayesian Robust Optimization for Imitation Learning
Daniel Brown (University of Texas at Austin) · Ashwin Balakrishna (University of California, Berkeley) · Zaynah Javed (UC Berkeley) · Satvik Sharma (UC Berkeley) · Jerry Zhu (UC Berkeley) · Marek Petrik (University of New Hampshire) · Anca Dragan (University of California, Berkeley) · Ken Goldberg (UC Berkeley)

Accelerating Safe Reinforcement Learning with Constraint-mismatched Baseline Policies
Jimmy (Tsung-Yen) Yang (Princeton University) · Justinian Rosca (Siemens Corp.) · Karthik Narasimhan (Princeton) · Peter Ramadge (Princeton)

First-Order Methods for Wasserstein Distributionally Robust MDP
Julien Grand-Clement (IEOR Department, Columbia University) · Christian Kroer (Columbia University)

Efficient Performance Bounds for Primal-Dual Reinforcement Learning from Demonstrations
Angeliki Kamoutsi (ETH Zurich) · Goran Banjac (ETH Zurich) · John Lygeros (ETH Zürich)

Safe Reinforcement Learning Using Advantage-Based Intervention
Nolan Wagener (Georgia Tech) · Ching-An Cheng (Microsoft Research) · Byron Boots (University of Washington)

model-free

Is Model-Free Learning Nearly Optimal for Non-Stationary RL?
Weichao Mao (University of Illinois at Urbana-Champaign) · Kaiqing Zhang (University of Illinois at Urbana-Champaign/MIT) · Ruihao Zhu (MIT) · David Simchi-Levi (MIT) · Tamer Basar (University of Illinois at Urbana-Champaign)

Counterfactual Credit Assignment in Model-Free Reinforcement Learning
Thomas Mesnard (DeepMind) · Theophane Weber (DeepMind) · Fabio Viola (DeepMind) · Shantanu Thakoor (DeepMind) · Alaa Saade (DeepMind) · Anna Harutyunyan (DeepMind) · Will Dabney (DeepMind) · Thomas Stepleton (DeepMind) · Nicolas Heess (DeepMind) · Arthur Guez (Google DeepMind) · Eric Moulines (Ecole Polytechnique) · Marcus Hutter (DeepMind) · Lars Buesing (Deepmind) · Remi Munos (DeepMind)

off-policy

Average-Reward off-policy Policy Evaluation with Function Approximation
Shangtong Zhang (University of Oxford) · Yi Wan (University of Alberta) · Richard Sutton (DeepMind / Univ Alberta) · Shimon Whiteson (University of Oxford)

Optimal off-policy Evaluation from Multiple Logging Policies
Nathan Kallus (Cornell University) · Yuta Saito (Tokyo Institute of Technology.) · Masatoshi Uehara (Cornell University)

Finite-Sample Analysis of off-policy Natural Actor-Critic Algorithm
sajad khodadadian (georgia institute of technology) · Zaiwei Chen (Georgia Institute of Technology) · Siva Maguluri (Georgia Tech)

off-policy Confidence Sequences
Nikos Karampatziakis (Microsoft) · Paul Mineiro (Microsoft) · Aaditya Ramdas (Carnegie Mellon University)

Deeply-Debiased off-policy Interval Estimation
Chengchun Shi (London School of Economics and Political Science) · Runzhe Wan (North Carolina State University) · Victor Chernozhukov (MIT) · Rui Song (North Carolina State University)

Learning Routines for Effective off-policy Reinforcement Learning
Edoardo Cetin (King's College London) · Oya Celiktutan (King's College London)

Doubly Robust off-policy Actor-Critic: Convergence and Optimality
Tengyu Xu (The Ohio State University) · Zhuoran Yang (Princeton University) · Zhaoran Wang (Northwestern U) · Yingbin LIANG (The Ohio State University)

State Relevance for off-policy Evaluation
Simon Shen (Harvard University) · Yecheng Ma (University of Pennsylvania) · Omer Gottesman (Harvard University) · Finale Doshi-Velez (Harvard University)

Bootstrapping Fitted Q-Evaluation for off-policy Inference
Botao Hao (Princeton University) · Xiang Ji (Princeton University) · Yaqi Duan (Princeton University) · Hao Lu (Princeton University) · Csaba Szepesvari (DeepMind/University of Alberta) · Mengdi Wang (Princeton University)

Data-efficient Hindsight off-policy Option Learning
Markus Wulfmeier (DeepMind) · Dushyant Rao (DeepMind) · Roland Hafner (DeepMind) · Thomas Lampe (DeepMind) · Abbas Abdolmaleki (DeepMind) · Tim Hertweck (DeepMind) · Michael Neunert (Google DeepMind) · Dhruva Tirumala Bukkapatnam (DeepMind) · Noah Siegel (DeepMind) · Nicolas Heess (DeepMind) · Martin Riedmiller (DeepMind)

on-policy

on-policy Reinforcement Learning for the Average-Reward Criterion
Yiming Zhang (New York University) · Keith Ross (New York University Shanghai)

distributional rl

GMAC: A Distributional Perspective on Actor-Critic Framework
Daniel Nam (KC Machine Learning Lab) · Younghoon Kim (KC-ML2) · Chan Youn Park (KC ML2)

OptiDICE: Offline Policy Optimization via Stationary Distribution Correction Estimation
Jongmin Lee (KAIST) · Wonseok Jeon (MILA, McGill University) · Byung-Jun Lee (Gauss Labs Inc.) · Joelle Pineau (McGill University / Facebook) · Kee-Eung Kim (KAIST)