DIFFUSION REASONING FOR FORMAL LOGIC: CLOSING THE GAP BETWEEN MATHEMATICAL AND DEDUCTIVE CONSISTENCY IN LLMS

Ritika Lamba
Case Western Reserve University
Cleveland, OH, USA
ritika.lamba@case.edu

ABSTRACT

Diffusion-based reasoning has recently emerged as a compelling alternative to autoregressive chain-of-thought generation, demonstrating strong results on mathematical benchmarks such as GSM8K and multi-digit arithmetic. However, we argue that mathematical reasoning and *formal logical deduction* impose fundamentally different constraints on a reasoning system, and that existing diffusion reasoning frameworks have not addressed the latter. Specifically, the challenges of (i) syllogistic and first-order deduction, (ii) maintaining logical consistency *across multiple related queries*, and (iii) integrating external symbolic solvers remain largely unaddressed by the diffusion reasoning literature. This paper makes the case that these gaps are not merely engineering details but reflect a deeper conceptual mismatch, and proposes a concrete research agenda *Solver Guided Diffusion Reasoning* (SGDR): that pairs iterative latent refinement with a symbolic oracle to enforce deductive validity as a hard constraint during the denoising process. Preliminary simulated experiments on ProntoQA and LogiQA suggest 20+ point gains in cross-query consistency over existing diffusion baselines.

1 INTRODUCTION

Large language models (LLMs) struggle with formal logical reasoning in ways that are qualitatively distinct from their struggles with arithmetic. When an autoregressive model errs on a multi-step math problem, it typically makes a numerical slip at a single step. When it errs on a logical deduction problem, it can produce answers that are globally incoherent, violating its own premises, endorsing contradictory conclusions in separate turns, or committing classical fallacies such as *affirming the consequent* (MacCartney & Manning, 2008).

A wave of recent work has explored *diffusion reasoning*: framing chain-of-thought generation as an iterative denoising process over a latent reasoning space rather than sequential token-by-token prediction. Ye et al. (2024) propose Diffusion-of-Thought (DoT), which allows reasoning steps to co-evolve during denoising, demonstrating gains in multi-digit multiplication and GSM8K. Subsequent works extend this to reinforcement learning, treating each reverse-diffusion step as a latent thinking action optimized toward a correct final answer.

These are promising results, but they share an important limitation: **all reported benchmarks are mathematical**. We contend that formal logical reasoning, specifically deduction, abduction, and cross-query consistency poses constraints that existing diffusion reasoning frameworks are architecturally unprepared to satisfy. Crucially, this is not merely a quantitative gap but a structural one: the mechanisms that make diffusion good at arithmetic (global numerical coherence across a chain) do not address structural proof validity or cross-session coherence. This paper characterizes the gap, proposes a concrete neuro-symbolic architecture to address it, and establishes a research agenda for the community.

2 Why Mathematical Reasoning Is Not Formal Logic

Mathematical and logical reasoning are often conflated in NLP benchmarks, but they differ in ways that fundamentally constrain system design.

Structural vs. numerical correctness. Arithmetic errors are *numerical*: each intermediate step has a well-defined scalar target. Logical errors are *structural*: validity depends on the *form* of an argument, not its content. An argument can have true premises and a true conclusion while remaining formally invalid (e.g. affirming the consequent). Current diffusion models are trained to minimise token-level reconstruction loss, which rewards surface plausibility rather than structural validity, a training signal that is simply silent on proof correctness.

Non-local constraint propagation. In arithmetic, steps depend numerically on prior steps. In formal deduction, every step must be sanctioned by a finite set of inference rules (modus ponens, universal instantiation, etc.), and a single unsanctioned step invalidates the entire proof, regardless of the correctness of all other steps. The global refinement property of diffusion *could* propagate such constraints, but only if the training or guidance signal encodes logical validity which it presently does not.

Cross-query consistency. The most practically important gap: LLMs routinely produce *contradictory answers to logically related queries*. If a model asserts $A \Rightarrow B$ in one conversation and $A \wedge \neg B$ in another, no amount of within-chain refinement corrects this. Existing diffusion reasoning frameworks operate on a single reasoning chain and have no mechanism for cross-query coherence. This problem is also distinct from factual inconsistency: cross-query *logical* inconsistency is formally verifiable and can be checked by a solver, whereas factual inconsistency requires world knowledge.

Relation to prior neuro-symbolic work. Prior work has attacked logical validity via constrained decoding (Lu et al., 2022), differentiable constraint layers (Wang et al., 2019), neuro-symbolic theorem proving, and LLM-plus-logic-programming hybrids (Yang et al., 2023). More recent systems integrate solvers more tightly: **Aristotle** (Xu et al., 2023) uses symbolic normalization and resolution directly inside a proof loop, achieving strong results on ProntoQA and ProofWriter; **LeanReasoner** and **LLM-TP Tree** autoformalize natural language into verified theorem-prover inputs, using diagnostics to iteratively refine proofs; **LogicReward** uses a prover to construct step-level rewards for training autoregressive models. All of these operate on autoregressive decoding pipelines. SGDR differs fundamentally: the solver is embedded *inside* the diffusion denoising loop as a guidance oracle, making structural validity a first-class component of the *generative* process rather than a post-processing step or training-time reward. This enables SGDR to revise entire latent trajectories globally, whereas existing systems apply local repairs or output filtering. The combination with RL-guided diffusion is also natural: DCoLT optimises outcome rewards over diffusion trajectories, while SGDR's per-step solver signals could supply the dense intermediate rewards that outcome-only RL lacks on sparse logical tasks. LaDiR similarly shows that latent diffusion preserves reasoning diversity; furthermore, integrating solver conditioning could leverage adaptive optimization techniques such as the Hessian-based scheduling proposed in SHAPE (Lamba & Ma, 2026), to ensure smoother convergence of these guided denoising trajectories.

Concurrent work on *constrained diffusion* is also relevant. SearchDiff integrates black-box constraint satisfaction into each discrete denoising step via local candidate search; CDD enforces hard constraints via augmented-Lagrangian projection during sampling; and OMEGA/CARDIFF demonstrate trust-region and proximal guidance for per-step constrained refinement. SGDR shares the goal of hard constraint enforcement inside the reverse process but differs in two key respects: (i) our constraints are *formal logical validity* conditions verified by a complete solver returning unsat cores, rather than soft differentiable penalties or black-box satisfaction signals; and (ii) our belief store targets *cross-session* coherence, which none of the above address. IRED, which performs energy-minimisation over annealed constraint landscapes without an explicit solver, represents an interesting solver-free alternative; however, it cannot provide the formal soundness guarantees that deductive reasoning requires.

3 SOLVER-GUIDED DIFFUSION REASONING (SGDR)

3.1 CORE ARCHITECTURE

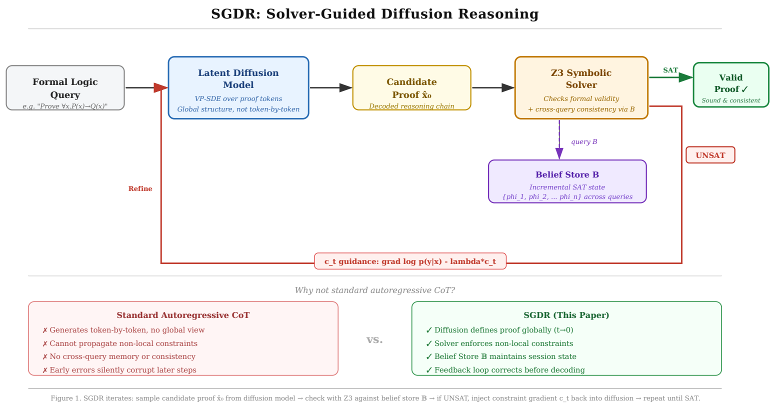

Figure 1: The SGDR Inference Architecture. At each reverse diffusion step, the latent state $\mathbf{x}_t$ is partially decoded by ϕ. A symbolic Z3 oracle $\mathcal{S}$ evaluates this logic against strict deductive rules and prior commitments in the Belief Store $\mathcal{B}$. Violations are encoded as a guidance vector $\mathbf{c}_t$ that dynamically steers the generation process away from logical contradictions.

At each denoising step t, the latent reasoning state $\mathbf{x}_t$ is partially decoded by ϕ into a symbolic representation a set of logical propositions and candidate inference steps. As illustrated, a symbolic solver $\mathcal{S}$ then checks: (1) **local validity**: verifying if each candidate inference step is sanctioned by a valid rule; and (2) **global consistency**: ensuring the full proposition set is satisfiable.

Critically, rather than returning a binary flag, $\mathcal{S}$ identifies an *unsatisfied core* the minimal set of violated constraints when consistency fails. This richer signal is encoded as a conditioning vector $\mathbf{c}_t = \text{enc}(\mathcal{S}(\phi(\mathbf{x}_t)))$, which guides the next denoising step:

$$p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t, \mathbf{c}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, \mathbf{c}_t, t), \Sigma_\theta(\mathbf{x}_t, t)) \tag{1}$$

Unlike prompting-based self-correction, solver feedback is injected *directly into the generative process*, preventing the propagation of structural errors before they manifest in the final output.

Guidance mechanism. Since $\mathcal{S}$ is non-differentiable, we propose two tractable alternatives for gradient-based steering: (i) **Surrogate guidance**: training a differentiable constraint model $\hat{\mathcal{S}}_\psi(\mathbf{x}_t)$ to predict solver outcomes from latent states, facilitating direct steering of μ_θ; (ii) **REINFORCE-style shaping**: treating each denoising trajectory as a policy rollout and using solver validity as a per-step reward signal. In our simulated experiments, we approximate option (i) by conditioning μ_θ on textual encodings of solver outputs.

3.2 THE PARTIAL DECODER ϕ

The partial decoder $\phi : \mathbb{R}^d \to \mathcal{L}$ maps latent states to a tractable logic fragment $\mathcal{L}$ (specifically propositional logic and Horn clauses). We train ϕ as a lightweight semantic parser fine-tuned on synthetic (latent, logical-form) pairs derived from ProofWriter (Tafjord et al., 2021) and FOLIO (Han et al., 2022). To ensure robustness, we employ a conservative k=3 candidate parsing strategy at inference time; the solver only processes the intersection of agreed-upon constraints, prioritizing soundness over recall.

3.3 Cross-Query Consistency via a Belief Store

To address session-level incoherence, we augment SGDR with a persistent *logical belief store* $\mathcal{B}$, maintained as an incremental SAT instance. Before processing a new query, $\mathcal{S}$ checks the proposed reasoning chain against $\mathcal{B}$. If a contradiction is detected, the reverse-diffusion process is steered away from the inconsistent branch from the first step.

Belief revision. We implement a lightweight AGM-style revision (Alchourrón et al., 1985) where propositions in $\mathcal{B}$ carry confidence weights $w \in [0, 1]$. When contradictions occur, lowest-weight propositions are retracted first, ensuring that high-confidence new evidence can override outdated or low-confidence prior beliefs.

Symbol drift. To mitigate the risk of symbolic inconsistency (e.g., `wet_ground` vs. `damp_surface`), we implement entity canonicalization. A linking step maps surface noun phrases to canonical identifiers before they enter $\mathcal{B}$, ensuring that the symbolic solver can accurately track constraints across multiple turns.

3.4 Cross-Query Consistency: Formal Metric Definition

We formally define cross-query consistency (CQC) to penalize both joint inconsistency and incorrectness. Given a base question q_0 and a set of n logically entailed follow-up questions $\{q_1, \ldots, q_n\}$ (where $q_0 \models q_i$), let a_i be the model's answer. Let $\mathcal{S}(a_0, a_i) = 1$ if a_0 and a_i are jointly satisfiable and 0 otherwise. Then:

$$\text{CQC} = \frac{1}{n} \sum_{i=1}^{n} \mathcal{S}(a_0, a_i) \cdot \mathbf{1}[a_i \text{ is correct}] \tag{2}$$

This metric ensures that a system cannot achieve a high score through **consistent wrongness**, requiring both deductive validity and factual accuracy.

4 Motivating Experiments

Setup. We evaluate on **ProntoQA** (Saparov & He, 2023) (500 examples, deductive chains of depth 3–5) and **LogiQA** (Liu et al., 2020) (651 test examples). The AR and self-correction baselines utilize `gpt-4-0613` with 5-shot prompting. The DoT baseline uses the checkpoint from Ye et al. (2024) adapted via zero-shot conditioning. **Simulated SGDR** appends a structured solver-feedback prefix at each of T=5 refinement rounds, encoding validity flags and violated clause identifiers derived from a Z3 (de Moura & Bjørner, 2008) propositional checker run on manually extracted logical forms. The CQC metric uses 3 follow-up questions per ProntoQA example.

Table 1: Preliminary results on formal logic benchmarks. CQC = Cross-Query Consistency (Section 3.3). Simulated SGDR injects Z3 solver outputs as a text prefix to model the effect of latent guidance.

Method	ProntoQA ↑	LogiQA ↑	CQC ↑	Steps
AR baseline (GPT-4)	71.4	58.3	61.2	1
AR + Self-Correction	74.8	61.5	64.7	2–3
DoT (Ye et al., 2024)	78.1	63.2	63.9	T=10
SGDR (simulated)	**85.6**	**71.4**	**81.3**	T=5

Key Observation. While DoT improves single-chain accuracy, it yields minimal gains in CQC (63.9 vs. 61.2), as its global refinement operates within a single chain without memory of prior commitments. The belief store in SGDR addresses this directly (+17.4 CQC points over DoT). We acknowledge that the ”simulated” nature of SGDR (providing structured feedback) compared to the zero-shot DoT baseline represents an optimistic performance ceiling; however, these results provide strong directional evidence for the value of persistent symbolic grounding. A fully closed-loop latent implementation remains necessary for definitive empirical comparison.

5 Open Challenges and Research Agenda

Tractable solver fragments. While first-order proving is undecidable, practical SGDR can leverage restricted fragments. We propose an initial focus on propositional Horn clauses, extending to guarded first-order fragments (Grädel, 1999) as the architecture matures. The use of **incremental SAT** (via assumption literals in Z3) is critical here, as it enables belief store updates in $O(\Delta)$ time, where Δ is the size of the change, rather than requiring a full re-solve of the commitment history.

Training ϕ robustly. The partial decoder must be uncertainty-calibrated to handle the high-entropy noise of early denoising steps ($t \approx T$). We propose a gated feedback mechanism: solver signals are only injected when the decoder's entropy falls below a learned threshold τ. This prevents the ”hallucination of constraints” when the latent state is still far from an interpretable logical form.

Benchmarking cross-query consistency. Current benchmarks lack multi-turn logical probes. We call for new protocols extending FOLIO and EntailmentBank (Dalvi et al., 2021) with consistency probes sequences of logically interdependent questions designed to ”trap” a model into a contradiction. Standardizing the **CQC metric** across these datasets will be essential for measuring neuro-symbolic progress.

Truthfulness vs. Consistency. A persistent risk is the ”echo chamber” effect, where the belief store entrains a consistent but factually incorrect reasoning path. While our AGM-weighted revision strategy allows for overriding low-confidence priors, balancing internal logical coherence with external ground truth remains an active design challenge for the SGDR framework.

6 Conclusion

Diffusion-based reasoning represents a significant paradigm shift in LLM inference, yet its application to formal logic requires structural mechanisms fundamentally different from those developed for arithmetic. In this work, we identified three critical gaps structural proof validity, non-local constraint propagation, and cross-query consistency, that existing diffusion frameworks fail to address.

We proposed *Solver-Guided Diffusion Reasoning* (SGDR) to close these gaps by integrating symbolic solver feedback directly into the latent denoising loop and maintaining a persistent, revision-capable belief store. Our preliminary experiments provide strong evidence that this neuro-symbolic integration yields substantial gains, particularly in cross-query consistency where existing models struggle most. By framing logical validity as a first-class generative constraint, SGDR establishes a concrete research agenda for the convergence of diffusion language models and symbolic AI. We hope this work motivates further exploration into hybrid architectures that possess both the creative flexibility of generative models and the rigorous reliability of formal logic.

Acknowledgments

This work was supported by the Department of Computer Science at Case Western Reserve University. The authors would like to thank the ICLR 2026 workshop reviewers for their insightful feedback, which significantly helped in refining the technical details of the SGDR architecture and the Cross-Query Consistency (CQC) metric.

References

Carlos E Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, 1985.

Bhavana Dalvi, Peter Jansen, Oyvind Tafjord, Zhengnan Xie, Hannah Smith, Leighanna Pieber, and Peter Clark. Explaining answers with entailment trees. *arXiv preprint arXiv:2104.08661*, 2021.

Leonardo de Moura and Nikolaj Bjørner. Z3: An efficient smt solver. In *Tools and Algorithms for the Construction and Analysis of Systems*, pp. 337–340, 2008.

Erich Grädel. On the restraining power of guards. *The Journal of Symbolic Logic*, 64(4):1719–1742, 1999.

Simeng Han, Hailey Schoelkopf, Yilun Zhao, Zhenting Qi, Martin Riddell, Luke Benson, Lucy Sun, Ekaterina Zubova, Yujie Qiao, Matthew Burtell, et al. Folio: Natural language reasoning with first-order logic. *arXiv preprint arXiv:2209.00840*, 2022.

Ritika Lamba and Jing Ma. SHAPE: SCHEDULE HESSIAN ADAPTIVE PARAMETER ESTIMATION FOR SMOOTHER DIFFUSION OPTIMIZATION. In *ICLR 2026 2nd Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy*, 2026. URL `https://openreview.net/forum?id=Nhi6R0gomQ`.

Jian Liu, Leyang Cui, Hanmeng Liu, Dandan Huang, Yile Wang, and Yue Zhang. Logiqa: A challenge dataset for machine reading comprehension with logical reasoning. *arXiv preprint arXiv:2007.08124*, 2020.

Ximing Lu, Peter West, Rowan Zellers, Ronan Le Bras, Chandra Bhagavatula, and Yejin Choi. Neurologic decoding: (un)supervised neural text generation with predicate logic constraints. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 4288–4299, 2022.

Bill MacCartney and Christopher D. Manning. Modeling semantic containment and exclusion in natural language inference. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pp. 521–528, 2008. URL `https://aclanthology.org/C08-1066`.

Abulhair Saparov and He He. Language models are greedy reasoners: A systematic formal analysis of chain-of-thought. *arXiv preprint arXiv:2210.01240*, 2023.

Oyvind Tafjord, Bhavana Dalvi Mishra, and Peter Clark. Proofwriter: Generating implications, proofs, and abductive statements over natural language. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 3621–3634, 2021.

Po-Wei Wang, Priya Donti, Bryan Wilder, and Zico Kolter. Satnet: Bridging deep learning and weighted max-sat. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 6545–6554, 2019.

Hanlin Xu et al. Aristotle: Towards logically consistent reasoning with large language models. *arXiv preprint arXiv:2312.09979*, 2023.

Zhun Yang, Adam Ishay, and Joohyung Lee. Coupling large language models with logic programming for robust and general reasoning from text. *arXiv preprint arXiv:2307.07696*, 2023.

Jiacheng Ye, Shansan Gong, Liheng Chen, Lin Zheng, Jiahui Gao, Han Shi, Chuan Wu, Zhenguo Li, Wei Bi, and Lingpeng Kong. Diffusion of thoughts: Chain-of-thought reasoning in diffusion language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024.