Novelty Assessment Report

Paper: LogicSR: A Unified Benchmark for Logical Discovery from Data

PDF URL: https://openreview.net/pdf?id=8ixcRzGuff

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-08

Abstract

Discovering underlying logical expressions from data is a critical task for interpretable AI and scientific discovery, yet it remains poorly served by existing research infrastructure. The field of Symbolic Regression (SR) primarily focuses on continuous mathematical functions, while Logic Synthesis (LS) is designed for exact, noise-free specifications, not for learning from incomplete or noisy data. This leaves a crucial gap for evaluating algorithms that can learn generalizable logical rules in realistic scenarios. To address this, we introduce LogicSR, a large-scale and comprehensive benchmark for logical symbolic regression. LogicSR is built from two sources: real-world problems from digital circuits and biological networks, and a novel synthetic data generator capable of producing a diverse set of complex logical formulas at scale. We use LogicSR to conduct a rigorous evaluation of 17 algorithms, spanning classical logic solvers, modern machine learning models, and Large Language Models (LLMs). Our findings reveal that the logical modeling capabilities and generalization robustness of these algorithms significantly depend on task scale and logical complexity, with current cutting-edge LLMs showing limited complex logical reasoning ability. LogicSR provides a robust foundation to benchmark progress, unify evaluation across disparate fields, and steer the future development of powerful neuro-symbolic systems.

Core Task Landscape

This paper addresses: **Discovering Logical Expressions from Data**

A total of **50 papers** were analyzed and organized into a taxonomy with **14 categories**.

Taxonomy Overview

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

- **Symbolic Regression and Formula Discovery**
- **Logical Rule Learning and Reasoning**
- **Deep Learning for Logical Tasks**
- **Applications and Domain-Specific Methods**

Complete Taxonomy Tree

- Discovering Logical Expressions from Data Survey Taxonomy
- Symbolic Regression and Formula Discovery
  - Neural-Guided Symbolic Regression (3 papers)
  - [10] Neural symbolic regression that scales (Luca Biggio, 2021) View paper
  - [28] Advanced Weakly-Supervised Formula Exploration for Neuro-Symbolic Mathematical Reasoning (Wu Yuxuan, 2025) View paper
  - [47] Generative pre-trained transformer for symbolic regression base in-context reinforcement learning (Li YanJie, 2024) View paper
  - Physics-Informed Symbolic Discovery (7 papers)
  - [3] Discovering Symbolic Models from Deep Learning with Inductive Biases (Miles Cranmer, 2022) View paper
  - [6] A neural symbolic model for space physics (Jie Ying, 2025) View paper
  - [17] Learning symbolic models for graph-structured physical mechanism (H Shi, 2022) View paper
  - [20] Machine learning symbolic equations for diffusion with physics-based descriptions (Konstantinos Papastamatiou, 2022) View paper
  - [25] Finite Expression Methods for Discovering Physical Laws from Data (Zhongyi Jiang, 2023) View paper
  - [40] Kolmogorov-Arnold Representation for Symplectic Learning: Advancing Hamiltonian Neural Networks (Wu Zongyu, 2025) View paper
  - [42] -Analyses with Symbolic Regression (H Bahl, 2025) View paper
  - Classical and Hybrid Symbolic Methods (3 papers)
  - [26] A coupled framework for symbolic turbulence models from deep-learning (Chitrarth Lav, 2023) View paper
  - [35] SYMBOL: Generating flexible black-box optimizers through symbolic equation learning (Chen Jiacheng, 2024) View paper
  - [44] Combining data and theory for derivable scientific discovery with AI-Descartes (Cristina Cornelio, 2023) View paper
- Logical Rule Learning and Reasoning
  - Knowledge Graph Rule Learning (6 papers)
  - [5] Differentiable Learning of Logical Rules for Knowledge Base Reasoning (Yang Fan, 2022) View paper
  - [7] ChatRule: Mining Logical Rules with Large Language Models for Knowledge Graph Reasoning (Linhao Luo, 2025) View paper
  - [19] Rlogic: Recursive logical rule learning from knowledge graphs (Kewei Cheng, 2022) View paper
  - [21] Logical Rule Learning (Kewei Cheng, 2024) View paper
  - [33] Bi-directional Learning of Logical Rules with Type Constraints for Knowledge Graph Completion (Kunxun Qi, 2024) View paper
  - [43] Leveraging Logical Rules in Knowledge Editing: A Cherry on the Top (Ali, 2024) View paper
  - Temporal Logic Learning (4 papers)

- [2] Learning temporal formulas from examples is hard (Mascle, 2023) View paper
- [11] TILP: Differentiable Learning of Temporal Logical Rules on Knowledge Graphs (Xiong, 2024) View paper
- [16] Retrieval-Augmented Mining of Temporal Logic Specifications from Data (Gaia Saveri, 2024) View paper
- [34] TLogic: Temporal Logical Rules for Explainable Link Forecasting on Temporal Knowledge Graphs (Hildebrandt, 2022) View paper
- Neuro-Symbolic Integration for Reasoning (7 papers)
- [9] Deep Symbolic Learning: Discovering Symbols and Rules from Perceptions (Alessandro Daniele, 2023) View paper
- [15] AI Reasoning in Deep Learning Era: From Symbolic AI to Neuralâ Symbolic AI (Baoyu Liang, 2025) View paper
- [23] Logic Tensor Networks: Deep Learning and Logical Reasoning from Data and Knowledge (Luciano Serafini, 2022) View paper
- [24] Harnessing Deep Neural Networks with Logic Rules (Zhiting Hu, 2022) View paper
- [27] Towards Deep Symbolic Reinforcement Learning (Garnelo, 2022) View paper
- [30] Probabilistic logic neural networks for reasoning (Qu Meng, 2019) View paper
- [48] Neuro-Symbolic Class Expression Learning (Ã aÄ lar Demir, 2023) View paper
- Classical Logic Induction and Mining (2 papers)
- [4] Discovering Classification Rules Using Variable-Valued Logic System VL1 (Michalski, 2025) View paper
- [49] Establishing logical rules from empirical data (John L. PFALTZ, 2008) View paper

- Deep Learning for Logical Tasks
  - Semantic Parsing and Logical Form Inference (2 papers)
  - [1] Inferring Logical Forms From Denotations (Panupong Pasupat, 2022) View paper
  - [14] Generating predicate logic expressions from natural language (Oleksii Levkovskyi, 2021) View paper
  - Neural Logical Reasoning (5 papers)
  - [13] Selection-inference: Exploiting large language models for interpretable logical reasoning (Creswell, 2022) View paper
  - [31] Exploring logical reasoning for referring expression comprehension (Ying Cheng, 2021) View paper
  - [37] Can llms reason with rules? logic scaffolding for stress-testing and improving llms (Choi, 2024) View paper
  - [45] Towards logiglue: A brief survey and a benchmark for analyzing logical reasoning capabilities of language models (Luo Man, 2023) View paper
  - [46] Neural logic reasoning (Shi Shaoyun, 2020) View paper
  - Deep Learning for Symbolic Mathematics (1 papers)
  - [41] Deep Learning for Symbolic Mathematics (Lample, 2022) View paper
- Applications and Domain-Specific Methods
  - Document and Text Analysis (2 papers)
  - [8] Rule by example: Harnessing logical rules for explainable hate speech detection (Chen Mei, 2024) View paper
  - [29] End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction (Du Jianfeng, 2024) View paper
  - Reinforcement Learning and Optimization (2 papers)
  - [18] Fortune: Formula-Driven Reinforcement Learning for Symbolic Table Reasoning in Language Models (Cao Lang, 2025) View paper
  - [39] Learning to Explore Paths for Symbolic Execution (Jingxuan He, 2021) View paper
  - Specialized Domain Applications ★ (5 papers)
  - [0] LogicSR: A Unified Benchmark for Logical Discovery from Data (Anon et al., 2026) View paper
  - [12] Mining logical arithmetic expressions from proper representations (Eitan Kosman, 2022) View paper
  - [22] Mining logical circuits in fungi (Nic Roberts, 2022) View paper
  - [36] Using symbolic machine learning to assess and model substance transport and decay in water distribution networks (Daniele Laucelli, 2024) View paper
  - [50] SWARMFLAWFINDER: Discovering and Exploiting Logic Flaws of Swarm Algorithms (Chi-jung Jung, 2022) View paper
  - Fairness and Interpretability (2 papers)
  - [32] Syllogistic reasoning: Logical decisions from a complex data base (Falmagne, 2015) View paper
  - [38] Proxy Attribute Discovery in Machine Learning Datasets via Inductive Logic Programming (Rafael GonÃ§alves, 2025) View paper

Narrative

Core task: Discovering logical expressions from data. The field encompasses a diverse set of approaches organized into four main branches. Symbolic Regression and Formula Discovery focuses on extracting mathematical and symbolic models directly from observations, often leveraging evolutionary algorithms or neural architectures to uncover closed-form equations. Logical Rule Learning and Reasoning emphasizes the induction of interpretable logical rules—ranging from propositional to first-order logic—that capture underlying patterns in structured or relational data. Deep Learning for Logical Tasks explores how neural networks can be trained to perform or assist in logical inference, blending differentiable modules with symbolic reasoning. Finally, Applications and Domain-Specific Methods tailors these techniques to specialized domains such as temporal logic synthesis, biological systems, water network modeling, and software verification, demonstrating how domain constraints guide the discovery process.

Recent work highlights contrasting trade-offs between interpretability and expressiveness. Some studies pursue purely symbolic outputs for transparency, while others integrate neural components to handle noisy or high-dimensional data at the cost of reduced clarity. Within the Applications and Domain-Specific Methods branch, LogicSR[0] sits alongside efforts like Logical Circuits Fungi[22] and Symbolic Water Networks[36], which apply logical or symbolic discovery to concrete engineering and biological problems. Compared to Mining Logical Arithmetic[12], which targets arithmetic rule extraction, LogicSR[0] emphasizes a broader logical framework suitable for diverse application contexts. Meanwhile, SwarmFlawFinder[50] illustrates how domain-specific heuristics can guide search in software analysis. Overall, LogicSR[0] occupies a niche where domain expertise and logical structure converge, bridging general symbolic regression methods with the practical demands of specialized fields.

Related Works in Same Category

The following **4 sibling papers** share the same taxonomy leaf node with the original paper:

1. Mining logical arithmetic expressions from proper representations

Authors: Eitan Kosman, Ilya Kolchinsky, Assaf Schuster, I. Kolchinsky, A. Schuster | **Year/Venue**: 2022 | **URL**: View paper

Abstract

Logical-arithmetic expressions are convenient for describing phenomena due to their expressiveness and comprehensibility. Therefore, we propose to target mining logical arithmetic expressions through a novel task called Logical-arithmetic expression mining (LAEM). Its goal is to discover expressive logical expressions that are representative for a database. It accepts a complex database as input and returns a set of representative expressions for the database. Driven by the success of machine le...

Relationship Analysis

Both papers belong to the Specialized Domain Applications category, focusing on discovering logical expressions from data in specific domains. While the original paper (LogicSR) presents a comprehensive benchmark for logical symbolic regression targeting Boolean functions in digital circuits and biological networks with extensive evaluation of 17 algorithms, the candidate paper (SEEN) proposes a specific mining algorithm using soft decision trees and representation learning to discover logical-arithmetic expressions from complex databases. The key difference is that LogicSR provides evaluation infrastructure and comparative analysis across methods, whereas SEEN contributes a novel algorithmic approach for expression discovery.

2. Mining logical circuits in fungi

Authors: Nic Roberts, Andrew Adamatzky | **Year/Venue**: 2022 | **URL**: View paper

Abstract

Living substrates are capable for nontrivial mappings of electrical signals due to the substrate nonlinear electrical characteristics. This property can be used to realise Boolean functions. Input logical values are represented by amplitude or frequency of electrical stimuli. Output logical values are decoded from electrical responses of living substrates. We demonstrate how logical circuits can be implemented in mycelium bound composites. The mycelium bound composites (fungal materials) are get...

Relationship Analysis

Both papers belong to the Specialized Domain Applications category, focusing on logical expression discovery in physical and biological systems. While the original paper (LogicSR) presents a comprehensive benchmark for discovering logical expressions from data across multiple domains including digital circuits and biological networks, the candidate paper demonstrates a specific experimental implementation of Boolean logic circuits in fungal mycelium composites. The key difference is that LogicSR provides a unified evaluation framework and benchmark suite for comparing logical discovery algorithms, whereas the candidate paper focuses on a novel biological substrate (fungi) for physically implementing and mining logical circuits through electrical signal transformation.

3. Using symbolic machine learning to assess and model substance transport and decay in water distribution networks

Authors: Daniele Laucelli, Laura Enriquez, Juan Saldarriaga, Orazio Giustolisi | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Drinking water infrastructures are systems of pipes which are generally networked. They play a crucial role in transporting and delivering clean water to people. The water quality analysis refers to the evaluation of the advective diffusion of any substance in drinking water infrastructures from source nodes. Such substances could be a contamination for the system or planned for the disinfection, e.g., chlorine. The water quality analysis is performed by integrating the differential equation in ...

Relationship Analysis

Both papers belong to the Specialized Domain Applications category, focusing on logical expression discovery for domain-specific systems. While the original paper (LogicSR) develops a unified benchmark for discovering logical expressions from data across digital circuits and biological networks using symbolic regression and machine learning methods, the candidate paper applies symbolic machine learning (Evolutionary Polynomial Regression) specifically to model substance transport and decay in water distribution networks. The key difference is that LogicSR provides a comprehensive benchmarking framework for evaluating logical discovery algorithms across multiple domains, whereas the candidate paper applies one specific symbolic learning technique (EPR) to solve a particular engineering problem in water infrastructure systems.

4. SWARMFLAWFINDER: Discovering and Exploiting Logic Flaws of Swarm Algorithms

Authors: Chi-jung Jung, Ali Ahad, Chi-Gon Jung, Yuseok Jeon, A. Ahad, et al. (6 authors total) | **Year/Venue**: 2022 | **URL**: View paper

Abstract

Inspired by swarms in nature, swarm robotics have been developed to conduct various challenging tasks such as environmental monitoring, disaster recovery, logistics, and even military operations. Despite the significant potential impact of the swarm on society, relatively little attention is given to adversarial scenarios against swarm robotics. In this paper, we explore a systematic approach to find logical flaws of the swarm robotics algorithms that adversaries can exploit. Specifically, we de...

Relationship Analysis

Both papers belong to the Specialized Domain Applications category, focusing on logical expression discovery in specific technical domains. While LogicSR develops a benchmark for discovering logical expressions from data in digital circuits and biological networks using symbolic regression and machine learning approaches, SWARMFLAWFINDER focuses on discovering and exploiting logical flaws in swarm robotics algorithms through automated testing and fuzzing techniques. The key difference is that LogicSR aims to learn generalizable logical rules from noisy/incomplete data for interpretable AI, whereas SWARMFLAWFINDER targets adversarial testing to find exploitable logic bugs in existing swarm control algorithms.

Contributions Analysis

Overall novelty summary. The paper introduces LogicSR, a benchmark for learning logical expressions from noisy, incomplete data—a task distinct from continuous symbolic regression and exact logic synthesis. Within the taxonomy, it resides in 'Specialized Domain Applications' alongside four sibling papers addressing biological networks, water distribution systems, and circuit design. This leaf represents a relatively sparse research direction (five papers total) focused on applying logical discovery to concrete engineering and scientific domains, suggesting the work targets an underserved niche rather than a crowded methodological space.

The taxonomy reveals that most logical expression discovery research concentrates in two neighboring branches: 'Symbolic Regression and Formula Discovery' (ten papers across three sub-areas) and 'Logical Rule Learning and Reasoning' (nineteen papers spanning knowledge graphs, temporal logic, and neuro-symbolic integration). LogicSR bridges these areas by addressing logical (not continuous) formulas while handling noisy data (unlike exact logic synthesis). The benchmark's dual focus on real-world circuits/biological networks and synthetic generation distinguishes it from purely domain-specific methods in its leaf and from general-purpose symbolic regression approaches that lack logical structure.

Among thirty candidates examined, none clearly refute the three core contributions. The LogicSR benchmark itself (ten candidates, zero refutations) appears novel as a dedicated evaluation framework for logical symbolic regression under noise. The synthetic data generator (ten candidates, zero refutations) shows no direct prior work in the limited search scope, though the analysis does not cover exhaustive generation literature. The cross-domain evaluation of seventeen algorithms (ten candidates, zero refutations) represents a substantial empirical effort, with no overlapping multi-algorithm comparisons identified in the examined papers. These statistics suggest originality within the search scope, though the limited candidate pool (thirty total) means undiscovered prior work remains possible.

Based on the top-thirty semantic matches and taxonomy structure, the work addresses a genuine gap between continuous symbolic regression and exact logic synthesis. The benchmark's combination of real-world and synthetic logical tasks, evaluated across classical solvers, ML models, and LLMs, appears distinctive within the examined literature. However, the analysis covers a narrow slice of potential prior work—broader searches in logic synthesis, program synthesis, or SAT-based learning communities might reveal additional relevant baselines or evaluation frameworks not captured here.

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

Contribution 1: LogicSR benchmark for logical symbolic regression

Description: The authors present LogicSR, a unified benchmark designed to evaluate algorithms that discover logical expressions from data. It combines real-world problems from digital circuits and biological networks with a novel synthetic data generator, addressing the gap between continuous symbolic regression and exact logic synthesis.

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. Large language models meet symbolic provers for logical reasoning evaluation

URL: View paper

Brief Assessment

Symbolic Provers LLMs[73] focuses on first-order logic (FOL) reasoning evaluation with chain-of-thought prompting, not on discovering logical expressions from data. The candidate addresses reasoning evaluation, while the original addresses symbolic regression for learning logical rules from incomplete/noisy data.

---

2. Recent Advances in Symbolic Regression

URL: View paper

Brief Assessment

Symbolic Regression Advances[72] is a survey paper reviewing symbolic regression methods broadly. It does not present a benchmark for logical symbolic regression or address the specific gap between continuous SR and exact logic synthesis that LogicSR targets.

---

3. EDGE: Evaluation Framework for Logical vs. Subgraph Explanations for Node Classifiers on Knowledge Graphs

URL: View paper

Brief Assessment

EDGE[78] focuses on evaluating explanations for knowledge graph node classifiers (comparing logical rule-based vs. subgraph-based explanations for GNN models), not on discovering logical expressions from data or symbolic regression benchmarks.

---

4. Deep Learning for Symbolic Mathematics

URL: View paper

Brief Assessment

Deep Symbolic Mathematics[41] focuses on symbolic integration and differential equations using seq2seq models, not on logical expression discovery from data or boolean function learning, which is the core focus of LogicSR.

---

5. DivLogicEval: A framework for benchmarking logical reasoning evaluation in large language models

URL: View paper

Brief Assessment

DivLogicEval[71] focuses on evaluating logical reasoning in natural language understanding for LLMs through multiple-choice questions, not on discovering logical expressions from data or symbolic regression tasks.

---

6. Logictree: Improving complex reasoning of LLMs via instantiated multi-step synthetic logical data

URL: View paper

Brief Assessment

LogicTree[61] focuses on synthesizing multi-step logical reasoning datasets for training LLMs, not on creating benchmarks for discovering logical expressions from data. The candidate addresses LLM reasoning enhancement through synthetic data generation, while the original paper presents an evaluation benchmark for logical discovery algorithms.

---

7. LogicBench: A Benchmark for Evaluation of Logical Reasoning

URL: View paper

Brief Assessment

LogicBench[75] focuses on evaluating logical reasoning capabilities of LLMs through natural language question-answering across propositional, first-order, and non-monotonic logics. In contrast, LogicSR addresses discovering logical expressions from data (symbolic regression), targeting algorithms that learn generalizable logical rules from incomplete or noisy datasets. These are fundamentally different tasks: LogicBench evaluates reasoning over given logical statements, while LogicSR evaluates expression discovery from input-output samples.

---

8. Contemporary Symbolic Regression Methods and their Relative Performance

URL: View paper

Brief Assessment

Contemporary Symbolic Regression[77] focuses on continuous mathematical functions and real-valued regression problems, not logical/boolean symbolic regression. The paper explicitly states 'sr research and its associated benchmarks have primarily focused on recovering continuous mathematical formulas involving operators such as addition, multiplication, and transcendental functions' and benchmarks methods on real-valued datasets, which is fundamentally different from LogicSR's focus on discovering logical expressions from boolean data.

---

9. Integrating Expert Knowledge into Logical Programs via LLMs

URL: View paper

Brief Assessment

Expert Knowledge Integration[74] focuses on evaluating LLMs' ability to translate expert knowledge (e.g., operational ranges) into executable Python code for engineering monitoring systems. LogicSR addresses discovering logical expressions from data for interpretable AI and scientific discovery. These are distinct problem formulations with different evaluation methodologies.

---

10. Rock: Cleaning Data by Embedding ML in Logic Rules

URL: View paper

Brief Assessment

Rock[76] focuses on data cleaning in relational databases using ML-embedded logic rules for entity resolution and conflict resolution, not on benchmarking algorithms for discovering logical expressions from data or symbolic regression tasks.

Contribution 2: Novel synthetic data generation algorithm

Description: The authors develop a two-stage synthesis process for generating large-scale, complex, and structurally diverse ground-truth logic networks. This algorithm uses truth table analysis, structured sampling, and graph-based composition to produce diverse, non-redundant logical formulas at scale.

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. Logical natural language generation from open-domain tables

URL: View paper

Brief Assessment

Logical Table Generation[64] focuses on generating natural language statements from tables with logical inference, not on synthetic generation of diverse logical expressions or logic networks as in the original paper.

2. Scalable anytime algorithms for learning fragments of linear temporal logic

URL: View paper

Brief Assessment

Temporal Logic Fragments[66] focuses on learning LTL formulas from trace data using directed formulas and boolean combinations, not on synthetic generation of diverse logical expressions. The paper's contribution is an enumeration algorithm for directed LTL formulas, not a data generation method for ground-truth logic networks.

3. Logic Augmented Generation

URL: View paper

Brief Assessment

Logic Augmented Generation[63] focuses on combining LLMs with semantic knowledge graphs for tasks like medical diagnostics and climate projections, not on synthetic data generation algorithms for logical expressions. The candidate addresses a completely different problem domain.

4. Enhancing reasoning capabilities of llms via principled synthetic logic corpus

URL: View paper

Brief Assessment

Synthetic Logic Corpus[69] focuses on generating logical reasoning samples for LLM training using symbolic logic principles, not on generating diverse logical expressions for symbolic regression benchmarks. The domains and objectives differ fundamentally.

5. Logictree: Improving complex reasoning of LLMs via instantiated multi-step synthetic logical data

URL: View paper

Brief Assessment

LogicTree[61] generates logical reasoning trees for LLM training via backward deduction and instantiation into natural language scenarios. The original paper's two-stage synthesis produces ground-truth logic networks with truth table analysis and graph-based composition for benchmarking purposes. These serve fundamentally different objectives and employ distinct technical approaches.

6. System for automatic generation of logical formulas

URL: View paper

Brief Assessment

Automatic Logical Formulas[68] focuses on random generation of SAT problem instances for testing solvers, not on generating diverse ground-truth logic networks with truth table analysis and graph-based composition for machine learning benchmarks.

7. DeLoSo: Detecting Logic Synthesis Optimization Faults Based on Configuration Diversity

URL: View paper

Brief Assessment

DeLoSo[70] focuses on detecting faults in logic synthesis optimization processes through configuration diversity testing, not on generating synthetic logical expressions or truth tables for benchmarking purposes.

8. SynLogic: Synthesizing Verifiable Reasoning Data at Scale for Learning Logical Reasoning and Beyond

URL: View paper

Brief Assessment

SynLogic[62] focuses on logical reasoning tasks (sudoku, game of 24, cipher) with rule-based generators for logic puzzles, not on general logic network synthesis for circuit design and biological networks as in the original paper.

9. Synthetic data generation for statistical testing

URL: View paper

Brief Assessment

Synthetic Testing Data[67] focuses on generating synthetic test data for statistical testing of data-intensive systems with logical validity constraints, not on generating diverse logical expressions or logic networks for symbolic regression benchmarks.

10. ABAC policy mining through affiliation networks and biclique analysis

URL: View paper

Brief Assessment

ABAC Policy Mining[65] focuses on generating synthetic examples for access control policy evaluation, not on creating diverse logical expressions or truth tables for symbolic regression tasks.

Contribution 3: Comprehensive cross-domain evaluation of 17 algorithms

Description: The authors conduct a rigorous evaluation of 17 algorithms spanning classical logic solvers, modern machine learning models, and large language models. The evaluation reveals capability boundaries and provides insights on scalability, noise robustness, and operator-set compatibility across current 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. Cross-domain Open-world Discovery

URL: View paper

Brief Assessment

Cross Domain Discovery[55] focuses on cross-domain open-world discovery for image classification with domain shifts, not on evaluating logic discovery algorithms across different domains and operator sets as in the original paper.

---

2. Truth Discovery Algorithms: An Experimental Evaluation

URL: View paper

Brief Assessment

Truth Discovery Algorithms[59] evaluates 12 truth discovery algorithms for data fusion and conflict resolution, not logic discovery algorithms. The domains (data veracity, source credibility) and algorithm types (truth discovery, probabilistic models) are fundamentally different from the original paper's focus on logical symbolic regression across classical solvers, ML models, and LLMs.

---

3. A Cross-Domain Evaluation of Approaches for Causal Knowledge Extraction

URL: View paper

Brief Assessment

Cross Domain Causal[57] evaluates 4 neural models for causal knowledge extraction across 4 datasets, not 17 algorithms for logical discovery. The domains (causal relation extraction vs. logical symbolic regression) and evaluation objectives are fundamentally different.

---

4. Survey and Evaluation of Causal Discovery Methods for Time Series

URL: View paper

Brief Assessment

Causal Discovery Survey[53] focuses on causal discovery methods for time series data, evaluating approaches like Granger causality and constraint-based methods. The ORIGINAL paper evaluates logical symbolic regression algorithms across different domains (circuits, biological networks), which is a fundamentally different task from temporal causal inference.

---

5. Latent logic tree extraction for event sequence explanation from llms

URL: View paper

Brief Assessment

Latent Logic Tree[56] focuses on extracting logic trees from event sequences using LLMs and temporal point processes, not on cross-domain evaluation of logic discovery algorithms. The candidate evaluates its own method against baselines for event prediction tasks, which is fundamentally different from the original paper's systematic benchmark evaluation of 17 diverse algorithms across logic synthesis, symbolic regression, and machine learning domains.

---

6. Security Key Management Protocol for Cross-domain Authentication of Internet of Vehicles

URL: View paper

Brief Assessment

Cross Domain Authentication[60] focuses on security key management and authentication protocols for Internet of Vehicles, not on evaluating logic discovery algorithms across domains. The domains referenced are network trust domains, not algorithmic evaluation domains.

---

7. Evaluating the Logical Reasoning Ability of ChatGPT and GPT-4

URL: View paper

Brief Assessment

ChatGPT Logical Reasoning[51] evaluates ChatGPT and GPT-4 on logical reasoning benchmarks (LogiQA, ReClor, NLI tasks), not a cross-domain evaluation of 17 diverse algorithms spanning logic solvers, ML models, and LLMs for logic discovery from data.

---

8. Criteria2Query: a natural language interface to clinical databases for cohort definition

URL: View paper

Brief Assessment

Criteria2Query[58] focuses on natural language interfaces for clinical database queries and cohort definition, not on evaluating logic discovery algorithms across domains. The technical domains and objectives are fundamentally different.

---

9. Revisiting Reinforcement Learning for LLM Reasoning from A Cross-Domain Perspective

URL: View paper

Brief Assessment

Cross Domain Reinforcement[54] focuses on RL training across six reasoning domains (math, code, science, logic, simulation, tabular) for LLMs, not on evaluating logic discovery algorithms across classical solvers, ML models, and LLMs as in the original paper.

---

10. Verifybench: A systematic benchmark for evaluating reasoning verifiers across domains

URL: View paper

Brief Assessment

VerifyBench[52] focuses on evaluating verifiers for reinforcement learning with verifiable reward (RLVR) across mathematics, physics, chemistry, and biology domains. The original paper evaluates logic discovery algorithms (logic solvers, ML models, LLMs) for symbolic regression tasks in boolean/circuit domains. These are fundamentally different evaluation objectives and application domains.

Appendix: Text Similarity Detection

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

References

- [0] LogicSR: A Unified Benchmark for Logical Discovery from Data View paper
- [1] Inferring Logical Forms From Denotations View paper
- [2] Learning temporal formulas from examples is hard View paper
- [3] Discovering Symbolic Models from Deep Learning with Inductive Biases View paper
- [4] Discovering Classification Rules Using Variable-Valued Logic System VL1 View paper
- [5] Differentiable Learning of Logical Rules for Knowledge Base Reasoning View paper
- [6] A neural symbolic model for space physics View paper
- [7] ChatRule: Mining Logical Rules with Large Language Models for Knowledge Graph Reasoning View paper
- [8] Rule by example: Harnessing logical rules for explainable hate speech detection View paper
- [9] Deep Symbolic Learning: Discovering Symbols and Rules from Perceptions View paper
- [10] Neural symbolic regression that scales View paper
- [11] TILP: Differentiable Learning of Temporal Logical Rules on Knowledge Graphs View paper
- [12] Mining logical arithmetic expressions from proper representations View paper
- [13] Selection-inference: Exploiting large language models for interpretable logical reasoning View paper
- [14] Generating predicate logic expressions from natural language View paper
- [15] AI Reasoning in Deep Learning Era: From Symbolic AI to Neuralâ□□Symbolic AI View paper
- [16] Retrieval-Augmented Mining of Temporal Logic Specifications from Data View paper
- [17] Learning symbolic models for graph-structured physical mechanism View paper
- [18] Fortune: Formula-Driven Reinforcement Learning for Symbolic Table Reasoning in Language Models View paper
- [19] Rlogic: Recursive logical rule learning from knowledge graphs View paper
- [20] Machine learning symbolic equations for diffusion with physics-based descriptions View paper
- [21] Logical Rule Learning View paper
- [22] Mining logical circuits in fungi View paper
- [23] Logic Tensor Networks: Deep Learning and Logical Reasoning from Data and Knowledge View paper
- [24] Harnessing Deep Neural Networks with Logic Rules View paper
- [25] Finite Expression Methods for Discovering Physical Laws from Data View paper
- [26] A coupled framework for symbolic turbulence models from deep-learning View paper
- [27] Towards Deep Symbolic Reinforcement Learning View paper
- [28] Advanced Weakly-Supervised Formula Exploration for Neuro-Symbolic Mathematical Reasoning View paper
- [29] End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction View paper
- [30] Probabilistic logic neural networks for reasoning View paper
- [31] Exploring logical reasoning for referring expression comprehension View paper
- [32] Syllogistic reasoning: Logical decisions from a complex data base View paper
- [33] Bi-directional Learning of Logical Rules with Type Constraints for Knowledge Graph Completion View paper
- [34] TLogic: Temporal Logical Rules for Explainable Link Forecasting on Temporal Knowledge Graphs View paper
- [35] SYMBOL: Generating flexible black-box optimizers through symbolic equation learning View paper
- [36] Using symbolic machine learning to assess and model substance transport and decay in water distribution networks View paper
- [37] Can llms reason with rules? logic scaffolding for stress-testing and improving llms View paper
- [38] Proxy Attribute Discovery in Machine Learning Datasets via Inductive Logic Programming View paper
- [39] Learning to Explore Paths for Symbolic Execution View paper
- [40] Kolmogorov-Arnold Representation for Symplectic Learning: Advancing Hamiltonian Neural Networks View paper
- [41] Deep Learning for Symbolic Mathematics View paper
- [42] -Analyses with Symbolic Regression View paper
- [43] Leveraging Logical Rules in Knowledge Editing: A Cherry on the Top View paper
- [44] Combining data and theory for derivable scientific discovery with AI-Descartes View paper
- [45] Towards logiglue: A brief survey and a benchmark for analyzing logical reasoning capabilities of language models View paper
- [46] Neural logic reasoning View paper
- [47] Generative pre-trained transformer for symbolic regression base in-context reinforcement learning View paper
- [48] Neuro-Symbolic Class Expression Learning View paper
- [49] Establishing logical rules from empirical data View paper
- [50] SWARMFLAWFINDER: Discovering and Exploiting Logic Flaws of Swarm Algorithms View paper
- [51] Evaluating the Logical Reasoning Ability of ChatGPT and GPT-4 View paper
- [52] Verifybench: A systematic benchmark for evaluating reasoning verifiers across domains View paper
- [53] Survey and Evaluation of Causal Discovery Methods for Time Series View paper
- [54] Revisiting Reinforcement Learning for LLM Reasoning from A Cross-Domain Perspective View paper
- [55] Cross-domain Open-world Discovery View paper
- [56] Latent logic tree extraction for event sequence explanation from llms View paper
- [57] A Cross-Domain Evaluation of Approaches for Causal Knowledge Extraction View paper
- [58] Criteria2Query: a natural language interface to clinical databases for cohort definition View paper
- [59] Truth Discovery Algorithms: An Experimental Evaluation View paper
- [60] Security Key Management Protocol for Cross-domain Authentication of Internet of Vehicles View paper
- [61] Logictree: Improving complex reasoning of LLMs via instantiated multi-step synthetic logical data View paper
- [62] SynLogic: Synthesizing Verifiable Reasoning Data at Scale for Learning Logical Reasoning and Beyond View paper
- [63] Logic Augmented Generation View paper
- [64] Logical natural language generation from open-domain tables View paper
- [65] ABAC policy mining through affiliation networks and biclique analysis View paper
- [66] Scalable anytime algorithms for learning fragments of linear temporal logic View paper
- [67] Synthetic data generation for statistical testing View paper
- [68] System for automatic generation of logical formulas View paper
- [69] Enhancing reasoning capabilities of llms via principled synthetic logic corpus View paper

- [70] DeLoSo: Detecting Logic Synthesis Optimization Faults Based on Configuration Diversity View paper
- [71] DivLogicEval: A framework for benchmarking logical reasoning evaluation in large language models View paper
- [72] Recent Advances in Symbolic Regression View paper
- [73] Large language models meet symbolic provers for logical reasoning evaluation View paper
- [74] Integrating Expert Knowledge into Logical Programs via LLMs View paper
- [75] LogicBench: A Benchmark for Evaluation of Logical Reasoning View paper
- [76] Rock: Cleaning Data by Embedding ML in Logic Rules View paper
- [77] Contemporary Symbolic Regression Methods and their Relative Performance View paper
- [78] EDGE: Evaluation Framework for Logical vs. Subgraph Explanations for Node Classifiers on Knowledge Graphs View paper