Novelty Assessment Report

Paper: Online Decision Making with Generative Action Sets

PDF URL: https://openreview.net/pdf?id=J8hn5LzNOE

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

With advances in generative AI, decision-making agents can now dynamically create new actions during online learning, but action generation typically incurs costs that must be balanced against potential benefits. We study an online learning problem where an agent can generate new actions at any time step by paying a one-time cost, with these actions becoming permanently available for future use. The challenge lies in learning the optimal sequence of two-fold decisions: which action to take and when to generate new ones, further complicated by the triangular tradeoffs among exploitation, exploration and creation. To solve this problem, we propose a doubly-optimistic algorithm that employs Lower Confidence Bounds (LCB) for action selection and Upper Confidence Bounds (UCB) for action generation. Empirical evaluation on healthcare question-answering datasets demonstrates that our approach achieves favorable generation-quality trade-offs compared to baseline strategies. From theoretical perspectives, we prove that our algorithm achieves the optimal regret of $O(T^{\frac{d}{d+2}}d^{\frac{d}{d+2}} + d\sqrt{T\log T})$, providing the first sublinear regret bound for online learning with expanding action spaces.

Core Task Landscape

This paper addresses: **online learning with dynamically expanding action spaces**

A total of **50 papers** were analyzed and organized into a taxonomy with **22 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Action Space Expansion Mechanisms and Theoretical Foundations**
- **Composite and Structured Action Space Decomposition**
- **Transfer Learning and Model Reuse Across Action Space Changes**
- **Constrained and Masked Action Space Methods**
- **Application-Driven Online Learning with Dynamic Action Spaces**
- **Time-Varying System Modeling and Adaptive Learning**
- **Optimization and Search Space Expansion Methods**

Complete Taxonomy Tree

- online learning with dynamically expanding action spaces Survey Taxonomy
- Action Space Expansion Mechanisms and Theoretical Foundations
  - Curriculum-Based and Progressive Action Space Growth ★ (2 papers)
  - [0] Online Decision Making with Generative Action Sets (Anon et al., 2026) View paper
  - [10] Growing action spaces (Gregory Farquhar, 2020) View paper
  - Lifelong and Continual Learning with Changing Action Sets (3 papers)
  - [21] Lifelong Learning with a Changing Action Set (Yash Chandak, 2019) View paper
  - [30] Online Continual Learning via Dynamic Expandable Recursive Model (Fei Ye, 2025) View paper
  - [36] Action-Adaptive Continual Learning: Enabling Policy Generalization under Dynamic Action Spaces (Pan, 2025) View paper
  - Adaptive Resolution and Discretization Strategies (2 papers)
  - [28] Growing Q-Networks: Solving Continuous Control Tasks with Adaptive Control Resolution (Seyde, 2024) View paper
  - [43] Variable resolution dynamic programming: Efficiently learning action maps in multivariate real-valued state-spaces (Moore, 1991) View paper
- Composite and Structured Action Space Decomposition
  - Action Branching and Hierarchical Decomposition (2 papers)
  - [2] A Sequential Decision Algorithm of Reinforcement Learning for Composite Action Space (Yuan. Gao, 2023) View paper
  - [9] Efficient dynamic spectrum anti-jamming access with large action space: An action space decomposition-based approach (Hao Han, 2024) View paper
  - Time-Varying Composite Action Spaces with Cooperation (2 papers)
  - [23] Structured Cooperative Reinforcement Learning With Time-Varying Composite Action Space (Wenhao Li, 2021) View paper
  - [27] SMAUG: A Sliding Multidimensional Task Window-Based MARL Framework for Adaptive Real-Time Subtask Recognition (Zhang Wen-jing, 2024) View paper
- Transfer Learning and Model Reuse Across Action Space Changes (2 papers)
  - [3] Transfer Learning in Deep Reinforcement Learning: Actor-Critic Model Reuse for Changed State-Action Space (Eric Veith, 2025) View paper
  - [46] Adapting the Behavior of Reinforcement Learning Agents to Changing Action Spaces and Reward Functions (RaÃºl de la Rosa, 2025) View paper

- Constrained and Masked Action Space Methods
  - Dynamic Action Masking and Feasibility Constraints (2 papers)
  - [24] Energy-efficient underwater acoustic communication based on Dyna-Q with an adaptive action space (Cheng Fan, 2023) View paper
  - [38] Magnetic Field Compensation Control for Spin-Exchange Relaxation-Free Comagnetometer Using Reinforcement Learning (Feng Li, 2023) View paper
  - Interval Restrictions and Collision Avoidance (2 papers)
  - [8] Automatic ship collision avoidance using deep reinforcement learning with LSTM in continuous action spaces (Ryohei Sawada, 2020) View paper
  - [39] Dynamic interval restrictions on action spaces in deep reinforcement learning for obstacle avoidance (Grams, 2023) View paper
- Application-Driven Online Learning with Dynamic Action Spaces
  - Multi-Agent Coordination and Network Optimization (4 papers)
  - [1] Multi-agent reinforcement learning based dynamic self-coordinated topology optimization for wireless mesh networks (Qingwei Tang, 2025) View paper
  - [4] Neural networks-based iterative learning control consensus for periodically time-varying multi-agent systems (Jiaxi Chen, 2024) View paper
  - [26] Fleet rebalancing for expanding shared e-mobility systems: A multi-agent deep reinforcement learning approach (Man Luo, 2023) View paper
  - [35] Cooperative Path Planning Method for Enhancing Ground-Units Survivability Based on Adaptive Q-Learning (Miao Guo, 2023) View paper
  - Manufacturing and Production Scheduling (4 papers)
  - [7] An efficient and adaptive design of reinforcement learning environment to solve job shop scheduling problem with soft actor-critic algorithm (Jinghua Si, 2024) View paper
  - [11] Deep Reinforcement Learning-Based Dynamic Reconfiguration Planning for Digital Twin-Driven Smart Manufacturing Systems With Reconfigurable Machine Tools (Jintang Huang, 2024) View paper
  - [32] Real-Time Scheduling Based on Simulation and Deep Reinforcement Learning with Featured Action Space (Shufang Xie, 2022) View paper
  - [42] Deep reinforcement learning for online combinatorial resource allocation with arbitrary state and action spaces (Zemuy, 2024)
  - Autonomous Systems and Robotics (5 papers)
  - [5] Flingbot: The unreasonable effectiveness of dynamic manipulation for cloth unfolding (Huy Ha, 2022) View paper
  - [29] Towards general single-utensil food acquisition with human-informed actions (EK Gordon, 2023) View paper
  - [40] Amortized Q-learning with Model-based Action Proposals for Autonomous Driving on Highways (Mirchevska, 2020) View paper
  - [45] Learning Robotic Manipulation Skills Using an Adaptive Force-Impedance Action Space (Ulmer, 2021) View paper
  - [50] Adaptive Space Expansion for Fast Motion Planning (Shenglei Shi, 2024) View paper
  - Communication and Network Resource Allocation (5 papers)
  - [15] Adaptive urban traffic signal control based on enhanced deep reinforcement learning (Min Wei, 2024) View paper
  - [17] Deep Reinforcement Learning with Coalition Action Selection for Online Combinatorial Resource Allocation with Arbitrary Action Space (Gebrekidan Tesfay Zemuy, 2024) View paper
  - [22] Deep reinforcement learning-based adaptive modulation for OFDM underwater acoustic communication system (Xuerong Cui, 2023) View paper
  - [25] Deep Reinforcement Learning With Communication Transformer for Adaptive Live Streaming in Wireless Edge Networks (Shuoyao Wang, 2022) View paper
  - [34] Online Learning Based Efficient Resource Allocation for LoRaWAN Network (Wang, 2025) View paper
  - Energy Management and Control Systems (2 papers)
  - [18] A deep reinforcement learning-based pid tuning strategy for nonlinear mimo systems with time-varying uncertainty (Hao Wang, 2024) View paper
  - [20] An Online Energy Management Strategy for Fuel Cell Vehicles Based on Fuzzy Q-Learning and Road Condition Recognition (Duo Yang, 2024) View paper
  - Cybersecurity and Network Defense (1 papers)
  - [33] How to Disturb Network Reconnaissance: A Moving Target Defense Approach Based on Deep Reinforcement Learning (Tao Zhang, 2023) View paper
  - Transportation and Route Planning (2 papers)
  - [41] Flexible Route Network Planning Adapting to Time-varying Air Traffic Using Reinforcement Learning (Yangjie Li, 2025) View paper
  - [48] A reinforcement learning framework for train rescheduling (Qi Shi, 2022) View paper
  - Educational and Personalized Learning Systems (1 papers)
  - [6] Towards Sustainable Learning in Online Education: A Reinforcement Learning Approach (C Zhai, 2025) View paper
- Time-Varying System Modeling and Adaptive Learning
  - Online Graph and Network Topology Learning (1 papers)
  - [12] Online Learning of Expanding Graphs (Samuel Rey, 2025) View paper
  - Adaptive Latent Space and Representation Learning (3 papers)
  - [14] Adaptive autoencoder latent space tuning for more robust machine learning beyond the training set for six-dimensional phase space diagnostics of a time-varying ultrafast electron-diffraction compact accelerator. (Alexander Scheinker, 2023) View paper
  - [31] Online Learning of a Probabilistic and Adaptive Scene Representation (Zike Yan, 2021) View paper
  - [37] An adaptive variable-parameter dynamic learning network for solving constrained time-varying QP problem. (Zhijun Zhang, 2025) View paper
  - Time-Varying Dynamics and Anomaly Detection (2 papers)
  - [13] Unified Flowing Normality Learning for Rotating Machinery Anomaly Detection in Continuous Time-Varying Conditions (Chenye Hu, 2024) View paper
  - [19] Learning Time-Varying Multi-Region Brain Communications via Scalable Markovian Gaussian Processes (Li Weihan, 2024) View paper
  - Structured Motion and Temporal Dynamics Representation (1 papers)
  - [16] FLD: Fourier Latent Dynamics for Structured Motion Representation and Learning (Li Chenhao, 2024) View paper
  - Distributed and Clustered Adaptive Learning (1 papers)

  ◦ [49] Distributed adaptive clustering learning over time-varying multitask networks (Qing Shi, 2021) View paper
- Optimization and Search Space Expansion Methods (2 papers)
  ◦ [44] A PSO-based multi-objective multi-label feature selection method in classification (Zhang Yon, 2017) View paper
  ◦ [47] Adaptive Expansion Bayesian Optimization for Unbounded Global Optimization (Wei Chen, 2020) View paper

Narrative

Core task: online learning with dynamically expanding action spaces. This field addresses scenarios where an agent must learn effective policies even as the set of available actions grows or changes over time. The taxonomy organizes research into several main branches: theoretical foundations and expansion mechanisms that formalize how action spaces evolve; composite and structured decomposition methods that break large or complex action sets into manageable subspaces; transfer learning and model reuse strategies that leverage prior knowledge when new actions appear; constrained and masked approaches that selectively enable or disable actions; application-driven studies spanning domains such as robotics, scheduling, and network control; time-varying system modeling for environments with inherent temporal dynamics; and optimization techniques that expand search spaces adaptively. Representative works like Growing Action Spaces[10] and Changing Action Set[21] illustrate early efforts to handle action-set variability, while more recent studies such as Actor-Critic Reuse[3] and Action-Adaptive Continual[36] explore how to efficiently transfer learned components across evolving action configurations.

A particularly active line of work focuses on curriculum-based and progressive growth strategies, where action spaces expand gradually to facilitate learning. Generative Action Sets[0] sits within this branch, emphasizing mechanisms that generate or reveal new actions in a structured manner rather than presenting the full action space at once. This contrasts with methods like Growing Q-Networks[28] and Adaptive Action Space[24], which dynamically adjust network architectures or action representations in response to observed task demands. Meanwhile, composite decomposition approaches such as Composite Action Space[2] tackle the combinatorial challenge of large action sets by factoring them into smaller components, and application-driven studies like Flingbot[5] demonstrate how domain-specific constraints shape expansion policies in robotic manipulation. The interplay between these directions highlights a central trade-off: whether to expand action spaces proactively via curriculum design or reactively based on environmental feedback, and how to balance exploration of new actions against exploitation of known strategies.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Growing action spaces

Authors: Gregory Farquhar, Laura Gustafson, Zeming Lin, Shimon Whiteson, Nicolas Usunier, et al. (6 authors total) | **Year/Venue**: 2020 | **URL**: View paper

Abstract

In complex tasks, such as those with large combinatorial action spaces, random exploration may be too inefficient to achieve meaningful learning progress. In this work, we use a curriculum of progressively growing action spaces to accelerate learning. We assume the environment is out of our control, but that the agent may set an internal curriculum by initially restricting its action space. Our approach uses off-policy reinforcement learning to estimate optimal value functions for multiple actio...

Relationship Analysis

Both papers belong to the curriculum-based and progressive action space growth category, using incremental expansion strategies to improve exploration efficiency. They overlap in addressing online learning with dynamically expanding action spaces through progressive curricula, where the original paper focuses on generative action creation with cost-benefit tradeoffs using a doubly-optimistic algorithm (LCB/UCB), while the candidate paper grows action spaces through hierarchical clustering and discretization refinement, particularly for multi-agent control scenarios. The key difference is that the original paper emphasizes costly generation decisions and create-to-reuse dynamics with theoretical regret bounds, whereas the candidate paper focuses on hierarchical action space restrictions without explicit generation costs, using off-action-space learning to transfer knowledge across restriction levels.

Contributions Analysis

Overall novelty summary. The paper introduces a create-to-reuse framework where an agent dynamically generates new actions during online learning by paying one-time costs, balancing exploitation, exploration, and creation. It resides in the Curriculum-Based and Progressive Action Space Growth leaf, which contains only two papers including this one. This leaf sits under Action Space Expansion Mechanisms and Theoretical Foundations, indicating a relatively sparse research direction focused on structured, incremental action space growth rather than reactive or application-driven expansion strategies.

The taxonomy reveals neighboring branches addressing related but distinct challenges: Lifelong and Continual Learning with Changing Action Sets handles catastrophic forgetting across evolving action spaces, while Adaptive Resolution and Discretization Strategies adjust granularity rather than generating entirely new actions. The paper's focus on cost-aware action generation distinguishes it from these directions, which either assume costless expansion or fixed discretization schemes. The broader Action Space Expansion Mechanisms branch emphasizes theoretical regret bounds and algorithmic mechanisms, positioning this work within foundational rather than application-specific research.

Among 29 candidates examined across three contributions, none were found to clearly refute the proposed approach. The create-to-reuse formulation examined 10 candidates with no refutable overlap, the doubly-optimistic algorithm examined 10 candidates with no refutable overlap, and the optimal regret bound examined 9 candidates with no refutable overlap. This limited search scope suggests that within the top-30 semantically similar papers, the specific combination of cost-aware action generation, doubly-optimistic confidence bounds, and sublinear regret guarantees appears underexplored, though the analysis does not cover the full literature landscape.

Based on the limited search scope of 29 candidates, the work appears to occupy a relatively novel position within its immediate research neighborhood. The sparse population of its taxonomy leaf and absence of refutable prior work among examined candidates suggest potential originality, though a more exhaustive literature review would be needed to confirm whether similar cost-aware generation frameworks exist in adjacent domains or under different terminologies.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Create-to-reuse problem formulation with expanding action spaces

Description: The authors introduce a novel online learning framework where agents can dynamically generate new actions at a fixed one-time cost, with generated actions becoming permanently available for future reuse. This formulation captures triangular tradeoffs among exploitation, exploration, and creation, distinguishing it from traditional fixed-action-space settings.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. CFWS: DRL-Based Framework for Energy Cost and Carbon Footprint Optimization in Cloud Data Centers

URL: View paper

Brief Assessment

CFWS[69] focuses on VM migration in cloud data centers with a flattened index action space for resource allocation, not on online learning with dynamically expanding action spaces where agents generate new actions at creation costs.

2. Growing Q-Networks: Solving Continuous Control Tasks with Adaptive Control Resolution

URL: View paper

Brief Assessment

Growing Q-Networks[28] focuses on discretizing continuous control spaces with adaptive resolution for robotics applications, not on online learning frameworks where agents generate new actions at fixed costs for future reuse. The candidate addresses control resolution refinement, while the original paper models strategic action creation with one-time generation costs and permanent availability.

3. Learning Robotic Manipulation Skills Using an Adaptive Force-Impedance Action Space

URL: View paper

Brief Assessment

Force-Impedance Action[45] addresses robotic manipulation with a fixed hierarchical action space combining RL and adaptive control. It does not involve dynamically generating new actions at runtime or expanding action spaces through creation decisions.

4. Reinforced Imitation in Heterogeneous Action Space

URL: View paper

Brief Assessment

Reinforced Heterogeneous Imitation[76] addresses imitation learning where expert and learner have different action spaces, but does not involve dynamically generating new actions at runtime with creation costs. The candidate focuses on learning from state-only expert observations with heterogeneous actions, not on expanding action spaces through costly generation.

5. Parameterized Deep Reinforcement Learning With Hybrid Action Space for Edge Task Offloading

URL: View paper

Brief Assessment

Hybrid Action Space[75] addresses discrete-continuous hybrid action spaces in edge task offloading, not the create-to-reuse framework with one-time creation costs and permanent action availability. The candidate focuses on optimizing offloading decisions within a predefined hybrid action space, whereas the original contribution involves dynamically generating new actions that become reusable assets.

6. A Deep Reinforcement Learning Approach for Multi-UAV Collaborative Coverage with Adaptive Step Size and Dynamic Reward Mechanism

URL: View paper

Brief Assessment

Multi-UAV Coverage[72] focuses on adaptive step size control in UAV path planning with a fixed action space that expands through step size variation, not on dynamically generating and permanently reusing new actions at a one-time cost as in the original paper's create-to-reuse framework.

7. AGFT: An Adaptive GPU Frequency Tuner for Real-Time LLM Inference Optimization

URL: View paper

Brief Assessment

AGFT[71] focuses on GPU frequency tuning for LLM inference using reinforcement learning with action space pruning. This is fundamentally different from the original paper's create-to-reuse framework where actions are generated at a cost and permanently added to an expanding library.

8. Generating learning sequences for decision makers through data mining and competence set expansion

URL: View paper

Brief Assessment

Competence Set Expansion[73] focuses on generating learning sequences for acquiring knowledge patterns from databases to help decision makers, not on online learning frameworks with dynamically expanding action spaces and creation costs. The candidate addresses knowledge acquisition sequencing in decision support systems, while the original paper studies sequential action generation with cost-benefit tradeoffs in online learning settings.

9. An adaptive learning-based approach for nearly optimal dynamic charging of electric vehicle fleets

URL: View paper

Brief Assessment

Dynamic Charging Fleet[70] addresses electric vehicle charging optimization with fixed action spaces (charging schedules). It does not involve dynamically generating new actions at runtime or expanding action spaces through creation costs, which is the core novelty of the original paper's create-to-reuse framework.

10. Deep Reinforcement Learning-Graph Neural Networks-Dynamic Clustering triplet for Adaptive Multi Energy Microgrid optimization

URL: View paper

Brief Assessment

Dynamic Clustering Triplet[74] focuses on microgrid energy optimization using DRL with graph neural networks for spatial relationships in energy systems. It does not address online learning with dynamically expanding action spaces through costly action generation, which is the core novelty of the original paper's create-to-reuse framework.

Contribution 2: Doubly-optimistic algorithm using LCB and UCB

Description: The authors develop an algorithm that employs LCB for action selection to balance exploitation and exploration, while using UCB-based probabilistic decisions for action generation. This double optimism principle enables the algorithm to maximize long-term value of new actions while controlling worst-case regret.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Fine-tuning offline policies with optimistic action selection

URL: View paper

Brief Assessment

Optimistic Action Selection[51] focuses on offline-to-online RL fine-tuning using optimistic action selection during environment interaction, not on online learning with expanding action spaces where actions can be generated at cost.

---

2. TRUST-ME: Trust-Based Resource Allocation and Server Selection in Multi-Access Edge Computing

URL: View paper

Brief Assessment

TRUST-ME[60] focuses on trust-based resource allocation in multi-access edge computing using optimistic Q-learning with UCB for server selection, not on online learning with expanding action spaces or the dual use of LCB/UCB for balancing creation and exploitation decisions.

---

3. Dyna-T: Dyna-Q and Upper Confidence Bounds Applied to Trees

URL: View paper

Brief Assessment

Dyna-T[57] focuses on combining Dyna-Q with UCT for model-based RL in fixed action spaces, not on online learning with expanding action sets or the dual use of LCB/UCB for action selection versus generation decisions.

---

4. Why so pessimistic? estimating uncertainties for offline rl through ensembles, and why their independence matters

URL: View paper

Brief Assessment

Pessimistic Ensembles[56] focuses on offline RL with pessimistic value estimation using LCB for Q-function ensembles, not on online learning with action generation. The candidate uses LCB to select among existing actions in a conservative manner, whereas the original paper uses LCB for action selection to balance exploitation/exploration and UCB for probabilistic action generation decisions in an expanding action space setting.

---

5. Wasserstein actor-critic: directed exploration via optimism for continuous-actions control

URL: View paper

Brief Assessment

Wasserstein Actor-Critic[55] uses LCB/UCB for continuous-action RL with uncertainty propagation via Wasserstein barycenters, not for online learning with expanding action spaces. The technical contexts differ fundamentally: WAC addresses exploration in fixed continuous action spaces, while the original paper addresses dynamic action generation with creation costs.

---

6. Principled Exploration via Optimistic Bootstrapping and Backward Induction

URL: View paper

Brief Assessment

Optimistic Bootstrapping[59] focuses on exploration in deep RL using UCB-bonus through bootstrapping for epistemic uncertainty estimation, not on action generation decisions with creation costs in expanding action spaces.

---

7. Uncertainty Quantification and Exploration for Reinforcement Learning

URL: View paper

Brief Assessment

Uncertainty Quantification Exploration[58] focuses on statistical uncertainty quantification and exploration policies for reinforcement learning in MDPs, not on online decision-making with generative action sets. The candidate uses confidence bounds for Q-value estimation and exploration policy design, while the original paper addresses action generation decisions in expanding action spaces.

---

8. Better exploration with optimistic actor critic

URL: View paper

Brief Assessment

Optimistic Actor Critic[54] uses LCB and UCB for different purposes in actor-critic RL (exploration policy vs. target policy updates), not for balancing action selection and generation decisions in expanding action spaces as in the original paper's create-to-reuse framework.

---

9. Information-theoretic confidence bounds for reinforcement learning

URL: View paper

Brief Assessment

Information-Theoretic Confidence[52] focuses on information-theoretic confidence bounds for bandits and MDPs, using UCB for action selection in a different context. The candidate does not address the specific create-to-reuse problem with expanding action spaces that the original paper tackles.

---

10. Learning the optimal control for evolving systems with converging dynamics

URL: View paper

Brief Assessment

Evolving Systems Control[53] uses optimistic-pessimistic bounds for a different problem (controlling systems with converging dynamics to stable states), not for action generation/selection in expanding action spaces. The technical mechanisms and problem settings are fundamentally different.

Contribution 3: Optimal regret bound with matching lower bound

Description: The authors prove their algorithm achieves expected regret of O(T^(d/(d+2)) d^(d/(d+2)) + d√(T log T)) where T is the time horizon and d is the covariate dimension. They establish this is optimal by proving a matching lower bound, providing the first sublinear regret guarantee for online learning with expanding action spaces.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Model-free online learning in unknown sequential decision making problems and games

URL: View paper

Brief Assessment

Model-Free Online Learning[62] addresses tree-form sequential decision making with unknown structure, achieving O(T^(3/4)) regret. The original paper studies online learning with expanding action spaces where actions can be generated at cost, a fundamentally different problem setting with different regret bounds O(T^(d/(d+2))).

2. Online Learning in Markov Decision Processes with Continuous Actions

URL: View paper

Brief Assessment

Continuous Actions MDP[68] addresses online learning in MDPs with continuous action spaces, which is fundamentally different from the original paper's create-to-reuse framework with expanding discrete action sets and generation costs.

3. Online Learning, Uniform Convergence, and a Theory of Interpretability

URL: View paper

Brief Assessment

Uniform Convergence Interpretability[66] focuses on batch learning theory (uniform convergence, fat-shattering dimension) and interpretable machine learning, not online learning with expanding action spaces or regret bounds for such settings.

4. Tight Regret Bounds for Infinite-armed Linear Contextual Bandits

URL: View paper

Brief Assessment

Infinite-Armed Contextual[61] addresses infinite-armed linear contextual bandits with fixed action sets, achieving O(√(d²T log T)) regret. The original paper studies a fundamentally different problem: online learning with dynamically expanding action spaces where new actions can be generated at cost c, achieving O(T^(d/(d+2)) d^(d/(d+2)) + d√(T log T)) regret with a matching Ω(T^(d/(d+2))) lower bound. These are distinct problem settings with different regret rates.

5. Adversarial Online Learning with Changing Action Sets: Efficient Algorithms with Approximate Regret Bounds

URL: View paper

Brief Assessment

Changing Action Sets[63] addresses online learning with sleeping experts where action availability varies, focusing on computational efficiency with approximate regret guarantees. The original paper studies a fundamentally different problem: online learning with expanding action spaces where new actions can be generated at a cost and permanently added to the library, requiring balancing exploitation, exploration, and creation decisions.

6. Online Learning Based Performance Optimization in Wireless Networks with Context Information

URL: View paper

Brief Assessment

Context Information Optimization[67] focuses on online antenna mode/beam selection in wireless networks using multi-armed bandit frameworks. The candidate does not address online learning with expanding action spaces or regret bounds for dynamically growing action sets, which is the core novelty claim of the original paper.

7. From External to Swap Regret 2.0: An Efficient Reduction for Large Action Spaces

URL: View paper

Brief Assessment

Swap Regret[64] addresses swap regret minimization in expert advice and game-theoretic settings, not online learning with expanding action spaces. The candidate's regret bounds apply to different problem structures (swap vs. external regret, fixed vs. dynamically generated actions).

8. Distributed No-Regret Learning for Multi-Stage Systems with End-to-End Bandit Feedback

URL: View paper

Brief Assessment

Multi-Stage Bandit[65] addresses multi-stage systems with distributed agents and end-to-end feedback, achieving O(T^(L/(L+1))) regret where L is the number of stages. The original paper studies online learning with expanding action spaces, achieving O(T^(d/(d+2))) regret where d is covariate dimension. These are fundamentally different problem settings with different structural parameters.

9. Online Learning of Expanding Graphs

URL: View paper

Brief Assessment

Expanding Graphs[12] addresses online network topology inference for growing graphs with spatiotemporal signals, not online decision-making with action generation costs. The regret analysis in Expanding Graphs[12] focuses on dynamic cumulative regret for graph structure learning, which is fundamentally different from the original paper's regret bounds for balancing exploitation, exploration, and creation decisions in action space expansion.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Online Decision Making with Generative Action Sets View paper
- [1] Multi-agent reinforcement learning based dynamic self-coordinated topology optimization for wireless mesh networks View paper
- [2] A Sequential Decision Algorithm of Reinforcement Learning for Composite Action Space View paper
- [3] Transfer Learning in Deep Reinforcement Learning: Actor-Critic Model Reuse for Changed State-Action Space View paper
- [4] Neural networks-based iterative learning control consensus for periodically time-varying multi-agent systems View paper
- [5] Flingbot: The unreasonable effectiveness of dynamic manipulation for cloth unfolding View paper
- [6] Towards Sustainable Learning in Online Education: A Reinforcement Learning Approach View paper
- [7] An efficient and adaptive design of reinforcement learning environment to solve job shop scheduling problem with soft actor-critic algorithm View paper

- [8] Automatic ship collision avoidance using deep reinforcement learning with LSTM in continuous action spaces View paper
- [9] Efficient dynamic spectrum anti-jamming access with large action space: An action space decomposition-based approach View paper
- [10] Growing action spaces View paper
- [11] Deep Reinforcement Learning-Based Dynamic Reconfiguration Planning for Digital Twin-Driven Smart Manufacturing Systems With Reconfigurable Machine Tools View paper
- [12] Online Learning of Expanding Graphs View paper
- [13] Unified Flowing Normality Learning for Rotating Machinery Anomaly Detection in Continuous Time-Varying Conditions View paper
- [14] Adaptive autoencoder latent space tuning for more robust machine learning beyond the training set for six-dimensional phase space diagnostics of a time-varying ultrafast electron-diffraction compact accelerator. View paper
- [15] Adaptive urban traffic signal control based on enhanced deep reinforcement learning View paper
- [16] FLD: Fourier Latent Dynamics for Structured Motion Representation and Learning View paper
- [17] Deep Reinforcement Learning with Coalition Action Selection for Online Combinatorial Resource Allocation with Arbitrary Action Space View paper
- [18] A deep reinforcement learning-based pid tuning strategy for nonlinear mimo systems with time-varying uncertainty View paper
- [19] Learning Time-Varying Multi-Region Brain Communications via Scalable Markovian Gaussian Processes View paper
- [20] An Online Energy Management Strategy for Fuel Cell Vehicles Based on Fuzzy Q-Learning and Road Condition Recognition View paper
- [21] Lifelong Learning with a Changing Action Set View paper
- [22] Deep reinforcement learning-based adaptive modulation for OFDM underwater acoustic communication system View paper
- [23] Structured Cooperative Reinforcement Learning With Time-Varying Composite Action Space View paper
- [24] Energy-efficient underwater acoustic communication based on Dyna-Q with an adaptive action space View paper
- [25] Deep Reinforcement Learning With Communication Transformer for Adaptive Live Streaming in Wireless Edge Networks View paper
- [26] Fleet rebalancing for expanding shared e-mobility systems: A multi-agent deep reinforcement learning approach View paper
- [27] SMAUG: A Sliding Multidimensional Task Window-Based MARL Framework for Adaptive Real-Time Subtask Recognition View paper
- [28] Growing Q-Networks: Solving Continuous Control Tasks with Adaptive Control Resolution View paper
- [29] Towards general single-utensil food acquisition with human-informed actions View paper
- [30] Online Continual Learning via Dynamic Expandable Recursive Model View paper
- [31] Online Learning of a Probabilistic and Adaptive Scene Representation View paper
- [32] Real-Time Scheduling Based on Simulation and Deep Reinforcement Learning with Featured Action Space View paper
- [33] How to Disturb Network Reconnaissance: A Moving Target Defense Approach Based on Deep Reinforcement Learning View paper
- [34] Online Learning Based Efficient Resource Allocation for LoRaWAN Network View paper
- [35] Cooperative Path Planning Method for Enhancing Ground-Units Survivability Based on Adaptive Q-Learning View paper
- [36] Action-Adaptive Continual Learning: Enabling Policy Generalization under Dynamic Action Spaces View paper
- [37] An adaptive variable-parameter dynamic learning network for solving constrained time-varying QP problem. View paper
- [38] Magnetic Field Compensation Control for Spin-Exchange Relaxation-Free Comagnetometer Using Reinforcement Learning View paper
- [39] Dynamic interval restrictions on action spaces in deep reinforcement learning for obstacle avoidance View paper
- [40] Amortized Q-learning with Model-based Action Proposals for Autonomous Driving on Highways View paper
- [41] Flexible Route Network Planning Adapting to Time-varying Air Traffic Using Reinforcement Learning View paper
- [42] Deep reinforcement learning for online combinatorial resource allocation with arbitrary state and action spaces
- [43] Variable resolution dynamic programming: Efficiently learning action maps in multivariate real-valued state-spaces View paper
- [44] A PSO-based multi-objective multi-label feature selection method in classification View paper
- [45] Learning Robotic Manipulation Skills Using an Adaptive Force-Impedance Action Space View paper
- [46] Adapting the Behavior of Reinforcement Learning Agents to Changing Action Spaces and Reward Functions View paper
- [47] Adaptive Expansion Bayesian Optimization for Unbounded Global Optimization View paper
- [48] A reinforcement learning framework for train rescheduling View paper
- [49] Distributed adaptive clustering learning over time-varying multitask networks View paper
- [50] Adaptive Space Expansion for Fast Motion Planning View paper
- [51] Fine-tuning offline policies with optimistic action selection View paper
- [52] Information-theoretic confidence bounds for reinforcement learning View paper
- [53] Learning the optimal control for evolving systems with converging dynamics View paper
- [54] Better exploration with optimistic actor critic View paper
- [55] Wasserstein actor-critic: directed exploration via optimism for continuous-actions control View paper
- [56] Why so pessimistic? estimating uncertainties for offline rl through ensembles, and why their independence matters View paper
- [57] Dyna-T: Dyna-Q and Upper Confidence Bounds Applied to Trees View paper
- [58] Uncertainty Quantification and Exploration for Reinforcement Learning View paper
- [59] Principled Exploration via Optimistic Bootstrapping and Backward Induction View paper
- [60] TRUST-ME: Trust-Based Resource Allocation and Server Selection in Multi-Access Edge Computing View paper
- [61] Tight Regret Bounds for Infinite-armed Linear Contextual Bandits View paper
- [62] Model-free online learning in unknown sequential decision making problems and games View paper
- [63] Adversarial Online Learning with Changing Action Sets: Efficient Algorithms with Approximate Regret Bounds View paper
- [64] From External to Swap Regret 2.0: An Efficient Reduction for Large Action Spaces View paper
- [65] Distributed No-Regret Learning for Multi-Stage Systems with End-to-End Bandit Feedback View paper
- [66] Online Learning, Uniform Convergence, and a Theory of Interpretability View paper
- [67] Online Learning Based Performance Optimization in Wireless Networks with Context Information View paper
- [68] Online Learning in Markov Decision Processes with Continuous Actions View paper
- [69] CFWS: DRL-Based Framework for Energy Cost and Carbon Footprint Optimization in Cloud Data Centers View paper
- [70] An adaptive learning-based approach for nearly optimal dynamic charging of electric vehicle fleets View paper
- [71] AGFT: An Adaptive GPU Frequency Tuner for Real-Time LLM Inference Optimization View paper

- [72] A Deep Reinforcement Learning Approach for Multi-UAV Collaborative Coverage with Adaptive Step Size and Dynamic Reward Mechanism View paper
- [73] Generating learning sequences for decision makers through data mining and competence set expansion View paper
- [74] Deep Reinforcement Learning-Graph Neural Networks-Dynamic Clustering triplet for Adaptive Multi Energy Microgrid optimization View paper
- [75] Parameterized Deep Reinforcement Learning With Hybrid Action Space for Edge Task Offloading View paper
- [76] Reinforced Imitation in Heterogeneous Action Space View paper