Novelty Assessment Report

Paper: Dual-Objective Reinforcement Learning with Novel Hamilton-Jacobi-Bellman Formulations
PDF URL: https://openreview.net/pdf?id=1SdPgRQrr5
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-29

Abstract

Hard constraints in reinforcement learning (RL) often degrade policy performance. Lagrangian methods offer a way to blend objectives with constraints, but require intricate reward engineering and parameter tuning. In this work, we extend recent advances that connect Hamilton-Jacobi (HJ) equations with RL to propose two novel value functions for dual-objective satisfaction. Namely, we address: 1) the Reach-Always-Avoid (RAA) problem – of achieving distinct reward and penalty thresholds – and 2) the Reach-Reach (RR) problem – of achieving thresholds of two distinct rewards. In contrast with temporal logic approaches, which typically involve representing an automaton, we derive explicit, tractable Bellman forms in this context via decomposition. Specifically, we prove that the RAA and RR problems may be rewritten as compositions of previously studied HJ-RL problems. We leverage our analysis to propose a variation of Proximal Policy Optimization (DO-HJ-PPO), and demonstrate that it produces distinct behaviors from previous approaches, out-competing a number of baselines in success, safety and speed across a range of tasks for safe-arrival and multi-target achievement.

Core Task Landscape

This paper addresses: **Dual-Objective Reinforcement Learning with Reach-Avoid Constraints**
A total of **50 papers** were analyzed and organized into a taxonomy with **13 categories**.

Taxonomy Overview

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

- **Theoretical Foundations and Formulations**
- **Algorithmic Approaches and Training Frameworks**
- **Application-Specific Implementations**
- **Game-Theoretic and Adversarial Formulations**

Complete Taxonomy Tree

- Dual-Objective Reinforcement Learning with Reach-Avoid Constraints Survey Taxonomy
- Theoretical Foundations and Formulations
  - Hamilton-Jacobi Reachability and Value Function Formulations ★ (4 papers)
  - [0] Dual-Objective Reinforcement Learning with Novel Hamilton-Jacobi-Bellman Formulations (Anon et al., 2026) View paper
  - [11] Safe Multi-Agent Reinforcement Learning via Approximate Hamilton-Jacobi Reachability (Kai Zhu, 2024) View paper
  - [28] Safety-Critical Human-Machine Shared Driving for Vehicle Collision Avoidance based on Hamilton-Jacobi reachability (Zhao Shiyue, 2025) View paper
  - [49] Dual-Objective Reinforcement Learning through novel Hamilton-Jacobi Bellman Formulations (W Sharpless, n.d.) View paper
  - Duality Theory and Optimization Guarantees (3 papers)
  - [6] Constrained reinforcement learning has zero duality gap (Paternain, 2019) View paper
  - [17] Compositional Policy Learning in Stochastic Control Systems with Formal Guarantees (Å½ikeliÄ, 2023) View paper
  - [47] Qualitative Multi-Objective Reachability for Ordered Branching MDPs (Etessami, 2020) View paper
  - Multi-Objective Problem Formulations (3 papers)
  - [25] Multi-objective safe reinforcement learning: the relationship between multi-objective reinforcement learning and safe reinforcement learning (Naoto Horie, 2019) View paper
  - [44] AI apology: interactive multi-objective reinforcement learning for human-aligned AI (Hadassah Harland, 2023) View paper
  - [50] Placing approach-avoidance conflict within the framework of multi-objective reinforcement learning (Enkhzaya Enkhtaivan, 2023) View paper
- Algorithmic Approaches and Training Frameworks
  - Dual-Agent and Modular Policy Architectures (4 papers)
  - [3] Safe Reinforcement Learning via Control-Theoretic Regularization: A Dual-Agent Framework with Hard Safety Guarantees (Ghosh, 2025) View paper
  - [9] Safe CoR: A Dual-Expert Approach to Integrating Imitation Learning and Safe Reinforcement Learning Using Constraint Rewards (Hyeokjin Kwon, 2024) View paper
  - [27] Safety Correction from Baseline: Towards the Risk-aware Policy in Robotics via Dual-agent Reinforcement Learning (Linrui Zhang, 2022) View paper
  - [41] Collaborative promotion: Achieving safety and task performance by integrating imitation reinforcement learning (Cai Zhang, 2024) View paper
  - Reward Engineering and Policy Optimization (4 papers)
  - [8] Safety optimized reinforcement learning via multi-objective policy optimization (Homayoun Honari, 2024) View paper
  - [12] Safe and balanced: A framework for constrained multi-objective reinforcement learning (Shangding Gu, 2025) View paper
  - [32] Reinforcement Learning Agent Training with Goals for Real World Tasks (Zhao Xuan, 2021) View paper

- ◦ [35] Game-Theoretic Constrained Policy Optimization for Safe Reinforcement Learning (Changxin Zhang, 2025) View paper
- ◦ Constrained RL Techniques (4 papers)
- ◦ [5] Motion Planner With Fixed-Horizon Constrained Reinforcement Learning for Complex Autonomous Driving Scenarios (Ke Lin, 2024) View paper
- ◦ [15] Safe Reinforcement Learning With Dual Robustness (Zeyang Li, 2024) View paper
- ◦ [20] Resource Allocation for Multi-target Radar Tracking via Constrained Deep Reinforcement Learning (Ziyang Lu, 2023) View paper
- ◦ [40] Enhanced Synchronization in Vehicular Digital Twin Networks: A Constrained Deep Reinforcement Learning Algorithm (H Wang, 2024) View paper
- ◦ Training Environment Design and Optimization (3 papers)
- ◦ [14] Learner engagement regulation of dual-user training based on deep reinforcement learning (Yang Yang, 2023) View paper
- ◦ [21] Optimized Deep Reinforcement Learning for Dual-Task Control in Deep-Sea Mining: Path Following and Obstacle Avoidance (Yulong Xue, 2025) View paper
- ◦ [36] Addressing Multi-Objective and domain adaptation challenges in Reinforcement Learning through case studies in multi-agent navigation and visual servoing (Ghalehseyf, 2023) View paper
- • Application-Specific Implementations
  - ◦ Autonomous Driving and Traffic Control (4 papers)
  - ◦ [1] Multi-objective deep reinforcement learning approach for adaptive traffic signal control system with concurrent optimization of safety, efficiency, and decarbonization at â⊡ (G Zhang, 2024) View paper
  - ◦ [10] Multi-objective deep reinforcement learning for variable speed limit control (Asmae Rhanizar, 2024) View paper
  - ◦ [18] Adaptive traffic signal safety and efficiency improvement by multi objective deep reinforcement learning approach (Shahin Mirbakhsh, 2024) View paper
  - ◦ [42] AoI-Aware Resource Allocation for C-V2X Networks via Multi-Agent Reinforcement Learning with Attention (Jiaao Chen, 2024) View paper
  - ◦ Robotics and Navigation (5 papers)
  - ◦ [16] Taming an autonomous surface vehicle for path following and collision avoidance using deep reinforcement learning (Eivind Meyer, 2020) View paper
  - ◦ [29] End-to-end Multi-Objective Deep Reinforcement Learning for Autonomous Navigation (Qiguang Shen, 2023) View paper
  - ◦ [30] A Dual-Functional Push-Grasping Strategy for Goal-Agnostic and Goal-Oriented Tasks Based on Deep Reinforcement Learning (Hongtao Liu, 2025) View paper
  - ◦ [37] A Multi-Objective Reinforcement Learning Based Controller for Autonomous Navigation in Challenging Environments (Amir Ramezani Dooraki, 2022) View paper
  - ◦ [46] Enhancing Robotic Navigation: An Evaluation of Single and Multi-Objective Reinforcement Learning Strategies (V. R. Young, 2023) View paper
  - ◦ Multi-Agent Systems and Coordination (6 papers)
  - ◦ [2] Federated reinforcement learning approach for detecting uncertain deceptive target using autonomous dual UAV system (Haythem Bany Salameh, 2023) View paper
  - ◦ [19] Coordinated defense allocation in reach-avoid scenarios with efficient online optimization (Liu Junwei, 2023) View paper
  - ◦ [26] Multi-Objective Dynamic Path Planning with Multi-Agent Deep Reinforcement Learning (Mengxue Tao, 2024) View paper
  - ◦ [31] Multi agent reinforcement learning for UAV collision avoidance (Nor Asilah Wati Abdul Hamid, 2024)
  - ◦ [33] Cooperative trajectory planning for multiple CAVs via multi-agent reinforcement learning (Xu Ming Ji, 2025) View paper
  - ◦ [45] Pedestrian simulation as multi-objective reinforcement learning (Naresh Balaji Ravichandran, 2018) View paper
  - ◦ Energy and Resource Management (3 papers)
  - ◦ [4] The practice of reinforcement learning algorithm in dynamic regulation and energy efficiency Optimization of wastewater treatment system in industrial parks (Tingting Wang, 2025) View paper
  - ◦ [7] A Multiagent Deep Reinforcement Learning-Enabled Dual-Branch Damping Controller for Multimode Oscillation (Guozhou Zhang, 2023) View paper
  - ◦ [24] Low-carbon economic dispatch optimization of a virtual power plant based on deep reinforcement learning in China's carbon market environment (Gengqi Wu, 2022) View paper
  - ◦ Domain-Specific Applications (4 papers)
  - ◦ [13] Intelligent Food Safety: A Prediction Model Based on Attention Mechanism and Reinforcement Learning (Mingxia Wu, 2024) View paper
  - ◦ [23] Q-learning based multi-objective clustering algorithm for cognitive radio ad hoc networks (Md Arman Hossen, 2019) View paper
  - ◦ [34] A Multi-Objective Hitch Avoidance Algorithm Using NSGA- II (Pratima Manhas, 2023) View paper
  - ◦ [39] LS-MolGen: Ligand-and-Structure Dual-Driven Deep Reinforcement Learning for Target-Specific Molecular Generation Improves Binding Affinity and Novelty. (Song Li, 2023) View paper
- • Game-Theoretic and Adversarial Formulations (4 papers)
  - ◦ [22] A Novel Multi-Objective Reinforcement Learning Algorithm for Pursuit-Evasion Game (Hu Penglin, 2025) View paper
  - ◦ [38] Obstacle Avoidance Terminal Guidance Law Design Considering Terminal Angle (Ruimin Jiang, 2025) View paper
  - ◦ [43] Multi-objective path-planning for autonomous agents using dynamic game theory (Selvakumar, 2019) View paper
  - ◦ [48] ENSURING SAFETY WHILE ENHANCING PERFORMANCE: ENCOURAGING REINFORCEMENT LEARNING BY ADDRESSING CONSTRAINTS AND â⊡ (Aghbolagh, 2024) View paper

Narrative

Core task: Dual-objective reinforcement learning with reach-avoid constraints addresses the challenge of training agents that must simultaneously optimize performance while guaranteeing safety through formal reach-avoid specifications. The field's structure reflects a blend of rigorous theory and practical deployment. Theoretical Foundations and Formulations establish the mathematical underpinnings, particularly through Hamilton-Jacobi reachability and value function methods that provide formal guarantees for constraint satisfaction. Algorithmic Approaches and Training Frameworks translate these foundations into learnable policies, exploring techniques ranging from constrained policy optimization to dual-agent architectures that decouple reward maximization from safety enforcement. Application-Specific Implementations demonstrate the breadth of domains—from autonomous navigation and traffic control to robotics and resource allocation—where dual objectives arise naturally. Game-Theoretic and Adversarial Formulations extend the framework to multi-agent settings, capturing competitive or cooperative dynamics where reach-avoid constraints must be satisfied under strategic interactions.

Several active lines reveal key trade-offs between computational tractability and safety guarantees. Works like HJ Reachability[11] and Dual Agent Safety[3] emphasize rigorous value-function approaches that offer strong theoretical assurances but can be computationally

intensive, while methods such as Safety Policy Optimization[8] and Safe CoR[9] pursue scalable training frameworks that balance empirical performance with constraint adherence. The original paper, Dual Objective HJB[0], sits squarely within the Hamilton-Jacobi reachability cluster, contributing formal value-function formulations for dual-objective problems. Compared to neighbors like Shared Driving Collision[28], which applies reach-avoid logic to human-robot interaction, Dual Objective HJB[0] focuses on the foundational mathematical characterization rather than domain-specific tuning. This positioning highlights an ongoing tension: how to preserve the elegance of HJ-based guarantees while enabling the flexibility needed for diverse real-world applications, a question that continues to drive research across all branches of the taxonomy.

Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

1. Safe Multi-Agent Reinforcement Learning via Approximate Hamilton-Jacobi Reachability

Authors: Kai Zhu, Fengbo Lan, Wenbo Zhao, Zhu Kai, Tao Zhang | **Year/Venue**: 2024 | **URL**: View paper

Abstract

â¦ Then CM3 [15] takes a further step by incorporating multi-objective and course learning â¦ Our method adeptly navigates towards the landmarks while avoiding collisions, even under â¦

Relationship Analysis

Both papers belong to the Hamilton-Jacobi Reachability and Value Function Formulations category, deriving HJ-based value functions and Bellman equations for reach-avoid problems. The original paper focuses on dual-objective problems (Reach-Always-Avoid and Reach-Reach) with novel decomposition theorems for single-agent settings, while the candidate paper applies HJ reachability to multi-agent reinforcement learning with a centralized training/decentralized execution framework. The key distinction is that the original paper addresses compositional single-agent problems through value decomposition, whereas the candidate paper extends HJ methods to multi-agent coordination and safety under non-cooperative conditions.

2. Safety-Critical Human-Machine Shared Driving for Vehicle Collision Avoidance based on Hamilton-Jacobi reachability

Authors: Zhao Shiyue, Zhang JunZhi, Shiyue Zhao, Zhou Rui, Junzhi Zhang, et al. (15 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Road safety continues to be a pressing global issue, with vehicle collisions imposing significant human, societal, and economic burdens. Human-machine shared collision avoidance in critical collision scenarios aims to aid drivers' accident avoidance through intervening only when necessary. Existing methods count on replanning collision-free trajectories and imposing human-machine tracking, which usually interrupts the driver's intent and increases the risk of conflict. This paper introduces a Re...

Relationship Analysis

Both papers belong to the Hamilton-Jacobi Reachability and Value Function Formulations category, utilizing HJ-based approaches for safety-critical control problems. The original paper derives novel HJ-Bellman equations for dual-objective RL problems (Reach-Always-Avoid and Reach-Reach) with theoretical decomposition results and applies them to general RL tasks using PPO. The candidate paper focuses specifically on vehicle collision avoidance using HJ reachability analysis to define Collision Avoidance Reachable Sets (CARS) and integrates this with human-machine shared control and RL for automotive safety applications, rather than developing new theoretical HJ formulations for compositional objectives.

3. Dual-Objective Reinforcement Learning through novel Hamilton-Jacobi Bellman Formulations

Authors: W Sharpless, D Hirsch, S Tonkens, NU Shinde | **URL**: View paper

Abstract

â¦ Abstract: Hard constraints in reinforcement learning (RL), â¦ a way to blend objectives with constraints, but often require â¦ for these simple reach/avoid/reach-avoid problem formulations. In â¦

⚠ Similarity Notice

These papers appear to be the same work or very closely related variants. Both papers have nearly identical titles ('Dual-Objective Reinforcement Learning with Novel Hamilton-Jacobi-Bellman Formulations'), present the same core contributions (RAA and RR problems, decomposition theorems, DOHJ-PPO algorithm), use identical mathematical formulations and notation, and describe the same experimental setups. The abstracts, problem definitions, theoretical results (Theorems 1 and 2), and algorithmic approaches are essentially identical, suggesting these are likely the same paper submitted to different venues or different versions of the same work.

Contributions Analysis

Overall novelty summary. The paper proposes novel value functions for two dual-objective problems—Reach-Always-Avoid (RAA) and Reach-Reach (RR)—within the Hamilton-Jacobi reachability framework. It resides in the 'Hamilton-Jacobi Reachability and Value Function Formulations' leaf, which contains four papers total, including the original work. This leaf represents a relatively focused research direction within the broader taxonomy of 50 papers across 13 leaf nodes, suggesting a specialized but not overcrowded area where formal HJ-based methods are actively developed.

The taxonomy reveals that the paper's leaf sits within 'Theoretical Foundations and Formulations,' adjacent to 'Duality Theory and Optimization Guarantees' (3 papers) and 'Multi-Objective Problem Formulations' (3 papers). These neighboring branches address complementary concerns—duality gaps, convergence properties, and Pareto optimality—while the HJ reachability leaf emphasizes explicit value-function derivations and Bellman forms. The paper's focus on decomposing dual-objective problems into compositions of simpler HJ-RL problems distinguishes it from multi-objective scalarization methods and from purely algorithmic approaches in the 'Algorithmic Approaches and Training Frameworks' branch.

Among 30 candidates examined, none were found to clearly refute any of the three contributions: the RAA/RR value functions (10 candidates, 0 refutable), the decomposition theorems (10 candidates, 0 refutable), and the DOHJ-PPO algorithm (10 candidates, 0 refutable). This suggests that within the limited search scope, the specific formulations and decomposition results appear distinct from prior work. The three sibling papers in the same leaf address HJ reachability and value functions but do not appear to cover the exact RAA and RR problem structures or their compositional characterizations.

Based on the top-30 semantic matches and the taxonomy structure, the work appears to occupy a relatively novel position within HJ-based dual-objective RL. The limited search scope means that more exhaustive examination—particularly of temporal logic and automaton-based methods mentioned in the abstract—could reveal additional overlaps. However, the explicit Bellman forms and decomposition approach seem to differentiate this work from both neighboring theoretical formulations and applied constraint-satisfaction methods.

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

Contribution 1: Novel value functions for Reach-Always-Avoid and Reach-Reach problems

Description: The authors introduce two new value functions for dual-objective satisfaction in reinforcement learning: the Reach-Always-Avoid (RAA) problem, which requires reaching a goal while perennially avoiding hazards, and the Reach-Reach (RR) problem, which requires reaching two distinct goals in either order. These formulations extend existing Hamilton-Jacobi reachability methods to more complex compositional tasks.

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. A Safe Navigation Algorithm for Differential-Drive Mobile Robots by Using Fuzzy Logic Reward Function-Based Deep Reinforcement Learning

URL: View paper

Brief Assessment

Fuzzy Reward Navigation[56] focuses on mobile robot navigation using fuzzy logic reward functions for obstacle avoidance and goal reaching, not on Hamilton-Jacobi value function formulations for dual-objective RL problems.

---

2. Risk averse robust adversarial reinforcement learning

URL: View paper

Brief Assessment

Risk Averse Adversarial[58] focuses on risk-averse reinforcement learning through adversarial training in autonomous driving contexts, not on Hamilton-Jacobi reachability methods or compositional reach-avoid value functions.

---

3. Safe Multi-Agent Navigation Guided by Goal-Conditioned Safe Reinforcement Learning

URL: View paper

Brief Assessment

Goal Conditioned Safety[52] focuses on safe multi-agent navigation using goal-conditioned RL with graph-based planning, not on introducing novel value functions for dual-objective satisfaction in RL. The candidate addresses navigation tasks rather than developing new Hamilton-Jacobi-Bellman formulations for compositional reach-avoid problems.

---

4. Boundary-aware value function generation for safe stochastic motion planning

URL: View paper

Brief Assessment

Boundary Aware Value[59] focuses on safe stochastic motion planning using finite element methods to satisfy boundary conditions for obstacle avoidance in continuous state spaces. It does not address dual-objective reinforcement learning with compositional reach-always-avoid or reach-reach formulations as defined in the original paper.

---

5. A Multiplicative Value Function for Safe and Efficient Reinforcement Learning

URL: View paper

Brief Assessment

Multiplicative Value Function[54] focuses on safe RL with a multiplicative value function combining safety and reward critics for constraint satisfaction, not on Hamilton-Jacobi reachability formulations for dual-objective compositional tasks like RAA and RR problems.

---

6. Lyapunov-Inspired Deep Reinforcement Learning for Robot Navigation in Obstacle Environments

URL: View paper

Brief Assessment

Lyapunov Robot Navigation[60] focuses on reward shaping using Lyapunov stability theory for obstacle avoidance in robot navigation, not on dual-objective value functions or Hamilton-Jacobi reachability formulations for compositional tasks.

---

7. SIGN: Safety-Aware Image-Goal Navigation for Autonomous Drones via Reinforcement Learning

URL: View paper

Brief Assessment

SIGN Drone Navigation[51] focuses on image-goal navigation for autonomous drones using reinforcement learning with continuous velocity control and safety modules. It does not address dual-objective value functions or Hamilton-Jacobi reachability formulations for compositional tasks.

---

8. Robot Mapless Navigation in VUCA Environments via Deep Reinforcement Learning

URL: View paper

Brief Assessment

VUCA Navigation[53] focuses on robot mapless navigation in social environments using deep RL with spatial-temporal reasoning and lidar maps. It does not address dual-objective value functions, Hamilton-Jacobi reachability methods, or compositional task formulations like RAA/RR problems.

---

9. Hierarchical Planning Through Goal-Conditioned Offline Reinforcement Learning

URL: View paper

Brief Assessment

Hierarchical Goal Planning[57] focuses on goal-conditioned offline RL for hierarchical planning in robotics, not on Hamilton-Jacobi value functions for dual-objective satisfaction problems like RAA and RR.

---

10. Safety and liveness guarantees through reach-avoid reinforcement learning

URL: View paper

Brief Assessment

Reach Avoid Guarantees[55] focuses on reach-avoid problems (reaching goals while avoiding hazards en route), not the Reach-Always-Avoid (perennially avoiding hazards after reaching goals) or Reach-Reach (reaching two distinct goals in either order) problems introduced in the original paper.

Contribution 2: Decomposition theorems for RAA and RR value functions

Description: The authors prove that the RAA and RR value functions can be decomposed into combinations of simpler reach, avoid, and reach-avoid value functions. Specifically, Theorem 1 shows RAA decomposes into avoid and reach-avoid problems, while Theorem 2 shows RR decomposes into three reach problems. This decomposition enables tractable solutions using existing methods.

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. Spatiotemporal Tubes based Controller Synthesis against Omega-Regular Specifications for Unknown Systems
URL: View paper

Brief Assessment

Spatiotemporal Tubes Controller[64] focuses on decomposing omega-regular specifications (recognized by Büchi automata) into reach-avoid tasks for controller synthesis, not on decomposing value functions in reinforcement learning contexts as the original paper does.

2. Collaborative Constrained Target-Reaching Control in a Multiplayer Reach-Avoid Game
URL: View paper

Brief Assessment

Collaborative Target Reaching[65] focuses on multiplayer reach-avoid games with adversarial agents and virtual defense channels, not on decomposing RAA/RR value functions into simpler reach, avoid, and reach-avoid problems as in the original paper.

3. Fast nonlinear controller synthesis using reachability analysis
URL: View paper

Brief Assessment

Fast Nonlinear Controller[69] focuses on controller synthesis for nonlinear systems using reachability analysis, not on decomposing reach-avoid problems into simpler reachability subproblems for reinforcement learning.

4. Defense penetration strategy for inferior USV based on reach-avoid game under complex target region
URL: View paper

Brief Assessment

USV Defense Penetration[61] focuses on reach-avoid game problems for unmanned surface vehicles in defense scenarios, not on decomposing dual-objective RL problems (RAA/RR) into simpler reach, avoid, and reach-avoid subproblems as in the original paper.

5. Optimal Strategies and Cooperative Teaming for 3-D Multiplayer Reach-Avoid Games
URL: View paper

Brief Assessment

Multiplayer Reach Avoid[62] focuses on decomposing multiplayer reach-avoid games into subgames for multi-agent pursuit-evasion scenarios in 3D space, not on decomposing dual-objective value functions (RAA/RR) into simpler reach, avoid, and reach-avoid components as in the original paper.

6. Compositional Policy Learning in Stochastic Control Systems with Formal Guarantees
URL: View paper

Brief Assessment

Compositional Policy Learning[17] focuses on compositional learning over graphs of probabilistic reach-avoid specifications in stochastic environments, not on decomposing RAA/RR value functions into simpler reach, avoid, and reach-avoid problems as the original paper does.

7. Compositional automata embeddings for goal-conditioned reinforcement learning
URL: View paper

Brief Assessment

Compositional Automata Embeddings[67] focuses on decomposing temporal tasks represented as automata into reach-avoid subproblems for goal-conditioned RL, not on decomposing value functions in Hamilton-Jacobi formulations. The technical approaches and problem domains are fundamentally different.

8. Reach-avoid analysis for stochastic differential equations
URL: View paper

Brief Assessment

Stochastic Reach Avoid[68] focuses on stochastic differential equations and barrier certificate methods for reach-avoid problems, not on decomposing dual-objective RL problems (RAA/RR) into simpler reach, avoid, and reach-avoid subproblems as proposed in the original paper.

9. DART-LLM: Dependency-Aware Multi-Robot Task Decomposition and Execution using Large Language Models
URL: View paper

Brief Assessment

DART LLM[66] focuses on multi-robot task decomposition using directed acyclic graphs for natural language instructions, not on decomposing reach-avoid value functions in reinforcement learning. The technical domains are entirely different.

10. LLM-Augmented Symbolic RL with Landmark-Based Task Decomposition
URL: View paper

Brief Assessment

LLM Symbolic RL[63] focuses on task decomposition using symbolic logic and LLMs for identifying subtasks in RL, not on decomposing reach-avoid value functions into mathematical compositions of simpler HJ-RL problems.

Contribution 3: DOHJ-PPO algorithm for dual-objective reinforcement learning

Description: The authors develop DOHJ-PPO, a novel algorithm that extends Proximal Policy Optimization to solve the RAA and RR problems. The algorithm bootstraps concurrently solved decompositions for coupling on-policy rollouts, using stochastic relaxations of the Bellman equations (SRBE and SRABE) to handle stochastic policies and dynamics.

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. Incentivizing safer actions in policy optimization for constrained reinforcement learning
URL: View paper

Brief Assessment

Incentivizing Safer Actions[70] focuses on constrained RL with penalty functions for single-objective optimization under safety constraints, not dual-objective problems like RAA/RR that require reaching multiple goals or avoiding hazards after goal achievement.

---

2. Augmented Proximal Policy Optimization for Safe Reinforcement Learning
URL: View paper

Brief Assessment

Augmented PPO[72] focuses on safe RL with Lagrangian-based constraint satisfaction in single-objective settings, not dual-objective problems like reach-always-avoid (RAA) or reach-reach (RR) that require Hamilton-Jacobi decompositions.

---

3. A Multi-Objective Intelligent Control Method for Active Distribution Power Network Using Proximal Policy Optimization Agents
URL: View paper

Brief Assessment

Active Distribution Control[75] applies standard PPO to multi-objective control in power distribution networks, not dual-objective constrained RL with Hamilton-Jacobi formulations. The candidate focuses on voltage regulation and power loss reduction in electrical grids, while the original develops novel HJ-based Bellman equations for reach-always-avoid and reach-reach problems with theoretical decomposition guarantees.

---

4. Multiâ Objective Reinforcement Learning for Automated Resilient Cyber Defence
URL: View paper

Brief Assessment

Cyber Defence MORL[76] focuses on multi-objective cyber defense using MOPPO and PCN algorithms for network security tasks, not on dual-objective Hamilton-Jacobi formulations for reach-always-avoid and reach-reach problems with decomposition-based PPO variants.

---

5. Mobile communications, computing, and caching resources allocation for diverse services via multi-objective proximal policy optimization
URL: View paper

Brief Assessment

Mobile Resource Allocation[78] focuses on multi-objective resource allocation in mobile networks using MO-PPO for Pareto optimization, not dual-objective constrained RL with Hamilton-Jacobi formulations for safety-critical tasks.

---

6. Pareto Envelope Augmented with Reinforcement Learning Multi-Objective Reinforcement Learning-Based Approach for Large-Scale Constrained Pressurized Water â
URL: View paper

Brief Assessment

Pareto Envelope RL[77] addresses multi-objective optimization in water distribution networks using PPO, but focuses on Pareto-optimal solutions rather than the specific dual-objective reach-always-avoid (RAA) and reach-reach (RR) problems with Hamilton-Jacobi formulations that DOHJ-PPO targets.

---

7. Mirror Descent Policy Optimization
URL: View paper

Brief Assessment

Mirror Descent Policy[74] focuses on trust-region optimization methods for standard RL objectives, not dual-objective problems involving reach-always-avoid or reach-reach formulations with Hamilton-Jacobi equations.

---

8. Multi-Objective Proximal Policy Optimization for Digital Twin-Assisted Computing Offloading in Internet of Vehicles
URL: View paper

Brief Assessment

Digital Twin Offloading[73] applies PPO to multi-objective optimization in vehicular edge computing, not dual-objective constrained RL with Hamilton-Jacobi formulations. The candidate focuses on IoV task offloading using MOEA/D combined with PPO, which is a different problem domain and methodology than the original's reach-always-avoid and reach-reach problems with novel Bellman equations.

---

9. TUC-PPO: Team Utility-Constrained Proximal Policy Optimization for Spatial Public Goods Games
URL: View paper

Brief Assessment

Team Utility Constrained[71] focuses on spatial public goods games with team utility constraints, not dual-objective reach-always-avoid or reach-reach problems. The candidate extends PPO for cooperative multi-agent settings rather than individual agent safety and target achievement.

---

10. â multi-objective optimization scheduling method for island-integrated energy systems based on meta-learning and enhanced proximal policy optimization
URL: View paper

Brief Assessment

Island Energy Scheduling[79] focuses on multi-objective optimization for energy systems using meta-learning and PPO, not dual-objective constrained RL with Hamilton-Jacobi formulations for reach-always-avoid and reach-reach problems.

Appendix: Text Similarity Detection

Textual similarity detection checked 33 papers and found 3 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Dual-Objective Reinforcement Learning through novel Hamilton-Jacobi Bellman Formulations

Detected in: Core Task (sibling)

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] Dual-Objective Reinforcement Learning with Novel Hamilton-Jacobi-Bellman Formulations View paper
- [1] Multi-objective deep reinforcement learning approach for adaptive traffic signal control system with concurrent optimization of safety, efficiency, and decarbonization at â⃞ View paper
- [2] Federated reinforcement learning approach for detecting uncertain deceptive target using autonomous dual UAV system View paper
- [3] Safe Reinforcement Learning via Control-Theoretic Regularization: A Dual-Agent Framework with Hard Safety Guarantees View paper
- [4] The practice of reinforcement learning algorithm in dynamic regulation and energy efficiency Optimization of wastewater treatment system in industrial parks View paper
- [5] Motion Planner With Fixed-Horizon Constrained Reinforcement Learning for Complex Autonomous Driving Scenarios View paper
- [6] Constrained reinforcement learning has zero duality gap View paper
- [7] A Multiagent Deep Reinforcement Learning-Enabled Dual-Branch Damping Controller for Multimode Oscillation View paper
- [8] Safety optimized reinforcement learning via multi-objective policy optimization View paper
- [9] Safe CoR: A Dual-Expert Approach to Integrating Imitation Learning and Safe Reinforcement Learning Using Constraint Rewards View paper
- [10] Multi-objective deep reinforcement learning for variable speed limit control View paper
- [11] Safe Multi-Agent Reinforcement Learning via Approximate Hamilton-Jacobi Reachability View paper
- [12] Safe and balanced: A framework for constrained multi-objective reinforcement learning View paper
- [13] Intelligent Food Safety: A Prediction Model Based on Attention Mechanism and Reinforcement Learning View paper
- [14] Learner engagement regulation of dual-user training based on deep reinforcement learning View paper
- [15] Safe Reinforcement Learning With Dual Robustness View paper
- [16] Taming an autonomous surface vehicle for path following and collision avoidance using deep reinforcement learning View paper
- [17] Compositional Policy Learning in Stochastic Control Systems with Formal Guarantees View paper
- [18] Adaptive traffic signal safety and efficiency improvement by multi objective deep reinforcement learning approach View paper
- [19] Coordinated defense allocation in reach-avoid scenarios with efficient online optimization View paper
- [20] Resource Allocation for Multi-target Radar Tracking via Constrained Deep Reinforcement Learning View paper
- [21] Optimized Deep Reinforcement Learning for Dual-Task Control in Deep-Sea Mining: Path Following and Obstacle Avoidance View paper
- [22] A Novel Multi-Objective Reinforcement Learning Algorithm for Pursuit-Evasion Game View paper
- [23] Q-learning based multi-objective clustering algorithm for cognitive radio ad hoc networks View paper
- [24] Low-carbon economic dispatch optimization of a virtual power plant based on deep reinforcement learning in China's carbon market environment View paper
- [25] Multi-objective safe reinforcement learning: the relationship between multi-objective reinforcement learning and safe reinforcement learning View paper
- [26] Multi-Objective Dynamic Path Planning with Multi-Agent Deep Reinforcement Learning View paper
- [27] Safety Correction from Baseline: Towards the Risk-aware Policy in Robotics via Dual-agent Reinforcement Learning View paper
- [28] Safety-Critical Human-Machine Shared Driving for Vehicle Collision Avoidance based on Hamilton-Jacobi reachability View paper
- [29] End-to-end Multi-Objective Deep Reinforcement Learning for Autonomous Navigation View paper
- [30] A Dual-Functional Push-Grasping Strategy for Goal-Agnostic and Goal-Oriented Tasks Based on Deep Reinforcement Learning View paper
- [31] Multi agent reinforcement learning for UAV collision avoidance
- [32] Reinforcement Learning Agent Training with Goals for Real World Tasks View paper
- [33] Cooperative trajectory planning for multiple CAVs via multi-agent reinforcement learning View paper
- [34] A Multi-Objective Hitch Avoidance Algorithm Using NSGA- II View paper
- [35] Game-Theoretic Constrained Policy Optimization for Safe Reinforcement Learning View paper
- [36] Addressing Multi-Objective and domain adaptation challenges in Reinforcement Learning through case studies in multi-agent navigation and visual servoing View paper
- [37] A Multi-Objective Reinforcement Learning Based Controller for Autonomous Navigation in Challenging Environments View paper
- [38] Obstacle Avoidance Terminal Guidance Law Design Considering Terminal Angle View paper
- [39] LS-MolGen: Ligand-and-Structure Dual-Driven Deep Reinforcement Learning for Target-Specific Molecular Generation Improves Binding Affinity and Novelty. View paper
- [40] Enhanced Synchronization in Vehicular Digital Twin Networks: A Constrained Deep Reinforcement Learning Algorithm View paper
- [41] Collaborative promotion: Achieving safety and task performance by integrating imitation reinforcement learning View paper
- [42] AoI-Aware Resource Allocation for C-V2X Networks via Multi-Agent Reinforcement Learning with Attention View paper
- [43] Multi-objective path-planning for autonomous agents using dynamic game theory View paper
- [44] AI apology: interactive multi-objective reinforcement learning for human-aligned AI View paper
- [45] Pedestrian simulation as multi-objective reinforcement learning View paper
- [46] Enhancing Robotic Navigation: An Evaluation of Single and Multi-Objective Reinforcement Learning Strategies View paper
- [47] Qualitative Multi-Objective Reachability for Ordered Branching MDPs View paper
- [48] ENSURING SAFETY WHILE ENHANCING PERFORMANCE: ENCOURAGING REINFORCEMENT LEARNING BY ADDRESSING CONSTRAINTS AND â⃞ View paper
- [49] Dual-Objective Reinforcement Learning through novel Hamilton-Jacobi Bellman Formulations View paper
- [50] Placing approach-avoidance conflict within the framework of multi-objective reinforcement learning View paper

- [51] SIGN: Safety-Aware Image-Goal Navigation for Autonomous Drones via Reinforcement Learning View paper
- [52] Safe Multi-Agent Navigation Guided by Goal-Conditioned Safe Reinforcement Learning View paper
- [53] Robot Mapless Navigation in VUCA Environments via Deep Reinforcement Learning View paper
- [54] A Multiplicative Value Function for Safe and Efficient Reinforcement Learning View paper
- [55] Safety and liveness guarantees through reach-avoid reinforcement learning View paper
- [56] A Safe Navigation Algorithm for Differential-Drive Mobile Robots by Using Fuzzy Logic Reward Function-Based Deep Reinforcement Learning View paper
- [57] Hierarchical Planning Through Goal-Conditioned Offline Reinforcement Learning View paper
- [58] Risk averse robust adversarial reinforcement learning View paper
- [59] Boundary-aware value function generation for safe stochastic motion planning View paper
- [60] Lyapunov-Inspired Deep Reinforcement Learning for Robot Navigation in Obstacle Environments View paper
- [61] Defense penetration strategy for inferior USV based on reach-avoid game under complex target region View paper
- [62] Optimal Strategies and Cooperative Teaming for 3-D Multiplayer Reach-Avoid Games View paper
- [63] LLM-Augmented Symbolic RL with Landmark-Based Task Decomposition View paper
- [64] Spatiotemporal Tubes based Controller Synthesis against Omega-Regular Specifications for Unknown Systems View paper
- [65] Collaborative Constrained Target-Reaching Control in a Multiplayer Reach-Avoid Game View paper
- [66] DART-LLM: Dependency-Aware Multi-Robot Task Decomposition and Execution using Large Language Models View paper
- [67] Compositional automata embeddings for goal-conditioned reinforcement learning View paper
- [68] Reach-avoid analysis for stochastic differential equations View paper
- [69] Fast nonlinear controller synthesis using reachability analysis View paper
- [70] Incentivizing safer actions in policy optimization for constrained reinforcement learning View paper
- [71] TUC-PPO: Team Utility-Constrained Proximal Policy Optimization for Spatial Public Goods Games View paper
- [72] Augmented Proximal Policy Optimization for Safe Reinforcement Learning View paper
- [73] Multi-Objective Proximal Policy Optimization for Digital Twin-Assisted Computing Offloading in Internet of Vehicles View paper
- [74] Mirror Descent Policy Optimization View paper
- [75] A Multi-Objective Intelligent Control Method for Active Distribution Power Network Using Proximal Policy Optimization Agents View paper
- [76] Multiâ￿￿Objective Reinforcement Learning for Automated Resilient Cyber Defence View paper
- [77] Pareto Envelope Augmented with Reinforcement Learning Multi-Objective Reinforcement Learning-Based Approach for Large-Scale Constrained Pressurized Water â￿¦ View paper
- [78] Mobile communications, computing, and caching resources allocation for diverse services via multi-objective proximal policy optimization View paper
- [79] â￿¦ multi-objective optimization scheduling method for island-integrated energy systems based on meta-learning and enhanced proximal policy optimization View paper