圆桌论道 | Neurips 2023强化学习相关350篇论文分类整理

作者：RLCN公众号　　　　　　　　二维码：

Fundamental RL

Mastering Atari Games with Generalized Weighted Path Consistency

Dengwei Zhao · Shikui Tu · Lei Xu

Learning Better with Less: Effective Augmentation for Sample-Efficient Visual Reinforcement Learning

Guozheng Ma · Linrui Zhang · Haoyu Wang · Lu Li · Zilin Wang · Zhen Wang · Li Shen · Xueqian Wang · Dacheng Tao

Boosting Verification of Deep Reinforcement Learning via Piece-Wise Linear Decision Neural Networks

Jiaxu Tian · Dapeng Zhi · Si Liu · Peixin Wang · Cheng Chen · Min Zhang

Improved Communication Efficiency in Federated Natural Policy Gradient via ADMM-based Gradient Updates

Guangchen Lan · Han Wang · James Anderson · Christopher Brinton · Vaneet Aggarwal

Policy Optimization for Continuous Reinforcement Learning

HANYANG ZHAO · Wenpin Tang · David Yao

Probabilistic Inference in Reinforcement Learning Done Right

Jean Tarbouriech · Tor Lattimore · Brendan O'Donoghue

On the Convergence and Sample Complexity Analysis of Deep Q-Networks with ϵ-Greedy Exploration

Shuai Zhang · Meng Wang · Hongkang Li · Miao Liu · Pin-Yu Chen · Songtao Lu · Sijia Liu · Keerthiram Murugesan · Subhajit Chaudhury

Reinforcement Learning with Fast and Forgetful Memory

Steven D Morad · Ryan Kortvelesy · Stephan Liwicki · Amanda Prorok

Rewarded soups: towards Pareto-optimality by interpolating weights fine-tuned on diverse rewards

Alexandre Rame · Guillaume Couairon · Corentin Dancette · Jean-Baptiste Gaya · Mustafa Shukor · Laure Soulier · Matthieu

Kerenlized Reinforcement Learning with Order Optimal Regret Bounds

Sattar Vakili · Julia Olkhovskaya

The Benefits of Being Distributional: Small-Loss Bounds for Reinforcement Learning

Kaiwen Wang · Kevin Zhou · Runzhe Wu · Nathan Kallus · Wen Sun

Optimal Convergence Rate for Exact Policy Mirror Descent in Discounted Markov Decision Processes

Emmeran Johnson · Ciara Pike-Burke · Patrick Rebeschini

Provably Efficient Algorithm for Nonstationary Low-Rank MDPs

Yuan Cheng · Jing Yang · Yingbin Liang

Finite-Time Analysis of Single-Timescale Actor-Critic

Xuyang Chen · Lin Zhao

General Munchausen Reinforcement Learning with Tsallis Kullback-Leibler Divergence

Lingwei Zhu · Zheng Chen · Matthew Schlegel · Martha White

Synthetic Experience Replay

Cong Lu · Philip Ball · Yee Whye Teh · Jack Parker-Holder

Dynamic Regret of Adversarial Linear Mixture MDPs

Long-Fei Li · Peng Zhao · Zhi-Hua Zhou

Instructing Goal-Conditioned Reinforcement Learning Agents with Temporal Logic Objectives

Wenjie Qiu · Wensen Mao · He Zhu

Nearly Tight Bounds For Differentially Private Multiway Cut

Mina Dalirrooyfard · Slobodan Mitrovic · Yuriy Nevmyvaka

A Theoretical Analysis of Optimistic Proximal Policy Optimization in Linear Markov Decision Processes

Han Zhong · Tong Zhang

Replicability in Reinforcement Learning

Amin Karbasi · Grigoris Velegkas · Lin Yang · Felix Zhou

Decision-Aware Actor-Critic with Function Approximation and Theoretical Guarantees
Sharan Vaswani · Amirreza Kazemi · Reza Babanezhad Harikandeh · Nicolas Le Roux
Improved Bayesian Regret Bounds for Thompson Sampling in Reinforcement Learning
Ahmadreza Moradipari · Mohammad Pedramfar · Modjtaba Shokrian Zini · Vaneet Aggarwal
"Why Not Looking backward?" A Robust Two-Step Method to Automatically Terminate Bayesian Optimization
Shuang Li · Ke Li · Wei Li
Fractal Landscapes in Policy Optimization
Tao Wang · Sylvia Herbert · Sicun Gao
Truncating Trajectories in Monte Carlo Policy Evaluation: an Adaptive Approach
Riccardo Poiani · Nicole Nobili · Alberto Maria Metelli · Marcello Restelli
Sample Efficient Reinforcement Learning in Mixed Systems through Augmented Samples and Its Applications to Queueing Networks
Honghao Wei · Xin Liu · Weina Wang · Lei Ying
Provably Robust Temporal Difference Learning for Heavy-Tailed Rewards
Semih Cayci · Atilla Eryilmaz
Performance Bounds for Policy-Based Average Reward Reinforcement Learning Algorithms
Yashaswini Murthy · Mehrdad Moharrami · R. Srikant
Variance-Reduced Gradient Estimation via Noise-Reuse in Online Evolution Strategies
Oscar Li · James Harrison · Jascha Sohl-Dickstein · Virginia Smith · Luke Metz
On the Properties of Kullback-Leibler Divergence Between Multivariate Gaussian Distributions
Yufeng Zhang · Jialu Pan · Wanwei Liu · Zhenbang Chen · Xinwang Liu · J Wang · Li Ken Li
RL-based Stateful Neural Adaptive Sampling and Denoising for Real-Time Path Tracing
Antoine Scardigli · Lukas Cavigelli · Lorenz K. Müller
Small batch deep reinforcement learning
Johan Obando Ceron · Marc Bellemare · Pablo Samuel Castro
Unsupervised Behavior Extraction via Random Intent Priors
Hao Hu · Jianing Ye · Yiqin Yang · Ziqing Mai · Chongjie Zhang
Efficient Policy Adaptation with Contrastive Prompt Ensemble for Embodied Agents
wonje choi · Woo Kyung Kim · SeungHyun Kim · Honguk Woo
Enhancing Generalization and Plasticity for Sample Efficient Reinforcement Learning
Hojoon Lee · Hanseul Cho · HYUNSEUNG KIM · DAEHOON GWAK · Joonkee Kim · Jaegul Choo · Se-Young Yun · Chulhee
End-to-End Meta-Bayesian Optimisation with Transformer Neural Processes
Alexandre Maraval · Matthieu Zimmer · Antoine Grosnit · Haitham Bou Ammar
Flexible Attention-Based Multi-Policy Fusion for Efficient Deep Reinforcement Learning
Zih-Yun Chiu · Yi-Lin Tuan · William Yang Wang · Michael Yip
Belief Projection-Based Reinforcement Learning for Environments with Delayed Feedback
Jangwon Kim · Hangyeol Kim · Jiwook Kang · Jongchan Baek · Soohee Han
Debiasing Conditional Stochastic Optimization
Lie He · Shiva Kasiviswanathan
Weakly Coupled Deep Q-Networks
Ibrahim El Shar · Daniel Jiang
State Regularized Policy Optimization on Data with Dynamics Shift
Zhenghai Xue · Qingpeng Cai · Shuchang Liu · Dong Zheng · Peng Jiang · Kun Gai · Bo An
Spectral Entry-wise Matrix Estimation for Low-Rank Reinforcement Learning
Stefan Stojanovic · Yassir Jedra · Alexandre Proutiere
Counterfactual-Augmented Importance Sampling for Semi-Offline Policy Evaluation
Shengpu Tang · Jenna Wiens
TD Convergence: An Optimization Perspective
Kavosh Asadi · Shoham Sabach · Yao Liu · Omer Gottesman · Rasool Fakoor
Regret Minimization via Saddle Point Optimization

Johannes Kirschner · Alireza Bakhtiari · Kushagra Chandak · Volodymyr Tkachuk · Csaba Szepesvari
Efficient RL with Impaired Observability: Learning to Act with Delayed and Missing State Observations
Minshuo Chen · Yu Bai · H. Vincent Poor · Mengdi Wang
Accelerating Value Iteration with Anchoring
Jongmin Lee · Ernest Ryu
Dynamics of Temporal Difference Reinforcement Learning
Blake Bordelon · Paul Masset · Henry Kuo · Cengiz Pehlevan
Taylor TD-learning
Michele Garibbo · Maxime Robeyns · Laurence Aitchison
Prioritizing Samples in Reinforcement Learning with Reducible Loss
Shivakanth Sujit · Somjit Nath · Pedro Braga · Samira Ebrahimi Kahou
Regret-Optimal Model-Free Reinforcement Learning for Discounted MDPs with Short Burn-In Time
Xiang Ji · Gen Li
Tackling Heavy-Tailed Rewards in Reinforcement Learning with Function Approximation: Minimax Optimal and Instance-Dependent Regret Bounds
Jiayi Huang · Han Zhong · Liwei Wang · Lin Yang
Online (Multinomial) Logistic Bandit: Improved Regret and Constant Computation Cost
Yu-Jie Zhang · Masashi Sugiyama
An Alternating Optimization Method for Bilevel Problems under the Polyak-Łojasiewicz Condition
Quan Xiao · Songtao Lu · Tianyi Chen
Complexity of Derivative-Free Policy Optimization for Structured $\mathcal{H}_\infty$ Control
Xingang Guo · Darioush Keivan · Geir Dullerud · Peter Seiler · Bin Hu
Pitfall of Optimism: Distributional Reinforcement Learning by Randomizing Risk Criterion
Taehyun Cho · Seungyub Han · Heesoo Lee · Kyungjae Lee · Jungwoo Lee
A Novel Framework for Policy Mirror Descent with General Parametrization and Linear Convergence
Carlo Alfano · Rui Yuan · Patrick Rebeschini
Efficient Hyper-parameter Optimization with Cubic Regularization
Zhenqian Shen · Hansi Yang · Yong Li · James Kwok · Quanming Yao
Offline RL
Offline Reinforcement Learning with Differential Privacy
Dan Qiao · Yu-Xiang Wang
Provably Efficient Offline Reinforcement Learning in Regular Decision Processes
Roberto Cipollone · Anders Jonsson · Alessandro Ronca · Mohammad Sadegh Talebi
Learning from Visual Observation via Offline Pretrained State-to-Go Transformer
Bohan Zhou · Ke Li · Jiechuan Jiang · Zongqing Lu
Percentile Criterion Optimization in Offline Reinforcement Learning
Cyrus Cousins · Elita Lobo · Marek Petrik · Yair Zick
(More) Sample-Efficient Offline RL with Options
Xiaoyan Hu · Ho-fung Leung
Mutual Information Regularized Offline Reinforcement Learning
Xiao Ma · Bingyi Kang · Zhongwen Xu · Min Lin · Shuicheng Yan
Efficient Diffusion Policies For Offline Reinforcement Learning
Bingyi Kang · Xiao Ma · Chao Du · Tianyu Pang · Shuicheng Yan
Budgeting Counterfactual for Offline RL
Yao Liu · Pratik Chaudhari · Rasool Fakoor
Understanding, Predicting and Better Resolving Q-Value Divergence in Offline-RL
Yang Yue · Rui Lu · Bingyi Kang · Shiji Song · Gao Huang
Conservative State Value Estimation for Offline Reinforcement Learning

Liting Chen · Jie Yan · Zhengdao Shao · Lu Wang · Qingwei Lin · Saravanakumar Rajmohan · Thomas Moscibroda · Dongmei Zhang
Adversarial Model for Offline Reinforcement Learning
Mohak Bhardwaj · Tengyang Xie · Byron Boots · Nan Jiang · Ching-An Cheng
Recovering from Out-of-sample States via Inverse Dynamics in Offline Reinforcement Learning
Ke Jiang · Jia-Yu Yao · Xiaoyang Tan
Importance Weighted Actor-Critic for Optimal Conservative Offline Reinforcement Learning
Hanlin Zhu · Paria Rashidinejad · Jiantao Jiao
Goal-Conditioned Predictive Coding as an Implicit Planner for Offline Reinforcement Learning
Zilai Zeng · Ce Zhang · Shijie Wang · Chen Sun
Off-Policy Evaluation for Human Feedback
Qitong Gao · Ge Gao · Juncheng Dong · Vahid Tarokh · Min Chi · Miroslav Pajic
Cal-QL: Calibrated Offline RL Pre-Training for Efficient Online Fine-Tuning
Mitsuhiko Nakamoto · Yuexiang Zhai · Anikait Singh · Max Sobol Mark · Yi Ma · Chelsea Finn · Aviral Kumar · Sergey Levine
Direct Preference-based Policy Optimization without Reward Modeling
Gaon An · Junhyeok Lee · Xingdong Zuo · Norio Kosaka · Kyung-Min Kim · Hyun Oh Song
Addressing Out-Of-Distribution Joint Actions in Offline Multi-Agent RL via Alternating Stationary Distribution Correction Estimation
Daiki E. Matsunaga · Jongmin Lee · Jaeseok Yoon · Stefanos Leonardos · Pieter Abbeel · Kee-Eung Kim
Bi-Level Offline Policy Optimization with Limited Exploration
Wenzhuo Zhou
Reward-agnostic Fine-tuning: Provable Statistical Benefits of Hybrid Reinforcement Learning
Gen Li · Wenhao Zhan · Jason Lee · Yuejie Chi · Yuxin Chen
Corruption-Robust Offline Reinforcement Learning with General Function Approximation
Chenlu Ye · Rui Yang · Quanquan Gu · Tong Zhang
Conformal Prediction for Uncertainty-Aware Planning with Diffusion Dynamics Model
Jiankai Sun · Yiqi Jiang · Jianing Qiu · Parth Nobel · Mykel J Kochenderfer · Mac Schwager
Offline Reinforcement Learning for Mixture-of-Expert Dialogue Management
Dhawal Gupta · Yinlam Chow · Azamat Tulepbergenov · Mohammad Ghavamzadeh · Craig Boutilier
Constrained Policy Optimization with Explicit Behavior Density For Offline Reinforcement Learning
Jing Zhang · Chi Zhang · Wenjia Wang · Bingyi Jing
Beyond Uniform Sampling: Offline Reinforcement Learning with Imbalanced Datasets
Zhang-Wei Hong · Aviral Kumar · Sathwik Karnik · Abhishek Bhandwaldar · Akash Srivastava · Joni Pajarinen · Romain Laroche · Abhishek Gupta · Pulkit Agrawal
Double Pessimism is Provably Efficient for Distributionally Robust Offline Reinforcement Learning: Generic Algorithm and Robust Partial Coverage
Jose Blanchet · Miao Lu · Tong Zhang · Han Zhong
Unified Off-Policy Learning to Rank: a Reinforcement Learning Perspective
Zeyu Zhang · Yi Su · Hui Yuan · Yiran Wu · Rishab Balasubramanian · Qingyun Wu · Huazheng Wang · Mengdi Wang
Reining Generalization in Offline Reinforcement Learning via Representation Distinction
Yi Ma · Hongyao Tang · Dong Li · Zhaopeng Meng
Offline Minimax Soft-Q-learning Under Realizability and Partial Coverage
Masatoshi Uehara · Nathan Kallus · Jason Lee · Wen Sun
Supported Value Regularization for Offline Reinforcement Learning
Yixiu Mao · Hongchang Zhang · Chen Chen · Yi Xu · Xiangyang Ji
One Risk to Rule Them All: Addressing Distributional Shift in Offline Reinforcement Learning via Risk-Aversion
Marc Rigter · Bruno Lacerda · Nick Hawes
VOCE: Variational Optimization with Conservative Estimation for Offline Safe Reinforcement Learning
Jiayi Guan · Guang Chen · Jiaming Ji · Long Yang · ao zhou · Zhijun Li · changjun jiang
Adversarial Counterfactual Environment Model Learning
Xiong-Hui Chen · Yang Yu · Zhengmao Zhu · ZhiHua Yu · Chen Zhenjun · Chenghe Wang · Yinan Wu · Rong-Jun Qin · Hongqiu Wu · Ruijin Ding · Huang Fangsheng
Look Beneath the Surface: Exploiting Fundamental Symmetry for Sample-Efficient Offline RL

Peng Cheng · Xianyuan Zhan · zhihao wu · Wenjia Zhang · Youfang Lin · Shou cheng Song · Han Wang · Li Jiang
ConSpec: honing in on critical steps for rapid learning and generalization in RL
Chen Sun · Wannan Yang · Thomas Jiralerspong · Dane Malenfant · Benjamin Alsbury-Nealy · Yoshua Bengio · Blake
SPQR: Controlling Q-ensemble Independence for Reinforcement Learning
Dohyeok Lee · Seungyub Han · Taehyun Cho · Jungwoo Lee
Provably Efficient Offline Goal-Conditioned Reinforcement Learning with General Function Approximation and Single-Policy Concentrability
Hanlin Zhu · Amy Zhang
Policy Finetuning in Reinforcement Learning via Design of Experiments using Offline Data
Ruiqi Zhang · Andrea Zanette
Learning to Influence Human Behavior with Offline Reinforcement Learning
Joey Hong · Sergey Levine · Anca Dragan
Offline RL with Discrete Proxy Representations for Generalizability in POMDPs
Pengjie Gu · Xinyu Cai · Dong Xing · Xinrun Wang · Mengchen Zhao · Bo An
Revisiting the Minimalist Approach to Offline Reinforcement Learning
Denis Tarasov · Vladislav Kurenkov · Alexander Nikulin · Sergey Kolesnikov
Survival Instinct in Offline Reinforcement Learning
Anqi Li · Dipendra Misra · Andrey Kolobov · Ching-An Cheng
CORL: Research-oriented Deep Offline Reinforcement Learning Library
Denis Tarasov · Alexander Nikulin · Dmitry Akimov · Vladislav Kurenkov · Sergey Kolesnikov
Hokoff: Real Game Dataset from Honor of Kings and its Offline Reinforcement Learning Benchmarks
Yun Qu · Boyuan Wang · Jianzhun Shao · Yuhang Jiang · Chen Chen · Zhenbin Ye · Liu Linc · Yang Feng · Lin Lai · Hongyang Qin · Minwen Deng · Juchao Zhuo · Deheng Ye · Qiang Fu · YANG GUANG · Wei Yang · Lanxiao Huang ·
Understanding and Addressing the Pitfalls of Bisimulation-based Representations in Offline Reinforcement Learning
Hongyu Zang · Xin Li · Leiji Zhang · Yang Liu · Baigui Sun · Riashat Islam · Remi Tachet des Combes · Romain Laroche
On Sample-Efficient Offline Reinforcement Learning: Data Diversity, Posterior Sampling and Beyond
Thanh Nguyen-Tang · Raman Arora
ReDS: Offline RL With Heteroskedastic Datasets via Support Constraints
Anikait Singh · Aviral Kumar · Quan Vuong · Yevgen Chebotar · Sergey Levine
Multi-agent/Game Theory
Arbitrarily Scalable Environment Generators via Neural Cellular Automata
Yulun Zhang · Matthew Fontaine · Varun Bhatt · Stefanos Nikolaidis · Jiaoyang Li
Connected Superlevel Set in (Deep) Reinforcement Learning and its Application to Minimax Theorems
Sihan Zeng · Thinh Doan · Justin Romberg
Offline Multi-Agent Reinforcement Learning with Implicit Global-to-Local Value Regularization
Xiangsen Wang · Haoran Xu · Yinan Zheng · Xianyuan Zhan
DIFFER:Decomposing Individual Reward for Fair Experience Replay in Multi-Agent Reinforcement Learning
Xunhan Hu · Jian Zhao · Wengang Zhou · Ruili Feng · Houqiang Li
IMP-MARL: a Suite of Environments for Large-scale Infrastructure Management Planning via MARL
Pascal Leroy · Pablo G. Morato · Jonathan Pisane · Athanasios Kolios · Damien Ernst
Information Design in Multi-Agent Reinforcement Learning
Yue Lin · Wenhao Li · Hongyuan Zha · Baoxiang Wang
Diverse Conventions for Human-AI Collaboration
Bidipta Sarkar · Andy Shih · Dorsa Sadigh
Conservative Offline Policy Adaptation in Multi-Agent Games
Chengjie Wu · Pingzhong Tang · Jun Yang · Yujing Hu · Tangjie Lv · Changjie Fan · Chongjie Zhang
Counterfactual Conservative Q Learning for Offline Multi-agent Reinforcement Learning
Jianzhun Shao · Yun Qu · Chen Chen · Hongchang Zhang · Xiangyang Ji
Efficient Adversarial Attacks on Online Multi-agent Reinforcement Learning
Guanlin Liu · Lifeng LAI

Dual Self-Awareness Value Decomposition Framework without Individual Global Max for Cooperative MARL
Zhiwei Xu · Bin Zhang · dapeng li · Guangchong Zhou · Zeren Zhang · Guoliang Fan
Multi-Agent Meta-Reinforcement Learning: Sharper Convergence Rates with Task Similarity
Weichao Mao · Haoran Qiu · Chen Wang · Hubertus Franke · Zbigniew Kalbarczyk · Ravishankar Iyer · Tamer Basar
Two Heads are Better Than One: A Simple Exploration Framework for Efficient Multi-Agent Reinforcement Learning
Jiahui Li · Kun Kuang · Baoxiang Wang · Xingchen Li · Long Chen · Fei Wu · Jun Xiao
Automatic Grouping for Efficient Cooperative Multi-Agent Reinforcement Learning
Yifan Zang · Jinmin He · Kai Li · Haobo Fu · Qiang Fu · Junliang Xing · Jian Cheng
Long-Term Fairness with Unknown Dynamics
Tongxin Yin · Reilly Raab · Mingyan Liu · Yang Liu
Exploiting hidden structures in non-convex games for convergence to Nash equilibrium
Iosif Sakos · Emmanouil-Vasileios Vlatakis-Gkaragkounis · Panayotis Mertikopoulos · Georgios Piliouras
Selectively Sharing Experiences Improves Multi-Agent Reinforcement Learning
Matthias Gerstgrasser · Tom Danino · Sarah Keren
RiskQ: Risk-sensitive Multi-Agent Reinforcement Learning Value Factorization
Siqi Shen · Chennan Ma · Chao Li · Weiquan Liu · Yongquan Fu · Songzhu Mei · Xinwang Liu · Cheng Wang
Learning to Ignore: Mutual-Information Regularized Multi-Agent Policy Iteration
Wang · Deheng Ye · Zongqing Lu
Robust Multi-Agent Reinforcement Learning via Adversarial Regularization: Theoretical Foundation and Stable
Alexander Bukharin · Yan Li · Yue Yu · Qingru Zhang · Zhehui Chen · Simiao Zuo · Chao Zhang · Songan Zhang · Tuo Zhao
Scalable Primal-Dual Actor-Critic Method for Safe Multi-Agent RL with General Utilities
Donghao Ying · Yunkai Zhang · Yuhao Ding · Alec Koppel · Javad Lavaei
Multi-Agent First Order Constrained Optimization in Policy Space
Youpeng Zhao · Yaodong Yang · Zhenbo Lu · Wengang Zhou · Houqiang Li
Decompose a Task into Generalizable Subtasks in Multi-Agent Reinforcement Learning
Zikang Tian · Ruizhi Chen · Fan Wu · Rui Zhang · Ling Li · Xing Hu · Shaohui Peng · Jiaming Guo · Zidong Du · Qi Guo · Yunji Chen
Team-PSRO for Learning Approximate TMECor in Large Team Games via Cooperative Reinforcement Learning
Stephen McAleer · Gabriele Farina · Gaoyue Zhou · Mingzhi Wang · Yaodong Yang · Tuomas Sandholm
Iteratively Learn Diverse Strategies with State Distance Information
Wei Fu · Weihua Du · Jingwei Li · Sunli Chen · Jingzhao Zhang · YI WU
Semantic HELM: An Interpretable Memory for Reinforcement Learning
Fabian Paischer · Thomas Adler · Markus Hofmarcher · Sepp Hochreiter
Hierarchical Multi-Agent Skill Discovery
Mingyu Yang · Yaodong Yang · Zhenbo Lu · Wengang Zhou · Houqiang Li
OFCOURSE: A Multi-Agent Reinforcement Learning Environment for Order Fulfillment
Yiheng Zhu · Yang Zhan · Xuankun Huang · Yuwei Chen · yujie Chen · Jiangwen Wei · Wei Feng · Yinzhi Zhou · Haoyuan Hu · Jieping Ye
Gigastep - One Billion Steps per Second Multi-agent Reinforcement Learning
Mathias Lechner · lianhao yin · Tim Seyde · Tsun-Hsuan Johnson Wang · Wei Xiao · Ramin Hasani · Joshua Rountree · Daniela Rus
Exponential Lower Bounds for Fictitious Play in Potential Games
Ioannis Panageas · Nikolas Patris · Stratis Skoulakis · Volkan Cevher
SMACv2: An Improved Benchmark for Cooperative Multi-Agent Reinforcement Learning
Benjamin Ellis · Jonathan Cook · Skander Moalla · Mikayel Samvelyan · Mingfei Sun · Anuj Mahajan · Jakob Foerster · Shimon Whiteson
On Imitation in Mean-field Games
Giorgia Ramponi · Pavel Kolev · Olivier Pietquin · Niao He · Mathieu Lauriere · Matthieu Geist
Computing Optimal Equilibria and Mechanisms via Learning in Zero-Sum Extensive-Form Games
Brian Zhang · Gabriele Farina · Ioannis Anagnostides · Federico Cacciamani · Stephen McAleer · Andreas Haupt · Andrea Celli · Nicola Gatti · Vincent Conitzer · Tuomas Sandholm
Exploration and Exploitation

Latent exploration for Reinforcement Learning
Alberto Silvio Chiappa · Alessandro Marin Vargas · Ann Huang · Alexander Mathis
Goal-conditioned Offline Planning from Curious Exploration
Marco Bagatella · Georg Martius
A Partially-Supervised Reinforcement Learning Framework for Visual Active Search
Anindya Sarkar · Nathan Jacobs · Yevgeniy Vorobeychik
Posterior Sampling with Delayed Feedback for Reinforcement Learning with Linear Function Approximation
Nikki Lijing Kuang · Ming Yin · Mengdi Wang · Yu-Xiang Wang · Yian Ma
Safe Exploration in Reinforcement Learning: A Generalized Formulation and Algorithms
Akifumi Wachi · Wataru Hashimoto · Xun Shen · Kazumune Hashimoto
Efficient Exploration in Continuous-time Model-based Reinforcement Learning
Lenart Treven · Jonas Hübotter · Bhavya · Florian Dorfler · Andreas Krause
Successor-Predecessor Intrinsic Exploration
Changmin Yu · Neil Burgess · Samuel J Gershman · Maneesh Sahani
ELDEN: Exploration via Local Dependencies
Zizhao Wang · Jiaheng Hu · Roberto Martín-Martín · Peter Stone
Deep Reinforcement Learning with Plasticity Injection
Evgenii Nikishin · Junhyuk Oh · Georg Ostrovski · Clare Lyle · Razvan Pascanu · Will Dabney · Andre Barreto
Breadcrumbs to the Goal: Supervised Goal Selection from Human-in-the-Loop Feedback
Marcel Torne Villasevil · Max Balsells I Pamies · Zihan Wang · Samedh Desai · Tao Chen · Pulkit Agrawal · Abhishek Gupta
On the Importance of Exploration for Generalization in Reinforcement Learning
Yiding Jiang · J. Zico Kolter · Roberta Raileanu
Reinforcement-Enhanced Autoregressive Feature Transformation: Gradient-steered Search in Continuous Space for Postfix Expressions
Dongjie Wang · Meng Xiao · Min Wu · pengfei wang · Yuanchun Zhou · Yanjie Fu
Learning to Search Feasible and Infeasible Regions of Routing Problems with Flexible Neural k-Opt
Yining Ma · Zhiguang Cao · Yeow Meng Chee
Accelerating Exploration with Unlabeled Prior Data
Qiyang Li · Jason Zhang · Dibya Ghosh · Amy Zhang · Sergey Levine
Optimistic Exploration in Reinforcement Learning Using Symbolic Model Estimates
Sarath Sreedharan · Michael Katz
Optimistic Active Exploration of Dynamical Systems
Bhavya · Lenart Treven · Cansu Sancaktar · Sebastian Blaes · Stelian Coros · Andreas Krause
Keep Various Trajectories: Promoting Exploration of Ensemble Policies in Continuous Control
Chao Li · Chen GONG · Qiang He · Xinwen Hou
Efficient Potential-based Exploration in Reinforcement Learning using Inverse Dynamic Bisimulation Metric
Yiming Wang · Ming Yang · Renzhi Dong · Binbin Sun · Leong Hou U · Furui Liu
Explore to Generalize in Zero-Shot RL
Ev Zisselman · Itai Lavie · Daniel Soudry · Aviv Tamar
Accelerating Reinforcement Learning with Value-Conditional State Entropy Exploration
Dongyoung Kim · Jinwoo Shin · Pieter Abbeel · Younggyo Seo
One Objective to Rule Them All: A Maximization Objective Fusing Estimation and Planning for Exploration
Zhihan Liu · Miao Lu · WEI XIONG · Han Zhong · Hao Hu · Shenao Zhang · Sirui Zheng · Zhuoran Yang · Zhaoran Wang
Diversify \& Conquer: Outcome-directed Curriculum RL via Out-of-Distribution Disagreement
Daesol Cho · Seungjae Lee · H. Jin Kim
Model-Free Active Exploration in Reinforcement Learning
Alessio Russo · Alexandre Proutiere
Large Language Model/RLHF
Fine-Grained Human Feedback Gives Better Rewards for Language Model Training

Zeqiu Wu · Yushi Hu · Weijia Shi · Nouha Dziri · Alane Suhr · Prithviraj (Raj) Ammanabrolu · Noah Smith · Mari Ostendorf · Hannaneh Hajishirzi

Optimizing Prompts for Text-to-Image Generation

Yaru Hao · Zewen Chi · Li Dong · Furu Wei

Language Models can Solve Computer Tasks

Geunwoo Kim · Pierre Baldi · Stephen McAleer

Blockwise Parallel Transformer for Large Models

Hao Liu · Pieter Abbeel

Chatting Makes Perfect - Chat-based Image Retrieval

Matan Levy · Rami Ben-Ari · Nir Darshan · Dani Lischinski

Self-Refine: Iterative Refinement with Self-Feedback

Aman Madaan · Niket Tandon · Prakhar Gupta · Skyler Hallinan · Luyu Gao · Sarah Wiegreffe · Uri Alon · Nouha Dziri · Shrimai Prabhumoye · Yiming Yang · Shashank Gupta · Bodhisattwa Prasad Majumder · Katherine Hermann · Sean Welleck · Amir Yazdanbakhsh · Peter Clark

Guiding Large Language Models via Directional Stimulus Prompting

Zekun Li · Baolin Peng · Pengcheng He · Michel Galley · Jianfeng Gao · Xifeng Yan

Large Language Model Is Semi-Parametric Reinforcement Learning Agent

Danyang Zhang · Lu Chen · Situo Zhang · Hongshen Xu · Zihan Zhao · Kai Yu

Principle-Driven Self-Alignment of Language Models from Scratch with Minimal Human Supervision

Zhiqing Sun · Yikang Shen · Qinhong Zhou · Hongxin Zhang · Zhenfang Chen · David Cox · Yiming Yang · Chuang Gan

SPRING: Studying Papers and Reasoning to play Games

Yue Wu · So Yeon Min · Shrimai Prabhumoye · Yonatan Bisk · Russ Salakhutdinov · Amos Azaria · Tom Mitchell · Yuanzhi Li

Teaching Cars to See in a Day: Unsupervised Object Discovery with Reward Fine-tuning

Katie Luo · Zhenzhen Liu · Xiangyu Chen · Yurong You · Sagie Benaim · Cheng Perng Phoo · Mark Campbell · Wen Sun · Bharath Hariharan · Kilian Weinberger

Reflexion: language agents with verbal reinforcement learning

Noah Shinn · Federico Cassano · Ashwin Gopinath · Karthik Narasimhan · Shunyu Yao

Direct Preference Optimization: Your Language Model is Secretly a Reward Model

Rafael Rafailov · Archit Sharma · Eric Mitchell · Christopher D Manning · Stefano Ermon · Chelsea Finn

OpenAGI: When LLM Meets Domain Experts

Yingqiang Ge · Wenyue Hua · Kai Mei · jianchao ji · Juntao Tan · Shuyuan Xu · Zelong Li · Yongfeng Zhang

BeaverTails: A Human-Preference Dataset for LLM Harmlessness Alignment

Jiaming Ji · Mickel Liu · Josef Dai · Xuehai Pan · Chi Zhang · Ce Bian · Boyuan Chen · Ruiyang Sun · Yizhou Wang · Yaodong Yang

OpenAssistant Conversations - Democratizing Large Language Model Alignment

Andreas Köpf · Yannic Kilcher · Dimitri von Rütte · Sotiris Anagnostidis · Zhi Rui Tam · Keith Stevens · Abdullah Barhoum · Duc Nguyen · Oliver Stanley · Richárd Nagyfi · Shahul ES · Sameer Suri · David Glushkov · Arnav Dantuluri · Andrew Maguire · Christoph Schuhmann · Huu Nguyen · Alexander Mattick

Language Model Alignment with Elastic Reset

Michael Noukhovitch · Samuel Lavoie · Florian Strub · Aaron Courville

LIMA: Less Is More for Alignment

Chunting Zhou · Pengfei Liu · Puxin Xu · Srinivasan Iyer · Jiao Sun · Yuning Mao · Xuezhe Ma · Avia Efrat · Ping Yu · LILI YU · Susan Zhang · Gargi Ghosh · Mike Lewis · Luke Zettlemoyer · Omer Levy

RRHF: Rank Responses to Align Language Models with Human Feedback without tears

Zheng Yuan · Hongyi Yuan · Chuanqi Tan · Wei Wang · Songfang Huang · Fei Huang

Aligning Language Models with Human Preferences via a Bayesian Approach

Jiashuo WANG · Haozhao Wang · Shichao Sun · Wenjie Li

Sequential Preference Ranking for Efficient Reinforcement Learning from Human Feedback

Minyoung Hwang · Gunmin Lee · Hogun Kee · Chan Woo Kim · Kyungjae Lee · Songhwai Oh

Is RLHF More Difficult than Standard RL?

Yuanhao Wang · Qinghua Liu · Chi Jin

AI safety/AI Alignment/Robust/Safe RL/Constrained/Uncertainty

Risk-Averse Model Uncertainty for Distributionally Robust Safe Reinforcement Learning
James Queeney · Mouhacine Benosman
Provably Safe Reinforcement Learning with Step-wise Violation Constraints
Nuoya Xiong · Yihan Du · Longbo Huang
Constraint-Conditioned Policy Optimization for Versatile Safe Reinforcement Learning
Yihang Yao · ZUXIN LIU · Zhepeng Cen · Jiacheng Zhu · Wenhao Yu · Tingnan Zhang · DING ZHAO
Generative Modelling of Stochastic Actions with Arbitrary Constraints in Reinforcement Learning
Changyu CHEN · Ramesha Karunasena · Thanh Nguyen · Arunesh Sinha · Pradeep Varakantham
On Dynamic Programming Decompositions of Static Risk Measures in Markov Decision Processes
Jia Lin Hau · Erick Delage · Mohammad Ghavamzadeh · Marek Petrik
Policy Gradient for Rectangular Robust Markov Decision Processes
Navdeep Kumar · Esther Derman · Matthieu Geist · Kfir Y. Levy · Shie Mannor
Anytime-Constrained Reinforcement Learning with Policy Prior
Jianyi Yang · Pengfei Li · Tongxin Li · Adam Wierman · Shaolei Ren
Seeing is not Believing: Robust Reinforcement Learning against Spurious Correlation
Wenhao Ding · Laixi Shi · Yuejie Chi · DING ZHAO
Distributional Model Equivalence for Risk-Sensitive Reinforcement Learning
Tyler Kastner · Murat Erdogdu · Amir-massoud Farahmand
Reduced Policy Optimization for Continuous Control with Hard Constraints
Shutong Ding · Jingya Wang · Yali Du · Ye Shi
FlowPG: Action-constrained Policy Gradient with Normalizing Flows
Janaka Brahmanage · Jiajing LING · Akshat Kumar
Trust Region-Based Safe Distributional Reinforcement Learning for Multiple Constraints
Dohyeong Kim · Kyungjae Lee · Songhwai Oh
Iterative Reachability Estimation for Safe Reinforcement Learning
Milan Ganai · Zheng Gong · Chenning Yu · Sylvia Herbert · Sicun Gao
Uncertainty Estimation for Safety-critical Scene Segmentation via Fine-grained Reward Maximization
Hongzheng Yang · Cheng Chen · Yueyao CHEN · Scheppach · Hon Chi Yip · DOU QI
Safety Gymnasium: A Unified Safe Reinforcement Learning Benchmark
Jiaming Ji · Borong Zhang · Jiayi Zhou · Xuehai Pan · Weidong Huang · Ruiyang Sun · Yiran Geng · Josef Dai · Yaodong
Natural Actor-Critic for Robust Reinforcement Learning with Function Approximation
Ruida Zhou · Tao Liu · Min Cheng · Dileep Kalathil · P. R. Kumar · Chao Tian
Multi-Modal Inverse Constrained Reinforcement Learning from a Mixture of Demonstrations
Guanren Qiao · Guiliang Liu · Pascal Poupart · Zhiqiang Xu
Honesty Is the Best Policy: Defining and Mitigating AI Deception
Francis Ward · Francesca Toni · Francesco Belardinelli · Tom Everitt
Sample-Efficient and Safe Deep Reinforcement Learning via Reset Deep Ensemble Agents
WOOJUN KIM · Yongjae Shin · Jongeui Park · Youngchul Sung
Bayesian Risk-Averse Q-Learning with Streaming Observations
Yuhao Wang · Enlu Zhou
An Alternative to Variance: Gini Deviation for Risk-averse Policy Gradient
Yudong Luo · Guiliang Liu · Pascal Poupart · Yangchen Pan
Learning from Active Human Involvement through Proxy Value Propagation
Zhenghao Peng · Wenjie Mo · Chenda Duan · Quanyi Li · Bolei Zhou
Meta learning/Transfer/Multi-task/Generalization
Contrastive Modules with Temporal Attention for Multi-Task Reinforcement Learning
Siming Lan · Rui Zhang · Qi Yi · Jiaming Guo · Shaohui Peng · Yunkai Gao · Fan Wu · Ruizhi Chen · Zidong Du · Xing Hu · xishan zhang · Ling Li · Yunji Chen
Learning Shared Safety Constraints from Multi-task Demonstrations

Konwoo Kim · Gokul Swamy · ZUXIN LIU · DING ZHAO · Sanjiban Choudhury · Steven Wu

Direction-oriented Multi-objective Learning: Simple and Provable Stochastic Algorithms

Peiyao Xiao · Hao Ban · Kaiyi Ji

Robust Knowledge Transfer in Tiered Reinforcement Learning

Jiawei Huang · Niao He

Diffused Task-Agnostic Milestone Planner

Mineui Hong · Minjae Kang · Songhwai Oh

Parameterizing Non-Parametric Meta-Reinforcement Learning Tasks via Subtask Decomposition

Suyoung Lee · Myungsik Cho · Youngchul Sung

Context Shift Reduction for Offline Meta-Reinforcement Learning

Yunkai Gao · Rui Zhang · Jiaming Guo · Fan Wu · Qi Yi · Shaohui Peng · Siming Lan · Ruizhi Chen · Zidong Du · Xing Hu · Qi Guo · Ling Li · Yunji Chen

Cross-Episodic Curriculum for Transformer Agents

Lucy Xiaoyang Shi · Yunfan Jiang · Jake Grigsby · Linxi Fan · Yuke Zhu

Train Hard, Fight Easy: Robust Meta Reinforcement Learning

Ido Greenberg · Shie Mannor · Gal Chechik · Eli Meirom

Multi-Objective Intrinsic Reward Learning for Conversational Recommender Systems

Zhendong Chu · Nan Wang · Hongning Wang

Cross-Domain Policy Adaptation via Value-Guided Data Filtering

Kang Xu · Chenjia Bai · Xiaoteng Ma · Dong Wang · Bin Zhao · Zhen Wang · Xuelong Li · Wei Li

Discovering General Reinforcement Learning Algorithms with Adversarial Environment Design

Matthew T Jackson · Minqi Jiang · Jack Parker-Holder · Risto Vuorio · Chris Lu · Greg Farquhar · Shimon Whiteson · Jakob Foerster

Provably Fast Convergence of Independent Natural Policy Gradient for Markov Potential Games

Youbang Sun · Tao Liu · Ruida Zhou · P. R. Kumar · Shahin Shahrampour

Doubly Robust Augmented Transfer for Meta-Reinforcement Learning

Yuankun Jiang · Nuowen Kan · Chenglin Li · Wenrui Dai · Junni Zou · Hongkai Xiong

Recurrent Hypernetworks are Surprisingly SOTA in Meta-RL

Jacob Beck · Risto Vuorio · Zheng Xiong · Shimon Whiteson

FAMO: Fast Adaptive Multitask Optimization

Bo Liu · Yihao Feng · Peter Stone · Qiang Liu

Efficient Symbolic Policy Learning via Gradient Descent

Jiaming Guo · Rui Zhang · Shaohui Peng · Qi Yi · Xing Hu · Ruizhi Chen · Zidong Du · xishan zhang · Ling Li · Qi Guo · Yunji

A Toolkit for Reliable Benchmarking and Research in Multi-Objective Reinforcement Learning

Florian Felten · Lucas N. Alegre · Ann Nowe · Ana Bazzan · El Ghazali Talbi · Grégoire Danoy · Bruno da Silva

Self-Supervised Reinforcement Learning that Transfers using Random Features

Boyuan Chen · Chuning Zhu · Pulkit Agrawal · Kaiqing Zhang · Abhishek Gupta

Neural MMO 2.0: A Massively Multi-task Addition to Massively Multi-agent Learning

Joseph Suarez · David Bloomin · Kyoung Choe · Hao Xiang Li · Ryan Sullivan · Nishaanth Kanna · Daniel Scott · Rose Shuman · Herbie Bradley · Louis Castricato · Phillip Isola · Chenghui Yu · Yuhao Jiang · Qimai Li · Jiaxin Chen · Xiaolong

Learning Pareto-Optimal Policies for Multi-Objective Joint Distribution

Xin-Qiang Cai · Pushi Zhang · Li Zhao · Jiang Bian · Masashi Sugiyama · Ashley Llorens

Model-based RL

Hyperbolic VAE via Latent Gaussian Distributions

Seunghyuk Cho · Juyong Lee · Dongwoo Kim

Task-aware world model learning with meta weighting via bi-level optimization

Huining Yuan · Hongkun Dou · Xingyu Jiang · Yue Deng

RePo: Resilient Model-Based Reinforcement Learning by Regularizing Posterior Predictability

Chuning Zhu · Max Simchowitz · Siri Gadipudi · Abhishek Gupta

Multi-Step Generalized Policy Improvement by Leveraging Approximate Models

Lucas N. Alegre · Ana Bazzan · Ann Nowe · Bruno da Silva
Reward Scale Robustness for Proximal Policy Optimization via DreamerV3 Tricks
Ryan Sullivan · Akarsh Kumar · Shengyi Huang · John Dickerson · Joseph Suarez
MoVie: Visual Model-Based Policy Adaptation for View Generalization
Sizhe Yang · Yanjie Ze · Huazhe Xu
STORM: Efficient Stochastic Transformer based World Models for Reinforcement Learning
Weipu Zhang · Gang Wang · Jian Sun · Yetian Yuan · Gao Huang
Posterior Sampling for Competitive RL: Function Approximation and Partial Observation
Shuang Qiu · Ziyu Dai · Han Zhong · Zhaoran Wang · Zhuoran Yang · Tong Zhang
The Curious Price of Distributional Robustness in Reinforcement Learning with a Generative Model
Laixi Shi · Gen Li · Yuting Wei · Yuxin Chen · Matthieu Geist · Yuejie Chi
Model-Based Reparameterization Policy Gradient Methods: Theory and Practical Algorithms
Shenao Zhang · Boyi Liu · Zhaoran Wang · Tuo Zhao
Optimal Exploration for Model-Based RL in Nonlinear Systems
Andrew Wagenmaker · Guanya Shi · Kevin Jamieson
Would I have gotten that reward? Long-term credit assignment by counterfactual contribution analysis
Alexander Meulemans · Simon Schug · Seijin Kobayashi · nathaniel daw · Gregory Wayne
Facing-off World Model Backbones: RNN, Transformer, and S4
Fei Deng · Junyeong Park · Sungjin Ahn
How to Fine-tune the Model: Unified Model Shift and Model Bias Policy Optimization
Hai Zhang · Hang Yu · Junqiao Zhao · Di Zhang · xiao zhang · Hongtu Zhou · Chang Huang · Chen Ye
Regularity as Intrinsic Reward for Free Play
Cansu Sancaktar · Justus Piater · Georg Martius
Adapt to Adapt: A Tempo-control Framework for Non-stationary Reinforcement Learning
Hyunin Lee · Yuhao Ding · Jongmin Lee · Ming Jin · Javad Lavaei · Somayeh Sojoudi
Representation Learning
When Do Transformers Shine in RL? Decoupling Memory from Credit Assignment
Tianwei Ni · Michel Ma · Benjamin Eysenbach · Pierre-Luc Bacon
Learning World Models with Identifiable Factorization
Yuren Liu · Biwei Huang · Zhengmao Zhu · Honglong Tian · Mingming Gong · Yang Yu · Kun Zhang
StateMask: Explaining Deep Reinforcement Learning through State Mask
Zelei Cheng · Xian Wu · Jiahao Yu · Wenbo Guo · Wenhai Sun · Xinyu Xing
Structured State Space Models for In-Context Reinforcement Learning
Chris Lu · Yannick Schroecker · Albert Gu · Emilio Parisotto · Jakob Foerster · Satinder Singh · Feryal Behbahani
A State Representation for Diminishing Rewards
Ted Moskovitz · Samo Hromadka · Ahmed Touati · Diana Borsa · Maneesh Sahani
For SALE: State-Action Representation Learning for Deep Reinforcement Learning
Scott Fujimoto · Wei-Di Chang · Edward Smith · Shixiang (Shane) Gu · Doina Precup · David Meger
State Sequences Prediction via Fourier Transform for Representation Learning
Mingxuan Ye · Yufei Kuang · Jie Wang · Yang Rui · Wengang Zhou · Houqiang Li · Feng Wu
Conditional Mutual Information for Disentangled Representations in Reinforcement Learning
Mhairi Dunion · Trevor McInroe · Kevin Sebastian Luck · Josiah Hanna · Stefano Albrecht
Learning dynamic attribute-factored world models for efficient multi-object reinforcement learning
Fan Feng · Sara Magliacane
H-InDex: Visual Reinforcement Learning with Hand-Informed Representations for Dexterous Manipulation
Yanjie Ze · Yuyao Liu · Ruizhe Shi · Jiaxin Qin · Zhecheng Yuan · Jiashun Wang · Xiaolong Wang · Huazhe Xu
State-Action Similarity-Based Representations for Off-Policy Evaluation
Brahma Pavse · Josiah Hanna

Distributional Policy Evaluation: a Maximum Entropy approach to Representation Learning
Riccardo Zamboni · Alberto Maria Metelli · Marcello Restelli
Reinforcement Learning with Simple Sequence Priors
Tankred Saanum · Noemi Elteto · Peter Dayan · Marcel Binz · Eric Schulz
$\texttt{TACO}$: Temporal Latent Action-Driven Contrastive Loss for Visual Reinforcement Learning
Ruijie Zheng · Xiyao Wang · Yanchao Sun · Shuang Ma · Jieyu Zhao · Huazhe Xu · Hal Daumé III · Furong Huang
Combinatorial Optimization
Adjustable Robust Reinforcement Learning for Online 3D Bin Packing
Yuxin Pan · Yize Chen · Fangzhen Lin
Efficient Meta Neural Heuristic for Multi-Objective Combinatorial Optimization
Jinbiao Chen · Zizhen Zhang · Te Ye · Zhiguang Cao · Siyuan Chen · Jiahai Wang
Resetting the Optimizer in Deep RL: An Empirical Study
Kavosh Asadi · Rasool Fakoor · Shoham Sabach
OneNet: Enhancing Time Series Forecasting Models under Concept Drift by Online Ensembling
yifan zhang · Qingsong Wen · xue wang · Weiqi Chen · Liang Sun · Zhang Zhang · Liang Wang · Rong Jin · Tieniu Tan
Neural Multi-Objective Combinatorial Optimization with Diversity Enhancement
Jinbiao Chen · Zizhen Zhang · Zhiguang Cao · Yaoxin Wu · Yining Ma · Te Ye · Jiahai Wang
DeepACO: Neural-enhanced Ant Systems for Combinatorial Optimization
Haoran Ye · Jiarui Wang · Zhiguang Cao · Helan Liang · Yong Li
Compositional Policy Learning in Stochastic Control Systems with Formal Guarantees
Krishnendu Chatterjee · Thomas Henzinger · Mathias Lechner · Abhinav Verma · Đorđe Žikelić
Swarm Reinforcement Learning for Adaptive Mesh Refinement
Niklas Freymuth · Philipp Dahlinger · Tobias Würth · Simon Reisch · Luise Kärger · Gerhard Neumann
Beyond Average Reward in Markov Decision Processes
Alexandre Marthe · Aurélien Garivier · Claire Vernade
Optimizing Solution-Samplers for Combinatorial Problems: The Landscape of Policy-Gradient Method
Constantine Caramanis · Dimitris Fotakis · Alkis Kalavasis · Vasilis Kontonis · Christos Tzamos
Suggesting Variable Order for Cylindrical Algebraic Decomposition via Reinforcement Learning
Fuqi Jia · Yuhang Dong · Minghao Liu · Pei Huang · Feifei Ma · Jian Zhang
Combinatorial Optimization with Policy Adaptation using Latent Space Search
Felix Chalumeau · Shikha Surana · Clément Bonnet · Nathan Grinsztajn · Arnu Pretorius · Alexandre Laterre · Thomas D
Winner Takes It All: Training Performant RL Populations for Combinatorial Optimization
Nathan Grinsztajn · Daniel Furelos-Blanco · Shikha Surana · Clément Bonnet · Thomas D Barrett
Environment and Benchmark
MetaBox: A Benchmark Platform for Meta-Black-Box Optimization with Reinforcement Learning
Zeyuan Ma · Hongshu Guo · Jiacheng Chen · Zhenrui Li · Guojun Peng · Yue-Jiao Gong · Yining Ma · Zhiguang Cao
Katakomba: Tools and Benchmarks for Data-Driven NetHack
Vladislav Kurenkov · Alexander Nikulin · Denis Tarasov · Sergey Kolesnikov
Pgx: Hardware-accelerated Parallel Game Simulators for Reinforcement Learning
Sotetsu Koyamada · Shinri Okano · Soichiro Nishimori · Yu Murata · Keigo Habara · Haruka Kita · Shin Ishii
RL-ViGen: A Reinforcement Learning Benchmark for Visual Generalization
Zhecheng Yuan · Sizhe Yang · Pu Hua · Can Chang · Kaizhe Hu · Xiaolong Wang · Huazhe Xu
Bridging RL Theory and Practice with the Effective Horizon
Cassidy Laidlaw · Stuart J Russell · Anca Dragan
Classical Simulation of Quantum Circuits: Parallel Environments and Benchmark
Xiao-Yang Liu · Zeliang Zhang
ScenarioNet: Open-Source Platform for Large-Scale Traffic Scenario Simulation and Modeling

Quanyi Li · Zhenghao Peng · Lan Feng · Zhizheng Liu · Chenda Duan · Wenjie Mo · Bolei Zhou
DynaDojo: An Extensible Platform for Benchmarking Sample Efficiency in Dynamical System Identification
Logan Bhamidipaty · Tommy Bruzzese · Rami Ratl Mrad · Caryn Tran · Maxinder S. Kanwal
SustainGym: Reinforcement Learning Environments for Sustainable Energy Systems
Christopher Yeh · Victor Li · Rajeev Datta · Julio Arroyo · Nicolas Christianson · Chi Zhang · Yize Chen · Mohammad Mehdi Hosseini · Azarang Golmohammadi · Yuanyuan Shi · Yisong Yue · Adam Wierman
KuaiSim: A Comprehensive Simulator for Recommender Systems
Kesen Zhao · Shuchang Liu · Qingpeng Cai · Xiangyu Zhao · Ziru Liu · Dong Zheng · Peng Jiang · Kun Gai
DriveMax: An Accelerated, Data-Driven Simulator for Large-Scale Autonomous Driving Research
Cole Gulino · Justin Fu · Wenjie Luo · George Tucker · Eli Bronstein · Yiren Lu · Jean Harb · Xinlei Pan · Yan Wang · Xiangyu Chen · John Co-Reyes · Rishabh Agarwal · Rebecca Roelofs · Yao Lu · Nico Montali · Paul Mougin · Zoey Yang · Brandyn White · Aleksandra Faust · Rowan McAllister · Dragomir Anguelov · Benjamin Sapp
TradeMaster: A Holistic Quantitative Trading Platform Empowered by Reinforcement Learning
Shuo Sun · Molei Qin · Wentao Zhang · Haochong Xia · Chuqiao Zong · Jie Ying · Yonggang Xie · Lingxuan Zhao · Xinrun Wang · Bo An
Minigrid & Miniworld: Modular & Customizable Reinforcement Learning Environments for Goal-Oriented Tasks
Maxime Chevalier-Boisvert · Bolun Dai · Mark Towers · Rodrigo Perez-Vicente · Lucas Willems · Salem Lahlou · Suman Pal · Pablo Samuel Castro · J Terry
Imitation Learning/Reward-free/Inverse
Understanding Expertise through Demonstrations: A Maximum Likelihood Framework for Offline Inverse Reinforcement Learning
Siliang Zeng · Chenliang Li · Alfredo Garcia · Mingyi Hong
RoboCLIP: One Demonstration is Enough to Learn Robot Policies
Sumedh Sontakke · Séb Arnold · Jesse Zhang · Karl Pertsch · Erdem Bıyık · Dorsa Sadigh · Chelsea Finn · Laurent Itti
Inverse Reinforcement Learning with the Average Reward Criterion
Feiyang Wu · Jingyang Ke · Anqi Wu
Action Inference by Maximising Evidence: Zero-Shot Imitation from Observation with World Models
Xingyuan Zhang · Philip Becker-Ehmck · Patrick van der Smagt · Maximilian Karl
Social Motion Prediction with Cognitive Hierarchies
Wentao Zhu · Jason Qin · Yuke Lou · Hang Ye · Xiaoxuan Ma · Hai Ci · Yizhou Wang
Inverse Preference Learning: Preference-based RL without a Reward Function
Joey Hejna · Dorsa Sadigh
Coherent Soft Imitation Learning
Joe Watson · Sandy Huang · Nicolas Heess
NetHack is Hard to Hack
Ulyana Piterbarg · Lerrel Pinto · Rob Fergus
Hybrid Policy Optimization from Imperfect Demonstrations
Hanlin Yang · Chao Yu · peng sun · Siji Chen
Control and Planning
Ensemble-based Deep Reinforcement Learning for Vehicle Routing Problems under Distribution Shift
Yuan Jiang · Zhiguang Cao · Yaoxin Wu · Wen Song · Jie Zhang
Managing Temporal Resolution in Continuous Value Estimation: A Fundamental Trade-off
Zichen Zhang · Johannes Kirschner · Junxi Zhang · Francesco Zanini · Alex Ayoub · Masood Dehghan · Dale Schuurmans
Policy Optimization in a Noisy Neighborhood: On Return Landscapes in Continuous Control
Nate Rahn · Pierluca D'Oro · Harley Wiltzer · Pierre-Luc Bacon · Marc Bellemare
Active Reinforcement Learning under Limited Visual Observability
Jinghuan Shang · Michael Ryoo
PID-Inspired Inductive Biases for Deep Reinforcement Learning in Partially Observable Control Tasks
Ian Char · Jeff Schneider
Decision Stacks: Flexible Reinforcement Learning via Modular Generative Models
Siyan Zhao · Aditya Grover

EDGI: Equivariant Diffusion for Planning with Embodied Agents

Johann Brehmer · Joey Bose · Pim de Haan · Taco Cohen

Thinker: Learning to Plan and Act

Stephen Chung · Ivan Anokhin · David Krueger

Self-Predictive Universal AI

Elliot Catt · Jordi Grau-Moya · Marcus Hutter · Matthew Aitchison · Tim Genewein · Gregoire Deletang · Kevin Li · Joel Veness

Diffusion Model

Reward-Directed Conditional Diffusion: Provable Distribution Estimation and Reward Improvement

Hui Yuan · Kaixuan Huang · Chengzhuo Ni · Minshuo Chen · Mengdi Wang

Reinforcement Learning for Fine-tuning Text-to-Image Diffusion Models

Ying Fan · Olivia Watkins · Yuqing Du · Hao Liu · Moonkyung Ryu · Craig Boutilier · Pieter Abbeel · Mohammad Ghavamzadeh · Kangwook Lee · Kimin Lee

Extracting Reward Functions from Diffusion Models

Felipe Nuti · Tim Franzmeyer · João Henriques

Generating Behaviorally Diverse Policies with Latent Diffusion Models

Shashank Hegde · Sumeet Batra · K.R. Zentner · Gaurav Sukhatme

Diffusion Model is an Effective Planner and Data Synthesizer for Multi-Task Reinforcement Learning

Haoran He · Chenjia Bai · Kang Xu · Zhuoran Yang · Weinan Zhang · Dong Wang · Bin Zhao · Xuelong Li

Adaptive Online Replanning with Diffusion Models

Siyuan Zhou · Yilun Du · Shun Zhang · Mengdi Xu · Yikang Shen · Wei Xiao · Dit-Yan Yeung · Chuang Gan

Reward shaping

Behavior Alignment via Reward Function Optimization

Dhawal Gupta · Yash Chandak · Scott Jordan · Philip Thomas · Bruno da Silva

GRD: A Generative Approach for Interpretable Reward Redistribution in Reinforcement Learning

Yudi Zhang · Yali Du · Biwei Huang · Ziyan Wang · Jun Wang · Meng Fang · Mykola Pechenizkiy

Video Prediction Models as Rewards for Reinforcement Learning

Alejandro Escontrela · Ademi Adeniji · Wilson Yan · Ajay Jain · Xue Bin Peng · Ken Goldberg · Youngwoon Lee · Danijar Hafner · Pieter Abbeel

A Long N-step Surrogate Stage Reward for Deep Reinforcement Learning

Junmin Zhong · Ruofan Wu · Jennie Si

ODPP: A Unified Algorithm Framework for Unsupervised Option Discovery based on Determinantal Point Process

Jiayu Chen · Vaneet Aggarwal · Tian Lan

Unsupervised Learning for Solving the Travelling Salesman Problem

Yimeng Min · Yiwei Bai · Carla Gomes

Hierarchical RL

Effectively Learning Initiation Sets in Hierarchical Reinforcement Learning

Akhil Bagaria · Ben Abbatematteo · Omer Gottesman · Matt Corsaro · Sreehari Rammohan · George Konidaris

Sample Complexity of Goal-Conditioned Hierarchical Reinforcement Learning

Arnaud Robert · Ciara Pike-Burke · Aldo Faisal

Natural Language-conditioned Reinforcement Learning with Task-related Language Development and Translation

Jingcheng Pang · Xin-Yu Yang · Si-Hang Yang · Xiong-Hui Chen · Yang Yu

Discovering Hierarchical Achievements in Reinforcement Learning via Contrastive Learning

Seungyong Moon · Junyoung Yeom · Bumsoo Park · Hyun Oh Song

Creating Multi-Level Skill Hierarchies in Reinforcement Learning

Joshua B. Evans · Özgür Şimşek

Online RL

Train Once, Get a Family: State-Adaptive Balances for Offline-to-Online Reinforcement Learning
Shenzhi Wang · Qisen Yang · Jiawei Gao · Matthieu Lin · HAO CHEN · Liwei Wu · Ning Jia · Shiji Song · Gao Huang
Online RL in Linearly q^π-Realizable MDPs Is as Easy as in Linear MDPs If You Learn What to Ignore
Gellért Weisz · András György · Csaba Szepesvari
When is Agnostic Reinforcement Learning Statistically Tractable?
Zeyu Jia · Gene Li · Alexander Rakhlin · Ayush Sekhari · Nati Srebro
No-Regret Online Reinforcement Learning with Adversarial Losses and Transitions
William Chang · Tiancheng Jin · Junyan Liu · Haipeng Luo · Chloé Rouyer · Chen-Yu Wei
Optimistic Natural Policy Gradient: a Simple Efficient Policy Optimization Framework for Online RL
Qinghua Liu · Gellért Weisz · András György · Chi Jin · Csaba Szepesvari

Pre-train

Pre-training Contextualized World Models with In-the-wild Videos for Reinforcement Learning
Jialong Wu · Haoyu Ma · Chaoyi Deng · Mingsheng Long
Parameter and Computation Efficient Transfer Learning for Vision-Language Pre-trained Models
Qiong Wu · Wei Yu · Yiyi Zhou · Shubin Huang · Xiaoshuai Sun · Rongrong Ji
Learning to Modulate pre-trained Models in RL
Thomas Schmied · Markus Hofmarcher · Fabian Paischer · Razvan Pascanu · Sepp Hochreiter
Can Pre-Trained Text-to-Image Models Generate Visual Goals for Reinforcement Learning?
Danika Gao · Kaizhe Hu · Guowei Xu · Huazhe Xu
BIRD: Generalizable Backdoor Detection and Removal for Deep Reinforcement Learning
Xuan Chen · Wenbo Guo · Guanhong Tao · Xiangyu Zhang · Dawn Song

Applications of RL

Task-driven metric learning for end-to-end model learning
Dishank Bansal · Ricky T. Q. Chen · Mustafa Mukadam · Brandon Amos
De novo Drug Design using Reinforcement Learning with Multiple GPT Agents
Xiuyuan Hu · Hao Zhang · Yang Zhao
Macro Placement by Wire-Mask-Guided Black-Box Optimization
Yunqi Shi · Ke Xue · Song Lei · Chao Qian

Continual RL

Prediction and Control in Continual Reinforcement Learning
Nishanth Anand · Doina Precup
A Definition of Continual Reinforcement Learning
David Abel · Andre Barreto · Benjamin Van Roy · Doina Precup · Hado van Hasselt · Satinder Singh
Rewiring Neurons in Non-Stationary Environments
Zhicheng Sun · Yadong Mu

Model-free RL

Efficient Model-Free Exploration in Low-Rank MDPs
Zak Mhammedi · Adam Block · Dylan J Foster · Alexander Rakhlin
Online Nonstochastic Model-Free Reinforcement Learning
Udaya Ghai · Arushi Gupta · Karan Singh · Wenhan Xia · Elad Hazan
Model-Free Reinforcement Learning with the Decision-Estimation Coefficient
Dylan J Foster · Noah Golowich · Jian Qian · Alexander Rakhlin · Ayush Sekhari

POMDP

Learning Generalizable Agents via Saliency-guided Features Decorrelation

Sili Huang · Yanchao Sun · Jifeng Hu · Siyuan Guo · Bo Yang · Hechang Chen · Yi Chang · Lichao Sun
ODE-based Recurrent Model-free Reinforcement Learning for POMDPs
Xuanle Zhao · Duzhen Zhang · Han Liyuan · Tielin Zhang · Bo Xu
Others
CQM: Curriculum Reinforcement Learning with a Quantized World Model
Seungjae Lee · Daesol Cho · Jonghae Park · H. Jin Kim
Waypoint Transformer: Reinforcement Learning via Supervised Learning with Intermediate Targets
Anirudhan Badrinath · Yannis Flet-Berliac · Allen Nie · Emma Brunskill
COOM: A Game Benchmark for Continual Reinforcement Learning
Tristan Tomilin · Meng Fang · Yudi Zhang · Mykola Pechenizkiy
Interpretable and Explainable Logical Policies via Neurally Guided Symbolic Abstraction
Quentin Delfosse · Hikaru Shindo · Devendra Dhami · Kristian Kersting
MAG-GNN: Reinforcement Learning Boosted Graph Neural Network
Lecheng Kong · Jiarui Feng · Hao Liu · Dacheng Tao · Yixin Chen · Muhan Zhang
Lending Interaction Wings to Recommender Systems with Conversational Agents
Jiarui Jin · Xianyu Chen · Fanghua Ye · Mengyue Yang · Yue Feng · Weinan Zhang · Yong Yu · Jun Wang
Koopman Kernel Regression
Petar Bevanda · Max Beier · Armin Lederer · Stefan Sosnowski · Eyke Hüllermeier · Sandra Hirche
Dynamics Generalisation via Adaptive Policies
Michael Beukman · Devon Jarvis · Richard Klein · Steven James · Benjamin Rosman
Replicable Reinforcement Learning
Eric Eaton · Marcel Hussing · Michael Kearns · Jessica Sorrell
SEENN: Towards Temporal Spiking Early Exit Neural Networks
Yuhang Li · Tamar Geller · Youngeun Kim · Priyadarshini Panda
Hierarchical Adaptive Value Estimation for Multi-modal Visual Reinforcement Learning
Yangru Huang · Peixi Peng · Yifan Zhao · Haoran Xu · Mengyue Geng · Yonghong Tian