SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Unified Neuro Symbolic Reasoning Algorithms Integrating First Order Logic With Transformers

E. Shalini[1*], K. Kanchana[2], Dr. Shinki Katyayani Pandey[3], Azizov Azizbek Sabitkhanovich[4]

[1*]Assistant Professor, Department of Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. E-mail: shalini@maher.ac.in

[2]Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. E-mail: kanchana@maher.ac.in; kanchana@maher.ac.in

[3]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India.
E-mail: u.shinkikatyayanipandey@kalingauniversity.ac.in, https://orcid.org/0009-0009-9316-5093

[4]Director of Research & Innovation Department, Faculty of Business administration, Turan International University, Namangan, Uzbekistan. E-mail: science@tiu-edu.uz, https://orcid.org/0000-0001-9210-7582

*Corresponding author: Email: shalini@maher.ac.in

Abstract

The integration of symbolic reasoning and deep learning is considered a revolutionary step in the field of artificial intelligence. This work proposes Unified Neuro-Symbolic Reasoning Algorithms (UNSRA), an innovative system that seamlessly combines first-order logic (FOL) and deep neural networks using transformer architectures. Conventional deep learning algorithms have great abilities to identify patterns but lack interpretability and logic-based reasoning, whereas classical symbolic AI approaches are capable of performing complex logic operations but cannot learn from raw data. This work takes advantage of both worlds by designing a two-path architecture in which neural networks deal with semantic representations and FOL components address logic-related tasks such as inference and rule implementation. The experimental results obtained using different datasets, such as Visual Question Answering, Knowledge Graph Completion, and Reasoning problems, show that the model outperforms existing neuro-symbolic models by 18% to 24% in logical consistency measures, while still competing effectively in benchmark tests. The findings of the interpretability study demonstrate that research have developed an approach that yields explainable decision-making processes that can be tracked from the neural layer to the symbolic layer, thus filling vital gaps in explainability in AI systems. Research contribution includes a flexible and modular system design that can be customized for different reasoning applications.

Keywords: First-order Logic, Interpretability, Knowledge Representation, Neural-Symbolic Integration, Transformer Networks.

1. Introduction

Deep learning techniques have been immensely successful in AI through the use of transformer models, which have transformed NLP and computer vision. But such models remain black boxes, using learned associations to perform decision-making without any logical inferences [1][4]. On the other hand, classic AI models have exhibited robustness in performing logical reasoning but necessitate extensive knowledge engineering and fail to adapt to new domains without undergoing re-training [2][6]. The key challenge associated with purely neural models is that they cannot ensure logical consistency and remain opaque in their reasoning process. Classic models, on the other hand, lack the flexibility and ability to learn [3][5].

The latest developments in the field of neuro-symbolic artificial intelligence have aimed at filling this gap; however, the current strategies suffer from either a superficial combination of symbolic and neural elements or lack of principled ways of combining them [6]. In this paper, research propose Unified Neuro-Symbolic Reasoning Algorithms (UNSRA) that provide a full-fledged solution for deep integration of first-order logic and transformer models. Studies novel contribution is an attention-based logic gate structure that facilitates bi-directional interactions between neural embeddings and symbolic expressions, enabling joint learning of symbolic rules and neural weights.

This paper makes three main contributions:

- To propose an innovative unified architecture where first-order logic is successfully incorporated into neural networks based on transformers via differentiable logic gates.
- To prove the effectiveness of study method via empirical results, showing its superiority in dealing with complicated reasoning problems while ensuring explainability.
- To develop a general framework which can be easily adapted for different applications.

2. Related Work

Neural-symbolic computing is one topic that has seen a great deal of research with pioneering works such as Logic Tensor Networks (LTN) and Neural-Symbolic Visual Question Answering models [7]. Garcez and Lamb presented a complete approach to neural-symbolic integration using differentiable logic programming [8]. Nevertheless, such methods usually involve embedding logic in neural networks or working separately using symbolic systems parallel to neural elements. The main distinction of the approach is in reaching the level of full bidirectional integration of the two computing methods [9]. Transformer-based models have emerged as the state-of-the-art approach for modeling sequential data and learning representations with structure [10]. The self-attention method, which allows parallel computation on sequences as well as modeling of long-range interactions, was first proposed by [11]. Improved versions such as BERT, GPT, and Vision Transformers extended their applications to various domains [12]. The research applies transformer models enhanced with logical reasoning capability using symbolic modules. First-order logic is an excellent formal system for knowledge representation and inference through logical reasoning [13]. The classical methodologies for logic programming and automated reasoning have provided sound mechanisms for executing rules and satisfying constraints. The system utilizes FOL concepts but modifies them to suit differentiable operations, thus facilitating gradient-based training of the logical rules.

3. Methodology

Architecture Overview

There are three main components of the UNSRA architecture, which include:

a) Neural Semantic Module, which is the transformer encoder to encode input features to produce distributed semantic representations;
b) Logical Reasoning Module, which is the first-order logic inferencing engine that keeps symbolic knowledge bases to do the reasoning process; and
c) Attention-Weighted Logic Gates, which connect neural and symbolic worlds in two directions.

Component	Function	Key Features
Neural Semantic Module	Process raw input into embeddings	Multi-head attention, positional encoding
Logic Reasoning Module	Perform FOL inference	Forward chaining, soft unification
Attention-Weighted Gates	Bridge neural-symbolic gap	E2L and L2E conversions

Neural Semantic Module

This neural module contains a multi-head self-attention transformer encoder which takes as input an arbitrary-length sequence [14]. Every entity in the input is first encoded into a semantic vector using a very large semantic space where the interactions between the entities can be captured using attention mechanisms. This results in contextual semantic representations whereby the semantic representation of every token in the sequence encodes information about the context.

Embedding attention weights are computed in equation (1) as:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V \qquad (1)$$

where Q, K, V represent query, key, and value projections respectively, and d_k is the dimensionality of the key space.

Logical Reasoning Module

The Logical Reasoning Module keeps an evolving knowledge base in the first-order logic representation format [15]. The job of this module is to apply rules and constraints, do deductions, and satisfy constraints. Rules in the knowledge base are represented by Horn clauses: $H \leftarrow B_1, B_2, \ldots, B_n$, where H stands for head predicate and $B_1 \ldots B_n$ stand for body predicates. For rule application, the Logical Reasoning Module uses the technique called forward chaining. In other words, the module derives new information until it reaches a fixed point or answers the query that it needs to answer. An important novelty introduced into the system is the differentiation of inference through soft unification and probabilistic rule application.

Attention-Weighted Logic Gates

The attention-weighted logic gates (AWLGs) play an important role in connecting symbolic and neural modules. Specifically, these gates enable the determination of symbolic reasoning contribution to output predictions as well as providing feedback from neural to symbolic modules in terms of neural confidence. There exist two different kinds of gates: embedding to logic gates (E2L) transfer information from the neural module to the symbolic one by translating attention weights into logical predicates; and logic to embedding gates (L2E) perform a reverse transformation.

Embedding-to-Logic conversion:

$$\mathrm{Pred}(x) = \mathrm{ThresholdGate}(\mathrm{Attention}(x), \tau) \qquad (2)$$

Logic-to-Embedding mapping:

$$\mathrm{Embed}'(x) = \mathrm{Embed}(x) + \lambda \cdot \mathrm{LogicReach}(x) \qquad (3)$$

where in equation (2) & (3), τ is a learnable threshold parameter, LogicReach represents the set of logical deductions reaching fact x, and λ is a learned gating parameter balancing neural and symbolic contributions.

4. Experimental Evaluation

UNSRA is evaluated on three different tasks:

(1) VQA on the GQA dataset, which involves understanding visual content as well as reasoning about object relationships;

(2) KG Completion on the Freebase and YAGO datasets, evaluating the capability of the model to infer missing relations from incomplete information;

(3) Complex Reasoning on the CLEVR-CoGenT dataset, requiring reasoning about spatial and numeric relations in a systematic way.

In the case of the GQA dataset, UNSRA obtains an accuracy of 87.3%, while the baseline neuro-symbolic models achieve only 71.2% accuracy. The higher accuracy of UNSRA indicates its better ability to reason over complex spatial and relational data. An important point here is that UNSRA outperforms baseline models by 23% in multi-step reasoning problems.

Method	VQA	KG Comp.	CLEVR	Interpretability
Pure Neural	76.1%	0.398	78.5%	62%
Baseline NS	71.2%	0.399	82.3%	71%
UNSRA (Ours)	87.3%	0.487	94.2%	89%

The attention weighted logic gates allow the system to dynamically manage the trade-off between neural perception and symbolic reasoning. In case of queries that relate to counting objects or establishing relationships, the system learns to give more importance to logical reasoning, whereas subjective assessment becomes more important for other kinds of questions.

The proposed method for knowledge graph completion yields MRR of 0.487 on Freebase and 0.521 on YAGO datasets, which amounts to 22% and 24% gains from previous neuro-symbolic approaches, respectively. The ability to learn hidden logical rules is evidence that the proposed unified framework works well. On the CLEVR-CoGenT dataset, which was created specifically to test compositional generalization, the UNSRA framework performs at 94.2%, while neural baselines score only 78.5%. It proves the importance of symbolic reasoning in obtaining compositional generalization ability, which is one of the weaknesses of neural frameworks.

An important benefit of UNSRA compared to pure neural models is interpretability. Decision making can be understood using both neural and symbolic processing. The study investigate decision making processes for each of the VQA cases and illustrate the interaction between visual embedding and logical rules leading to an answer. Consider the question 'What color is the object to the left of the red cube?'. In this case, the reasoning process followed by the model is as follows: (1) neural detection of the red cube, (2) use of the 'left-of' rule, (3) neural detection of the target object, and (4) neural detection of its color. The degree of interpretability is measured in terms of explanation accuracy, or how well the explanations capture the true underlying decision-making process of the model. Using the LIME perturbation test, study show that the UNSRA explanation reaches an accuracy of 89%, compared to only 62% for the black-box neural network.

5. Discussion

In addition, the performance gains in UNSRA are attributed to a number of reasons: (1) Symbolic reasoning ensures that spurious associations are not learned in the network, which would be the case in any neural-only approach; (2) Neural learning allows rules to be automatically learned by the system from data instead of having them defined manually; (3) Attention-weighted gates help in dynamically weighting paradigms according to the needs of the task [18]. Nevertheless, there are certain computational issues associated with the proposed framework [17]. Since the process of symbolic inference is iterative, there is increased inference time, which averages about 2.3x slower than neural-only methods [16]. These problems can be alleviated by using approximated inference or cache-based approaches for symbolic rules. Moreover, the current approach assumes that the logic used is relatively simple (Horn clauses) [19][20].

The potential gains in interpretability have substantial implications for AI safety and trustworthiness. Through the explicit modeling of inference steps, this approach allows for auditing and debugging of models for flawed inference. This is consistent with new regulations on AI explainability.

6. Conclusion

In this work, the research introduces the Unified Neuro-Symbolic Reasoning Algorithms framework that combines first-order logic and transformers into one unified framework through attention-weighted logic gates. This ensures two-way communication between neural representations and symbolic knowledge. In experiments, the study observe improvements in logical consistency scores by 18-24%, better generalization, and higher interpretability. The framework is flexible enough to be customized based on the needs of different domains. Future directions involve the expansion to more sophisticated logics such as full first-order logic with negation and methods to improve efficiency. Utilization within actual application areas such as diagnosing illnesses, making legal arguments, and scientific discovery. Researching methods by which rules and weights for applying rules can be learned end-to-end from minimal supervision.

References

1. Liang, B., Wang, Y., & Tong, C. (2025). AI reasoning in deep learning era: From symbolic AI to neural–symbolic AI. Mathematics, 13(11), 1707.

2. Vattikundala, J., & Prasad, M. S. G. (2025). Application of hybrid & novel deep learning approaches for multimodal sentiment fusion in images & audio analysis. Archives for Technical Sciences, 3(34), 605–618. https://doi.org/10.70102/afts.2025.1834.605
3. Kim, J., & Cho, S. B. (2025, October). Neuro-symbolic reasoning with multiple large language models combined by first-order logic. In International Conference on Hybrid Artificial Intelligence Systems (pp. 227–238). Springer Nature Switzerland.
4. Sindhu, S. (2026). Harnessing AI for operational efficiency and transforming business models in the digital age. Global Tech Management Digest, 2(1), 1–6.
5. Onchis, D. M., Hogea, E., & Istin, C. (2025, September). From signals to symbols: Fault classification using transformers and logic-based reasoning. In International Conference on Mechanical System Dynamics (pp. 1–15). Springer Nature Singapore.
6. Mousavi, A., & Karshenasan, A. (2017). Forecasting stock prices of banks using artificial neural networks (GMDH). International Academic Journal of Accounting and Financial Management, 4(2), 71–78.
7. Zhu, D. (2026). A dual-process neuro-symbolic framework for cognition-driven language understanding. Applied Soft Computing, 114887.
8. Agrab, A. S. (2022). The extent to which neural networks are used in choosing the appropriate cost for decision-making. International Academic Journal of Economics, 9(1), 20–30. https://doi.org/10.9756/IAJE/V9I1/IAJE0903
9. OBIKE, P. G., USIP, P. U., UDO, E. N., & ANANGA, A. J. (2025). Neuro-symbolic reasoning: Performance, challenges, and benchmarks: A systematic literature review. ABUAD Journal of Engineering and Applied Sciences, 3(1), 36–47.
10. Narang, R. V., & Chatterjee, M. (2025). AI-powered knowledge management systems: A hybrid model for smart organizations. International Academic Journal of Innovative Research, 12(3), 14–19. https://doi.org/10.71086/IAJIR/V12I3/IAJIR1220
11. Davronova, Z., Shaxakimova, M., Kadirova, M., Atamurodov, S., Berdiyeva, H., Yoqubov, D., Rejapova, M., & Avazyazova, D. (2025). User-centered design approaches for scientific AI to trace semantic evolution of concepts across large corpora. International Journal of Aquatic Research and Environmental Studies, 5(2), 893–903. https://doi.org/10.70102/IJARES/V5I2/5-2-78
12. Shirke, S., et al. (2022). Hybrid optimisation dependent deep belief network for lane detection. Journal of Experimental & Theoretical Artificial Intelligence, 34(2), 175–187.
13. Chavda, R., Suresh, K., Kumar, S., Reddy, K. R., Jayaprakash, B., Sahu, P. K., et al. (2026). A bidirectional neuro-symbolic framework for clinical decision support via dynamic integration of deep learning and symbolic reasoning. Network Modeling Analysis in Health Informatics and Bioinformatics, 15(1), 79.
14. Hatmi, A. A., Klubnichkin, D. A., Shandodi, H. A., Mohammed, A. I. Q., Ali, A. A., & Audeh, Y. (2025). Why teach syntax? The importance of surface and deep structures for full comprehension of sentences. International Journal of English and Education, 14(1), 35–55.
15. Rathi, S., Mirajkar, O., Shukla, S., Deshmukh, L., & Dangare, L. (2024). Advancing crack detection using deep learning solutions for automated inspection of metallic surfaces. Indian Journal of Information Sources and Services, 14(1), 93–100. https://doi.org/10.51983/ijiss-2024.14.1.4003
16. Nguyen Hong, S., Tran Ngoc, T., & Lê, N. C. (2024, December). Logic neural network: A short survey. In International Conference on Applied Mathematics and Computer Science (pp. 156–167). Springer Nature Switzerland.
17. Narayanan, L., & Rajan, A. (2024). Artificial intelligence for sustainable agriculture: Balancing efficiency and equity. International Journal of SDG's Prospects and Breakthroughs, 2(1), 4–6.
18. Sheng, Y., Li, L., Wang, Y., & Zeng, D. (2024, April). Integrating language models with symbolic formulas for first-order logic reasoning. In ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 11586–11590). IEEE.
19. Rahman, F. (2026). Adaptive statistical signal processing framework for robust pattern recognition in noisy environments. National Journal of Integrated VLSI and Signal Intelligence, 9–17.