

Novelty Assessment Report

Paper: GAR: Generative Adversarial Reinforcement Learning for Formal Theorem Proving

PDF URL: <https://openreview.net/pdf?id=1MUZsrjxi9>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

Solving math problems through verifiable languages such as Lean has significantly impacted both the mathematics and computer science communities. Current state-of-the-art models are often trained with expensive online Reinforcement Learning (RL) or expert iteration. However, these approaches rely on fixed problem sets, which causes inefficient training and limits the model to tackle complex problems. To overcome these limitations, we propose **GAR**: Generative Adversarial Reinforcement learning, a comprehensive RL training framework that jointly trains the problem composer and solver in an adversarial loop. **GAR** introduces an implicit curriculum learning mechanism, which aligns task difficulty with the prover's evolving capability. It thereby improves the training efficiency and enables stronger performance of proving advanced theorems. Experiments show that with **GAR** training, Goedel-Prover-V2-8B and DeepSeek-Prover-V2-7B achieve an average relative improvement in pass@32 of **4.20%** on MiniF2F-Test benchmark, while DeepSeek-Prover-V2's pass@32 on ProofNet-Test increases from 22.58% to **25.81%**. Beyond formal proving, **GAR** establishes a general RL paradigm for co-evolution of problem generation and solving under verifiable environments.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **formal theorem proving with reinforcement learning**

A total of **50 papers** were analyzed and organized into a taxonomy with **23 categories**.

Taxonomy Overview

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

- **Core RL Algorithms and Training Frameworks**
- **Proof Search Strategies and Inference Selection**
- **Neural Architectures and State Representations**
- **Data Generation and Augmentation**
- **Domain-Specific Systems and Applications**
- **Benchmarks, Evaluation, and Analysis**
- **Cross-Domain and Interdisciplinary Approaches**

Complete Taxonomy Tree

- formal theorem proving with reinforcement learning Survey Taxonomy
- Core RL Algorithms and Training Frameworks
 - Policy Optimization and Expert Iteration (5 papers)
 - [1] Olympiad-level formal mathematical reasoning with reinforcement learning. (Thomas Hubert, 2025) [View paper](#)
 - [2] Reinforcement learning for guiding the e theorem prover (Jack McKeown, 2023) [View paper](#)
 - [3] Reinforcement learning of theorem proving (Kaliszyk, 2018) [View paper](#)
 - [8] ABEL: Sample efficient online reinforcement learning for neural theorem proving (F Gloeckle, 2024) [View paper](#)
 - [13] A deep reinforcement learning approach to first-order logic theorem proving (Maxwell Crouse, 2021) [View paper](#)
 - Adversarial and Curriculum Learning ★ (4 papers)
 - [0] GAR: Generative Adversarial Reinforcement Learning for Formal Theorem Proving (Anon et al., 2026) [View paper](#)
 - [22] Learning Interestingness in Automated Mathematical Theory Formation (George Tsoukalas, 2025) [View paper](#)
 - [45] Towards finding longer proofs (Zombori, 2021) [View paper](#)
 - [48] Learning to find proofs and theorems by learning to refine search strategies: The case of loop invariant synthesis (Laurent, 2022) [View paper](#)
 - Reward Modeling and Verifier Integration (4 papers)
 - [10] CriticLean: Critic-Guided Reinforcement Learning for Mathematical Formalization (Yao Yifan, 2025) [View paper](#)
 - [15] Leanabell-prover-v2: Verifier-integrated reasoning for formal theorem proving via reinforcement learning (Liu Yahui, 2025) [View paper](#)
 - [24] RL Tango: Reinforcing Generator and Verifier Together for Language Reasoning (Zha Kaiwen, 2025) [View paper](#)
 - [49] Rewarding the Unlikely: Lifting GRPO Beyond Distribution Sharpening (He, 2025) [View paper](#)
 - Generative Flow Networks and Alternative RL Paradigms (1 papers)
 - [30] Proof Flow: Preliminary Study on Generative Flow Network Language Model Tuning for Formal Reasoning (Ho, 2024) [View paper](#)
- Proof Search Strategies and Inference Selection
 - Monte Carlo Tree Search and Lookahead (2 papers)
 - [21] Local Look-Ahead Guidance via Verifier-in-the-Loop for Automated Theorem Proving (Rajae, 2025) [View paper](#)
 - [29] Deepseek-prover-v1. 5: Harnessing proof assistant feedback for reinforcement learning and monte-carlo tree search (Ren Z Z, 2024) [View paper](#)

- Hierarchical Proof Decomposition (3 papers)
- [6] Seed-Prover: Deep and Broad Reasoning for Automated Theorem Proving (Chen, 2025) [View paper](#)
- [9] Deepseek-prover-v2: Advancing formal mathematical reasoning via reinforcement learning for subgoal decomposition (Ren Z Z, 2025) [View paper](#)
- [11] Formal Theorem Proving by Rewarding LLMs to Decompose Proofs Hierarchically (Dong, 2024) [View paper](#)
- Metaheuristic and Strategy Selection (1 papers)
- [12] Search Strategy Selection for Automated Theorem Proving (McKeown, 2024) [View paper](#)
- Iterative Refinement and Self-Correction (3 papers)
- [5] QEDCartographer: Automating formal verification using reward-free reinforcement learning (Alex Sanchez-Stern, 2025) [View paper](#)
- [18] ProofNet++: A Neuro-Symbolic System for Formal Proof Verification with Self-Correction (Ambati, 2025) [View paper](#)
- [35] Goedel-Prover-V2: Scaling Formal Theorem Proving with Scaffolded Data Synthesis and Self-Correction (Lin Yong, 2025) [View paper](#)
- Neural Architectures and State Representations
 - Graph Neural Networks and Invariant Representations (1 papers)
 - [23] An Ensemble Approach for Automated Theorem Proving Based on Efficient Name Invariant Graph Neural Representations (Fokoue, 2023) [View paper](#)
 - Attention-Based and Transformer Architectures (1 papers)
 - [17] Holist: An environment for machine learning of higher order logic theorem proving (Kshitij Bansal, 2019) [View paper](#)
- Data Generation and Augmentation
 - Synthetic Problem and Conjecture Generation (3 papers)
 - [16] A Combinatorial Identities Benchmark for Theorem Proving via Automated Theorem Generation (Xiong Beibei, 2025) [View paper](#)
 - [41] STP: Self-play LLM Theorem Provers with Iterative Conjecturing and Proving (Dong, 2025) [View paper](#)
 - [44] LeanConjecturer: Automatic Generation of Mathematical Conjectures for Theorem Proving (Kasaura, 2025) [View paper](#)
 - Autoformalization and Natural Language Translation (3 papers)
 - [7] DeepTheorem: Advancing LLM Reasoning for Theorem Proving Through Natural Language and Reinforcement Learning (Zhang Zi-yin, 2025) [View paper](#)
 - [26] FormARL: Enhancing Autoformalization with no Labeled Data (Huang Yanxing, 2025) [View paper](#)
 - [43] Mathesis: Towards Formal Theorem Proving from Natural Languages (Xuejun Yu, 2025) [View paper](#)
 - Proof Synthesis and Data Augmentation (2 papers)
 - [28] Learning equational theorem proving (Piepenbrock, 2021) [View paper](#)
 - [42] Goedel-Prover: A Frontier Model for Open-Source Automated Theorem Proving (Lin Yong, 2025) [View paper](#)
- Domain-Specific Systems and Applications
 - Lean 4 Theorem Provers (2 papers)
 - [19] Leanabell-Prover: Posttraining Scaling in Formal Reasoning (Zhang Jingyuan, 2025) [View paper](#)
 - [27] Kimina-Prover Preview: Towards Large Formal Reasoning Models with Reinforcement Learning (Wang Haiming, 2025) [View paper](#)
 - HOL and Higher-Order Logic Systems (1 papers)
 - [36] Tacticzero: Learning to prove theorems from scratch with deep reinforcement learning (Wu Minchao, 2021) [View paper](#)
 - First-Order and Saturation-Based Provers (3 papers)
 - [34] Project proposal: A modular reinforcement learning based automated theorem prover (Shminke, 2022) [View paper](#)
 - [50] Gym-saturation: an OpenAI Gym environment for saturation provers (Boris Shminke, 2022) [View paper](#)
 - Propositional and Specialized Logic Systems (2 papers)
 - [46] Prolog Technology Reinforcement Learning Prover: (System Description) (Zombori, 2020) [View paper](#)
 - [47] Automated Theorem Proving in Intuitionistic Propositional Logic by Deep Reinforcement Learning (Mitsuru Kusumoto, 2022) [View paper](#)
 - Software Verification and Formal Methods Integration (2 papers)
 - [20] Re:Form - Reducing Human Priors in Scalable Formal Software Verification with RL in LLMs: A Preliminary Study on Dafny (Yan, 2025) [View paper](#)
 - [38] Proof2Silicon: Prompt Repair for Verified Code and Hardware Generation via Reinforcement Learning (Wan JiaXin, 2025) [View paper](#)
- Benchmarks, Evaluation, and Analysis
 - Benchmark Datasets and Evaluation Frameworks (1 papers)
 - [14] Neural Theorem Proving: Generating and Structuring Proofs for Formal Verification (Eiers, 2025) [View paper](#)
 - Efficiency and Cost Analysis (1 papers)
 - [40] EconProver: Towards More Economical Test-Time Scaling for Automated Theorem Proving (Li, 2025) [View paper](#)
 - Theoretical Foundations and Formal Verification of RL (2 papers)
 - [4] Verification using formalised mathematics and theorem proving of reinforcement and deep learning (Chevallier, 2024) [View paper](#)
 - [25] Markov Decision Processes with Formal Verification: Mathematical Guarantees for Safe Reinforcement Learning (AZMI, 2021) [View paper](#)
- Cross-Domain and Interdisciplinary Approaches
 - General Mathematical Reasoning and Learning (2 papers)
 - [32] Reinforcement Learning in Mathematical Research (Miquel Noguera I Alonso, 2024) [View paper](#)
 - [39] Peano: learning formal mathematical reasoning (Gabriel Poesia, 2022) [View paper](#)
 - Formal Methods and RL Integration (2 papers)
 - [31] Exploring and Evaluating Interplays of BPPy with Deep Reinforcement Learning and Formal Methods (Yaacov, 2025) [View paper](#)
 - [33] Formal methods meet machine learning (F3ML) (Kim G. Larsen, 2022) [View paper](#)

Narrative

Core task: formal theorem proving with reinforcement learning. The field has matured into a rich ecosystem organized around several complementary dimensions. At the highest level, work divides into core RL algorithms and training frameworks (exploring policy gradient

methods, value-based approaches, and curriculum strategies), proof search strategies and inference selection (addressing how agents navigate large search spaces), neural architectures and state representations (designing encoders for logical formulas and proof states), data generation and augmentation (creating training corpora from existing libraries or synthetic problems), domain-specific systems targeting particular proof assistants or mathematical domains, benchmarks and evaluation methodologies, and cross-domain or interdisciplinary efforts that bridge theorem proving with program synthesis or hardware verification. Early foundational efforts such as RL Theorem Proving[3] and Holist[17] established core paradigms, while recent systems like Deepseek-prover-v2[9], QEDCartographer[5], and Seed-Prover[6] demonstrate how these branches intertwine in practice—combining sophisticated search with learned representations and large-scale data.

Within this landscape, adversarial and curriculum learning approaches represent a particularly active line of inquiry, seeking to guide exploration toward increasingly challenging or informative subgoals rather than relying on static datasets. GAR[0] exemplifies this direction by dynamically adjusting training difficulty, contrasting with works like Longer Proofs[45] that focus on scaling search depth, or Loop Invariant Synthesis[48] which applies RL to program verification subtasks. Meanwhile, Learning Interestingness[22] explores how to prioritize novel or underexplored proof states, a theme that complements GAR[0]'s curriculum emphasis but tackles the problem from a different angle—rewarding exploration of rare trajectories rather than structured difficulty progression. These curriculum-driven methods sit at the intersection of core RL innovation and proof search strategy, addressing the perennial challenge of sample efficiency and generalization in domains where supervision is sparse and search spaces are vast.

Related Works in Same Category

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

1. Learning Interestingness in Automated Mathematical Theory Formation

Authors: George Tsoukalas, Rahul Saha, Amitayush Thakur, Sabrina Reguyal, Swarat Chaudhuri | **Year/Venue:** 2025 | **URL:** [View paper](#)

Abstract

We take two key steps in automating the open-ended discovery of new mathematical theories, a grand challenge in artificial intelligence. First, we introduce $\text{\emph{FERMAT}}$, a reinforcement learning (RL) environment that models concept discovery and theorem-proving using a set of symbolic actions, opening up a range of RL problems relevant to theory discovery. Second, we explore a specific problem through $\text{\emph{FERMAT}}$: automatically scoring the $\text{\emph{interestingness}}$ of mathematical objects...

Relationship Analysis

Both papers belong to the Adversarial and Curriculum Learning category, employing co-evolutionary training approaches in formal theorem proving. While GAR focuses on adversarial reinforcement learning between a problem composer (statement fuser) and prover to dynamically adjust theorem difficulty in Lean4, FERMAT introduces a reinforcement learning environment for automated mathematical theory formation with evolutionary algorithms for learning interestingness measures. The key difference is that GAR targets proof generation with adversarial training loops, whereas FERMAT addresses the broader challenge of open-ended mathematical concept discovery and theory formation.

2. Towards finding longer proofs

Authors: Zombori, Zsolt, CsizsÁrrik, Adrián, Zsolt Zombori, et al. (15 authors total) | **Year/Venue:** 2021 | **URL:** [View paper](#)

Abstract

We present a reinforcement learning (RL) based guidance system for automated theorem proving geared towards Finding Longer Proofs (FLoP). FLoP focuses on generalizing from short proofs to longer ones of similar structure. To achieve that, FLoP uses state-of-the-art RL approaches that were previously not applied in theorem proving. In particular, we show that curriculum learning significantly outperforms previous learning-based proof guidance on a synthetic dataset of increasingly difficult arith...

Relationship Analysis

Both papers belong to the Adversarial and Curriculum Learning category, employing curriculum learning to progressively adapt problem difficulty to the prover's capability. While the original paper (GAR) uses adversarial training between a problem composer (statement fuser) and prover to co-evolve both components jointly, the candidate paper (FLoP) focuses on curriculum learning with a fixed problem set, starting proofs near completion and gradually moving backwards, without adversarial generation of new problems.

3. Learning to find proofs and theorems by learning to refine search strategies: The case of loop invariant synthesis

Authors: Laurent, Jonathan, Platzer, AndrÁ© | **Year/Venue:** 2022 | **URL:** [View paper](#)

Abstract

We propose a new approach to automated theorem proving where an AlphaZero-style agent is self-training to refine a generic high-level expert strategy expressed as a nondeterministic program. An analogous teacher agent is self-training to generate tasks of suitable relevance and difficulty for the learner. This allows leveraging minimal amounts of domain knowledge to tackle problems for which training data is unavailable or hard to synthesize. As a specific illustration, we consider loop invarian...

Relationship Analysis

Both papers belong to the Adversarial and Curriculum Learning category, employing co-evolution of problem generation and solving capabilities through adversarial training loops. The original paper (GAR) focuses on formal theorem proving in Lean with a statement fuser that generates harder theorems to train a prover, while the candidate paper addresses loop invariant synthesis for imperative programs using a teacher agent that generates verification tasks to train a solver agent. The key difference is that GAR operates on mathematical theorem proving with natural language statement fusion and autoformalization, whereas the candidate paper targets program verification with nondeterministic search strategies expressed as programs.

Contributions Analysis

Overall novelty summary. The paper proposes GAR, a generative adversarial reinforcement learning framework that jointly trains a problem composer and solver in an adversarial loop with implicit curriculum learning. This work resides in the 'Adversarial and Curriculum Learning' leaf of the taxonomy, which contains four papers total including the original submission. This leaf sits within the broader 'Core RL Algorithms and Training Frameworks' branch, indicating a moderately populated research direction focused on training paradigms rather than proof search mechanics or neural architectures. The taxonomy reveals this is an active but not overcrowded area, with sibling leaves exploring policy optimization, reward modeling, and alternative RL paradigms.

The taxonomy structure shows GAR's leaf neighbors include 'Policy Optimization and Expert Iteration' (five papers on PPO and expert iteration methods) and 'Reward Modeling and Verifier Integration' (four papers on critic models and formal verifiers). The scope note for GAR's leaf explicitly covers 'adversarial training loops or curriculum learning to co-evolve problem generation and solving capabilities,' distinguishing it from standard supervised or policy gradient approaches in adjacent categories. Nearby branches address 'Proof Search Strategies' (MCTS, hierarchical decomposition) and 'Data Generation' (synthetic problem generation, autoformalization), suggesting GAR bridges training methodology with data creation concerns that typically occupy separate research threads.

Among fifteen candidates examined across three contributions, only one refutable pair emerged. The core GAR framework contribution examined zero candidates (likely due to its novelty as an integrated system). Statement Fusion for generating formal theorems examined five candidates with no refutations, suggesting this technical component has limited direct overlap in the search scope. The general RL paradigm for co-evolution examined ten candidates and found one potential refutation, indicating some conceptual precedent exists within the limited search. The statistics suggest the framework's integration of adversarial training with curriculum learning in theorem proving is relatively unexplored among the top-fifteen semantic matches, though the search scope cannot confirm exhaustive novelty.

Based on thirty candidates examined through semantic search and citation expansion, the work appears to occupy a sparsely populated intersection of adversarial training and formal theorem proving. The taxonomy confirms this direction has fewer papers than policy optimization or proof search categories, and the low refutation rate across contributions aligns with this positioning. However, the limited search scope means potentially relevant work in adjacent areas (e.g., curriculum learning in general RL, adversarial training in other domains) may not have been captured.

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

Contribution 1: GAR: Generative Adversarial Reinforcement Learning framework

Description: The authors introduce GAR, a novel reinforcement learning framework that simultaneously optimizes both a theorem prover and a problem composer (statement fuser) through adversarial training. This joint optimization establishes an implicit curriculum learning mechanism that dynamically adjusts problem difficulty to match the prover's evolving capabilities, improving training efficiency and enabling stronger performance on advanced theorems.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

Contribution 2: Statement Fusion technique for generating formal theorems

Description: The authors develop a statement fusion method that combines pairs of natural language mathematical statements to create more challenging problems. This technique deliberately separates natural language fusion from formal language formalization, allowing the generation of progressively harder theorems that adapt to the prover's current skill level rather than relying on fixed problem sets.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Tree-Based Premise Selection for Lean4

URL: [View paper](#)

Brief Assessment

Tree-Based Premise Selection[55] focuses on premise retrieval using structural tree representations and similarity metrics (WL kernel, tree edit distance). It does not address theorem generation or statement fusion techniques.

2. Adaptive human-machine theorem proving system

URL: [View paper](#)

Brief Assessment

Adaptive Human-Machine[53] focuses on human-machine collaboration in theorem proving with adaptive learning from expert interactions, not on generating formal theorems through statement fusion. The candidate discusses training systems to adapt to expert reasoning styles rather than fusing mathematical statements to create progressively harder problems.

3. DRIFT: Decompose, Retrieve, Illustrate, then Formalize Theorems

URL: [View paper](#)

Brief Assessment

DRIFT[54] focuses on decomposing informal mathematical statements into sub-queries for retrieval-augmented autoformalization, not on fusing pairs of statements to generate progressively harder theorems for training provers.

4. Multimodal Extraction of Proofs and Theorems from the Scientific Literature

URL: [View paper](#)

Brief Assessment

Multimodal Extraction[52] focuses on extracting existing mathematical statements and proofs from PDF documents, not on generating new formal theorems by fusing statements. The candidate addresses document information extraction, while the original contribution concerns synthetic theorem generation for training.

5. StepFun-Formalizer: Unlocking the Autoformalization Potential of LLMs through Knowledge-Reasoning Fusion

URL: [View paper](#)

Brief Assessment

StepFun-Formalizer[51] focuses on autoformalization (translating natural language to formal language) rather than generating new formal theorems through statement fusion. The candidate addresses informal-to-formal translation with reasoning trajectories, not the adaptive fusion of statement pairs to create progressively harder problems as in the original paper.

Contribution 3: General RL paradigm for co-evolution in verifiable environments

Description: Beyond formal theorem proving, the authors establish GAR as a general reinforcement learning paradigm where problem generators and solvers co-evolve through adversarial training in environments with automatic verification. This framework provides a foundation for applying similar adversarial co-training approaches to other reasoning-intensive domains that have verifiable outcomes.

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. Scenario co-evolution for reinforcement learning on a grid world smart factory domain

URL: [View paper](#)

Brief Assessment

Scenario Co-evolution[63] focuses on a grid world smart factory domain with evolutionary algorithms for scenario generation, not formal theorem proving with verifiable formal languages. The domains and verification mechanisms differ fundamentally.

2. Generative Adversarial Reasoner: Enhancing LLM Reasoning with Adversarial Reinforcement Learning

URL: [View paper](#)

Prior Art Analysis

Generative Adversarial Reasoner[59] demonstrates a prior implementation of co-evolution between problem generators (discriminators) and solvers (reasoners) through adversarial reinforcement learning in verifiable environments. The candidate paper explicitly describes joint training of a discriminator and reasoner through adversarial RL, where both components co-evolve: the discriminator evaluates reasoning steps and the reasoner learns to produce steps the discriminator endorses. This establishes the same general paradigm of adversarial co-training in verifiable settings (mathematical reasoning with automatic verification) that the original paper claims as novel.

Evidence

Evidence 1 - **Rationale:** Both papers describe adversarial co-training frameworks where a generator/reasoner and evaluator/discriminator co-evolve through RL in verifiable environments. The candidate explicitly implements joint training with adversarial dynamics. - **Original:** furthermore, thegarcontributes a general rl paradigm for the co-evolution of problem generation and solving in verifiable environments, which offers new directions for adversarial co-training in other reasoning-intensive domains. - **Candidate:** we proposegenerative adversarial reasoner (gar), which incorporates a compute-efficient review schedule and an adversarial co-training framework. specifically, we partition each reasoning chain into logically complete slices of comparable length. the discriminator evaluates each slice for logical so...

Evidence 2 - **Rationale:** Both papers describe joint optimization of two components through adversarial training with dynamic co-adaptation, establishing the same general paradigm of co-evolution. - **Original:** garjointly optimizes the prover and the problem composer, termed as the statement fuser, through adversarial training. this process establishes an implicit curriculum learning that dynamically adjusts statement difficulty to match the prover's development. - **Candidate:** concretely, we optimize the llm reasoner and an llm-based discriminator together: the discriminator judges the logical soundness of each intermediate reasoning step and explains its judgment, while the reasoner learns to produce steps the discriminator consistently endorses for valid logic. this co...

Evidence 3 - **Rationale:** The candidate demonstrates the same general paradigm of joint training between evaluator and solver in verifiable environments (mathematical reasoning with automatic answer verification), establishing prior work in this area. - **Original:** beyond theorem proving,garoffers a general rl paradigm of co-evolution of the problem generation and solving under a verifiable environment. it provides a foundation for adversarial co-training in other reasoning-intensive domains. - **Candidate:** our approach jointly trains the reasoner and a discriminator under an on-policy scheme, yielding dense slice-level rewards with a computeefficient review schedule and alignment regularization to curb reward hacking, addressing prm mis-specification and noisy signals without sacrificing efficiency

3. A Co-Evolution Algorithm With Dueling Reinforcement Learning Mechanism for the Energy-Aware Distributed Heterogeneous Flexible Flow-Shop Scheduling Problem

URL: [View paper](#)

Brief Assessment

Co-Evolution Dueling RL[58] focuses on production scheduling optimization in manufacturing systems, not on co-evolution of problem generation and solving in verifiable reasoning environments like formal theorem proving.

4. AdvEvo-MARL: Shaping Internalized Safety through Adversarial Co-Evolution in Multi-Agent Reinforcement Learning

URL: [View paper](#)

Brief Assessment

AdvEvo-MARL[62] focuses on safety in multi-agent systems through adversarial co-evolution of attackers and defenders, not on problem generation and solving in verifiable environments like formal theorem proving.

5. Redefining Adversarial Dynamics: Co-Evolution of Attack and Defense Strategies in AI-Enabled Power Cyber-Physical Systems

URL: [View paper](#)

Brief Assessment

Adversarial Co-Evolution[65] focuses on cybersecurity in power cyber-physical systems with game-theoretic adversarial dynamics, not on reinforcement learning paradigms for problem generation and solving in verifiable environments like formal theorem proving.

6. Adversarial image generation using evolution and deep learning

URL: [View paper](#)

Brief Assessment

Adversarial Image Evolution[61] focuses on artist-critic coevolution for image generation using evolutionary computation and CNNs, not reinforcement learning. The original paper establishes a general RL paradigm for formal theorem proving with verifiable outcomes, which is fundamentally different from image generation tasks without formal verification.

7. EvaDrive: Evolutionary Adversarial Policy Optimization for End-to-End Autonomous Driving

URL: [View paper](#)

Brief Assessment

EvaDrive[57] focuses on autonomous driving trajectory planning through adversarial multi-objective RL, not formal theorem proving or verifiable reasoning environments. The domains and verification mechanisms are fundamentally different.

8. Reinforcement Learning for Human-AI Collaboration: Challenges, Mechanisms, and Methods

URL: [View paper](#)

Brief Assessment

RL Human-AI Collaboration[60] discusses adversarial reinforcement learning and co-evolution in multi-agent contexts, but does not address verifiable environments or the specific paradigm of problem generator-solver co-evolution through adversarial training with automatic verification that the original paper establishes.

9. Learning to generate unit test via adversarial reinforcement learning

URL: [View paper](#)

Brief Assessment

Unit Test Generation[56] focuses on adversarial co-training of unit test generators and code generators for software testing, not formal theorem proving. The verifiable environment (unit test execution) and domain (programming) differ fundamentally from GAR's formal mathematical reasoning context.

10. Machine learning in adversarial environments

URL: [View paper](#)

Brief Assessment

ML Adversarial Environments[64] focuses on adversarial manipulation of machine learning systems in security contexts (spam filtering, intrusion detection, web spam), not on co-evolution of problem generators and solvers through reinforcement learning in verifiable environments like formal theorem proving.

Appendix: Text Similarity Detection

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

References

- [0] GAR: Generative Adversarial Reinforcement Learning for Formal Theorem Proving [View paper](#)
- [1] Olympiad-level formal mathematical reasoning with reinforcement learning. [View paper](#)
- [2] Reinforcement learning for guiding the e theorem prover [View paper](#)
- [3] Reinforcement learning of theorem proving [View paper](#)
- [4] Verification using formalised mathematics and theorem proving of reinforcement and deep learning [View paper](#)
- [5] QEDCartographer: Automating formal verification using reward-free reinforcement learning [View paper](#)
- [6] Seed-Prover: Deep and Broad Reasoning for Automated Theorem Proving [View paper](#)
- [7] DeepTheorem: Advancing LLM Reasoning for Theorem Proving Through Natural Language and Reinforcement Learning [View paper](#)
- [8] ABEL: Sample efficient online reinforcement learning for neural theorem proving [View paper](#)
- [9] Deepseek-prover-v2: Advancing formal mathematical reasoning via reinforcement learning for subgoal decomposition [View paper](#)
- [10] CriticLean: Critic-Guided Reinforcement Learning for Mathematical Formalization [View paper](#)
- [11] Formal Theorem Proving by Rewarding LLMs to Decompose Proofs Hierarchically [View paper](#)
- [12] Search Strategy Selection for Automated Theorem Proving [View paper](#)
- [13] A deep reinforcement learning approach to first-order logic theorem proving [View paper](#)
- [14] Neural Theorem Proving: Generating and Structuring Proofs for Formal Verification [View paper](#)
- [15] Leanabell-prover-v2: Verifier-integrated reasoning for formal theorem proving via reinforcement learning [View paper](#)
- [16] A Combinatorial Identities Benchmark for Theorem Proving via Automated Theorem Generation [View paper](#)
- [17] Holist: An environment for machine learning of higher order logic theorem proving [View paper](#)
- [18] ProofNet++: A Neuro-Symbolic System for Formal Proof Verification with Self-Correction [View paper](#)
- [19] Leanabell-Prover: Posttraining Scaling in Formal Reasoning [View paper](#)
- [20] Re:Form - Reducing Human Priors in Scalable Formal Software Verification with RL in LLMs: A Preliminary Study on Dafny [View paper](#)
- [21] Local Look-Ahead Guidance via Verifier-in-the-Loop for Automated Theorem Proving [View paper](#)
- [22] Learning Interestingness in Automated Mathematical Theory Formation [View paper](#)
- [23] An Ensemble Approach for Automated Theorem Proving Based on Efficient Name Invariant Graph Neural Representations [View paper](#)
- [24] RL Tango: Reinforcing Generator and Verifier Together for Language Reasoning [View paper](#)
- [25] Markov Decision Processes with Formal Verification: Mathematical Guarantees for Safe Reinforcement Learning [View paper](#)
- [26] FormaRL: Enhancing Autoformalization with no Labeled Data [View paper](#)
- [27] Kimina-Prover Preview: Towards Large Formal Reasoning Models with Reinforcement Learning [View paper](#)
- [28] Learning equational theorem proving [View paper](#)
- [29] Deepseek-prover-v1.5: Harnessing proof assistant feedback for reinforcement learning and monte-carlo tree search [View paper](#)
- [30] Proof Flow: Preliminary Study on Generative Flow Network Language Model Tuning for Formal Reasoning [View paper](#)
- [31] Exploring and Evaluating Interplays of BPPy with Deep Reinforcement Learning and Formal Methods [View paper](#)
- [32] Reinforcement Learning in Mathematical Research [View paper](#)
- [33] Formal methods meet machine learning (F3ML) [View paper](#)
- [34] Project proposal: A modular reinforcement learning based automated theorem prover [View paper](#)
- [35] Goedel-Prover-V2: Scaling Formal Theorem Proving with Scaffolded Data Synthesis and Self-Correction [View paper](#)
- [36] Tacticzero: Learning to prove theorems from scratch with deep reinforcement learning [View paper](#)
- [37] Project proposal: A modular reinforcement learning based automated theorem prover [View paper](#)
- [38] Proof2Silicon: Prompt Repair for Verified Code and Hardware Generation via Reinforcement Learning [View paper](#)
- [39] Peano: learning formal mathematical reasoning [View paper](#)
- [40] EconProver: Towards More Economical Test-Time Scaling for Automated Theorem Proving [View paper](#)
- [41] STP: Self-play LLM Theorem Provers with Iterative Conjecturing and Proving [View paper](#)
- [42] Goedel-Prover: A Frontier Model for Open-Source Automated Theorem Proving [View paper](#)
- [43] Mathesis: Towards Formal Theorem Proving from Natural Languages [View paper](#)
- [44] LeanConjecturer: Automatic Generation of Mathematical Conjectures for Theorem Proving [View paper](#)
- [45] Towards finding longer proofs [View paper](#)
- [46] Prolog Technology Reinforcement Learning Prover: (System Description) [View paper](#)
- [47] Automated Theorem Proving in Intuitionistic Propositional Logic by Deep Reinforcement Learning [View paper](#)
- [48] Learning to find proofs and theorems by learning to refine search strategies: The case of loop invariant synthesis [View paper](#)
- [49] Rewarding the Unlikely: Lifting GRPO Beyond Distribution Sharpening [View paper](#)
- [50] Gym-saturation: an OpenAI Gym environment for saturation provers [View paper](#)
- [51] StepFun-Formalizer: Unlocking the Autoformalization Potential of LLMs through Knowledge-Reasoning Fusion [View paper](#)
- [52] Multimodal Extraction of Proofs and Theorems from the Scientific Literature [View paper](#)
- [53] Adaptive human-machine theorem proving system [View paper](#)
- [54] DRIFT: Decompose, Retrieve, Illustrate, then Formalize Theorems [View paper](#)
- [55] Tree-Based Premise Selection for Lean4 [View paper](#)
- [56] Learning to generate unit test via adversarial reinforcement learning [View paper](#)
- [57] EvaDrive: Evolutionary Adversarial Policy Optimization for End-to-End Autonomous Driving [View paper](#)
- [58] A Co-Evolution Algorithm With Dueling Reinforcement Learning Mechanism for the Energy-Aware Distributed Heterogeneous Flexible Flow-Shop Scheduling Problem [View paper](#)

- [59] Generative Adversarial Reasoner: Enhancing LLM Reasoning with Adversarial Reinforcement Learning [View paper](#)
- [60] Reinforcement Learning for Human-AI Collaboration: Challenges, Mechanisms, and Methods [View paper](#)
- [61] Adversarial image generation using evolution and deep learning [View paper](#)
- [62] AdvEvo-MARL: Shaping Internalized Safety through Adversarial Co-Evolution in Multi-Agent Reinforcement Learning [View paper](#)
- [63] Scenario co-evolution for reinforcement learning on a grid world smart factory domain [View paper](#)
- [64] Machine learning in adversarial environments [View paper](#)
- [65] Redefining Adversarial Dynamics: Co-Evolution of Attack and Defense Strategies in AI-Enabled Power Cyber-Physical Systems [View paper](#)