12-2-2022

Context-aware Collaborative Neuro-Symbolic Inference in Internet of Battlefield Things

Tarek Abdelzaher

Nathaniel D. Bastian
Army Cyber Institute, U.S. Military Academy, nathaniel.bastian@westpoint.edu

Susmit Jha

Lance Kaplan

Mani Srivastava

See next page for additional authors

Authors
Tarek Abdelzaher, Nathaniel D. Bastian, Susmit Jha, Lance Kaplan, Mani Srivastava, and Venugopal Veeravalli

Context-aware Collaborative Neuro-Symbolic Inference in IoBTs

Tarek Abdelzaher
Department of Computer Science
University of Illinois, Urbana-Champaign
zaher@illinois.edu

Nathaniel D. Bastian
Army Cyber Institute
U.S. Military Academy
nathaniel.bastian@westpoint.edu

Susmit Jha
Computer Science Laboratory
SRI International
susmit.jha@sri.com

Lance Kaplan
DEVCOM Army Research Laboratory
U.S. Army Futures Command
lance.m.kaplan.civ@army.mil

Mani Srivastava
ECE Department
University of California, Los Angeles
mbs@ee.ucla.edu

Venugopal V. Veeravalli
ECE Department
University of Illinois, Urbana-Champaign
vvv@illinois.edu

***Abstract*—IoBTs must feature collaborative, context-aware, multi-modal fusion for real-time, robust decision-making in adversarial environments. The integration of machine learning (ML) models into IoBTs has been successful at solving these problems at a small scale (e.g., AiTR), but state-of-the-art ML models grow exponentially with increasing temporal and spatial scale of modeled phenomena, and can thus become brittle, untrustworthy, and vulnerable when interpreting large-scale tactical edge data. To address this challenge, we need to develop principles and methodologies for uncertainty-quantified neuro-symbolic ML, where learning and inference exploit symbolic knowledge and reasoning, in addition to, multi-modal and multi-vantage sensor data. The approach features integrated neuro-symbolic inference, where symbolic context is used by deep learning, and deep learning models provide atomic concepts for symbolic reasoning. The incorporation of high-level symbolic reasoning improves data efficiency during training and makes inference more robust, interpretable, and resource-efficient. In this paper, we identify the key challenges in developing context-aware collaborative neuro-symbolic inference in IoBTs and review some recent progress in addressing these gaps.**

***Index Terms*—Neuro-symbolic inference, robust learning**

I. Introduction

A central challenge in deploying artificial intelligence (AI) and machine learning (ML) in IoBTs is the lack of principled approaches for context-aware (collaborative) multi-modal fusion and robust decision-making in adversarial environments. The scale and speed of acquisition, assessment, aggregation, state estimation, and decision-making in IoBTs operating in a rapidly-evolving high-tempo and adversarial battlefield environment necessitate the use of human-on-the-loop artificial intelligence (AI) across the span of short-time scale perception to high-level C3I decision-support. But AI/ML models are known to be vulnerable to adversarial attacks and lack generalization – a problem that increases in severity as the span of modeled phenomena increases in time and in space. Unlike commercial applications, the failure of AI/ML in IoBTs can have catastrophic consequences. At the same time, the resources available for executing AI/ML models may be significantly more constrained. Hence, a responsible, safe, and ethical use of machine intelligence in the context of multi-modal fusion (within the MDO effects loop in IoBTs) requires innovations that substantially improve efficiency, while at the same time offering uncertainty/confidence quantification in results. This challenge requires the development of new methods to significantly advance both resource efficiency and confidence estimation in distributed IoBTs. We posit that neuro-symbolic learning and inference (that combines symbolic reasoning with uncertainty-quantified deep learning) can achieve this end, and identify challenges, develop hypotheses, and present initial indicative results.

While traditional machine learning typically relies on purely bottom-up inference from sensors treating each observation as independent uncorrelated input, our neuro-symbolic ML approach interleaves bottom-up inference with top-down predictions from the learned neuro-symbolic context. Any surprise arising from the mismatch between the top-down prediction and the bottom-up inference is used for self-supervised training and continual adaptation. This neuro-symbolic approach is aided by techniques for uncertainty quantification to detect out-of-distribution (OOD) and novel inputs. The top-down inference using context and well-calibrated uncertainty quantification facilitates distributed inference, where edge sensor resources are queried to corroborate inferences and do not need to stream data continuously. Moreover, the exploitation of symbolic knowledge dramatically reduces the required neural network model size while allowing models to reason about phenomena that extend over substantial ranges in time and/or in space. Our hybrid neuro-symbolic approach to machine learning is thus particularly suited for heterogeneous distributed IoBT nodes, allowing us to select the appropriate combination of symbolic reasoning and deep learning from the spectrum of neuro-symbolic methods depending on the resource limitations, modality of the data, and the availability of background data. Thus, neuro-symbolic machine learning can provide tactical edge coordination to increase scalability, corroboration, and context-aware intelligence.

Funding ack to be added in the final version

II. Challenges and Key Research Questions

In this section, we identify the main challenges in developing a context-aware collaborative neuro-symbolic inference in IoBTs and summarize these challenges into two central research questions.

Research Question 1. *How do we develop scalable context-aware neuro-symbolic learning and inference that span the spectrum of combinations of symbolic reasoning with data-driven deep learning, and enables the selection of an appropriate combination for a given resource-constrained IoBT?*

Inference in IoBTs requires collaborative multi-modal fusion where sensors are not continuously streaming data but, instead, there is a context-dependent sharing of data over a dynamic and contested IoBT network. Traditional machine intelligence techniques have successfully fused data and shared insights when the measured phenomena were localized in space and time (e.g., Aided Target Recognition). However, the requirements of such models grow exponentially as the size of the modeled phenomenon increases, leading to substantial scalability problems. Neuro-symbolic methods naturally represent knowledge at varying levels of abstraction, allowing the symbolic representation to model a larger and interpretable context that can be fused with background knowledge to allow reasoning over longer spatial and temporal scales. Thus, these approaches can be used for context-aware sensing and exploit the flexibility of heterogeneous, multi-modal, and multi-vantage data in IoBTs. The use of context and corroboration of predictions from different data sources and modalities also makes inference robust.

While specific neuro-symbolic ML approaches have been developed previously, the heterogeneity of IoBT nodes requires a flexible neuro-symbolic framework that can span different levels of the combination of symbolic reasoning and data-driven learning. At one end of the spectrum, we can consider mostly symbolic inference using rules over entities that are identified using deep learning, and at the other end, we can consider methods where logical rules are compiled into the deep learning presentation either as a regularizing loss function or a differentiable DNN representation. The selection of appropriate combinations for different inference tasks and resource constrained IoBT nodes would help create an adaptive IoBT.

Research Question 2. *How do we create and integrate neuro-symbolic ML models into IoBTs with guarantees on their predictions, robustness to new environments, and resilience to adversarial examples?*

The battlefield environment and context of an IoBT will change rapidly, and hence, the responsible deployment of ML in IoBTs necessitates the detection of OOD and novel inputs, and quick adaptation of the ML models. Beyond this lack of robustness, DNNs are also susceptible to adversarial attacks. Physically realizable attacks can exploit this vulnerability without cyberattacks. An adversary will readily exploit any vulnerability on a battlefield; hence, ML in IoBTs must be resilient to adversarial attacks. Traditional approaches to building high-assurance systems using formal methods and control theory are insufficient to reason about ML models. These shortcomings create unique challenges for ensuring the robustness and resilience of ML-enabled IoBTs. In particular, the inference in neuro-symbolic models does not follow the usual acyclic propagation in feedforward neural networks, which creates new challenges for uncertainty quantification. Thus, there is a need to develop methods to analyze the prediction performance and robustness of neuro-symbolic models.

We formulate the following two hypothesis that forms the basis of our technical approach to address the above research questions:

- A multi-layered neuro-symbolic architecture inspired by Predictive Processing (PP) - a theory of mind, will enable context-aware data-efficient robust ML models with tight integration between symbolic reasoning and deep learning that can be tailored to resource constraints on an IoBT node and the needs for specific task and modality.
- Robust and resilient ML models must be able to detect "surprise" by observing their own inference patterns when subjected to novel or adversarial inputs. Information-theoretic approaches can be combined with computationally lightweight runtime monitors to detect such surprises and yield guarantees on IoBT's robustness and resilience.

We elaborate on these two hypotheses below. First, the PP-inspired architecture relies on building a "world model" that captures context (such as spatiotemporal relationships) and uses this context to hypothesize and confirm predictions over the sensor data. We use a hierarchical representation of the world model, varying from neural models (capturing local context but at a higher level of detail) to symbolic abstract models (capturing broader contexts that may extend more broadly in space and time). Different learning and inference tasks and modalities running on heterogeneous IoBT nodes can benefit from varying extent of symbolic knowledge. For example, tracking a vehicle can exploit well-understood physics models while detecting a vehicle of a particular type can rely on "is-part-of" relationships such as "wheel is-part-of vehicle" or spatial co-occurrences such as vehicles co-occur with roads. This choice of the right level of fusing symbolic knowledge with deep learning can also enable resource-efficient inferences on different IoBT nodes with varying computational power.

Second, the safe and trustworthy integration of neuro-symbolic ML models into IoBTs requires methods to characterize the generalization and robustness of these models, and the creation of runtime monitors that detect when the environment has evolved outside the training context and the models cannot be trusted. We hypothesize that runtime monitors can observe the inference pattern of the learning model to detect inputs that are surprising to the model, and such monitors are more effective in detecting OOD and adversarial inputs when compared with traditional anomaly detection approaches that attempt to learn the training distribution without taking into account the model. Developing such monitors will require sound and systematic generation and a combination of statistics to measure the surprise of a model. Further, the deployment of the ML models and these runtime monitors on the IoBT edge nodes require these models to be computationally lightweight. Thus, generating fast-to-evaluate monitors is critical to making ML models more robust to novel and adversarial inputs without sacrificing inference speed.

III. Related Work

A plethora of ad-hoc neuro-symbolic approaches have been proposed in the literature. A common approach is to use differentiable approximations of logical formulas, such as the CLIP system [10] (developed in 1990s and precursor to Logic Tensor Networks [24]), ILP [8], DFOL [9], and our prior work Neuroplex [30]. Such approximations without guarantees on the gap between original and differentiable representations can be created through either direct fuzzy logic-style compilation, such as in LTN [28] or via projected-gradient/mirror descent, such as program induction [6], and our prior work [11]. Another approach arising from Statistical Relational AI [18] is to use the categorical output of deep neural networks as atomic predicates of a probabilistic and/or logical reasoning system, such as DeepProbLog [26] and our prior work DeepProbCEP [22]. These approaches have shallow integration between neural and symbolic representations with limited scalability, and without any assurance guarantees.

Apart from the focus on assurance, we argue the need for iterative training and inference in neuro-symbolic learning and inference. Most neuro-symbolic methods are at one of the two ends of the spectrum of how logical symbolic models are integrated with deep learning. At one end are the techniques such as DeepProbLog [26], and program induction [6], where the symbolic inference using rules or programs are layered on the top of entities recognized by deep learning models. At the other end, there are methods such as LTN [28] and TNNs [11] that compile logic into deep learning representations (as a regularizing loss or differential DNN representation).We hypothesize that an ideal neuro-symbolic approach would not use such a layered architecture. Instead, the neuro-symbolic inference needs to be bidirectional – learning-based bottom-up push that is uncertainty-driven and reasoning-based symbolic top-down pull that is decision-driven. This allows a more flexible architecture where both the neural and symbolic layers make an integrated inference. This is crucial for assurance and robustness as it avoids any assumption by symbolic layer on the correctness of the neural inferences, or vice versa.

Several ML models have been recently proposed based on hierarchical predictive processing (PP) including our own prior work on Trinity [1], [14], [16], [23], which was one of the first practical implementation of PP. These recent attempts to implement PP or HPP [3], [12], [19] are not neuro-symbolic and unable to benefit from symbolic knowledge or reasoning. We advocate a neuro-symbolic architecture with bidirectional flow of information supported by equilibration [2] for training and inference, instead of stitching separately trained layers.

IV. Technical Approach and Initial Results

A cognitive architecture aimed at continuous learning, reasoning, comprehension, and robust inference in contested environments such as IoBTs must provide the necessary goal-directed knowledge representation, inference, and decision-making mechanisms to allow distributed inference where edges nodes of the IoBT can quickly adapt. The heterogeneity of IoBT nodes in their resource limitations and modality of data requires an ML paradigm that supports a flexible cognitive architecture with varying knowledge representation, goal structure, and inference methods. Motivated by this need to consider varying levels of information, we have developed TrinityAI [1], [4], [13]–[17], [23], [29] framework for trustworthy, resilient, and interpretable AI. We sketch this framework and the planned extensions that will address the challenges identified in this paper for developing a context-aware collaborative neuro-symbolic learning and inference approach.

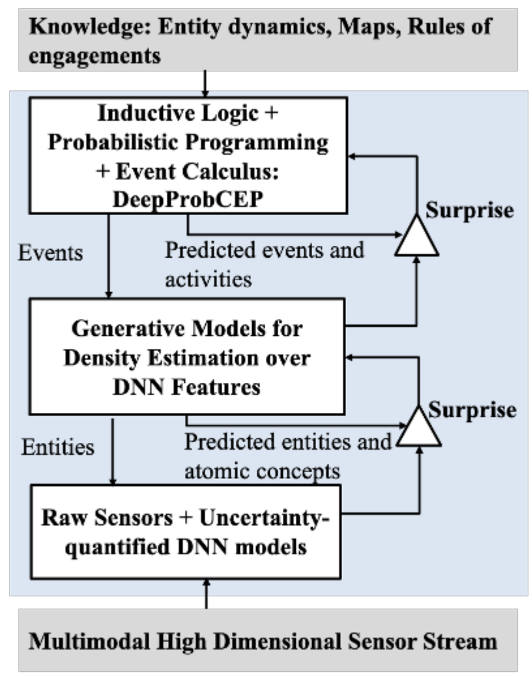

Fig. 1. Overview of the neuro-symbolic TrinityAI Architecture

TrinityAI (Figure 1 uses an architecture akin to dual processing theory [7] in human psychology (often referred to as System-1 and System-2) where System-1 is a reflex neural layer, and System-2 is a more deliberative symbolic layer. System-1 is built on a generative distribution learning model (based on our recently proposed Principal Component Flows [4]) that learns the conditional distribution of intermediate features from high dimensional data. System-2 is a deliberative symbolic AI implemented using Graph Neural Networks [20] or ProbLog [5] – a combination of inductive logic [25] logical modeling and probabilistic programming. Based on our prior work [27], we customize ProbLog with predicates corresponding to elements in event calculus for inference tasks related to perception and situation understanding. A key technical component of our approach is a principled approach to quantifying uncertainty of such a neuro-symbolic ML architecture.

Another significant challenge is to explore how the interactions between the reflex layer (System-1) and the deliberative longer timescale symbolic layer (System-2) should inform data communication in the distributed IoBT system. Addressing this challenge will result in algorithms and theoretical foundations for tactical edge coordination (over a dynamic and contested network) that determine what data to share, at what level of abstraction, and with which entities to optimize mission performance. Our approach to building neuro-symbolic models can be viewed as an example of “analysis by synthesis” [31], meaning that we formulate hypotheses (i.e., candidate world models) and favor those whose predictions match the input data.

In our recent work on TrinityAI [23], we built such a neuro-symbolic model. The first layer in Trinity currently uses a feature density model (implemented via either simple statistical models such as Mahalanobis distance or using more expressive deep learning-based distribution models such as extensions of normalizing flows [4]) that learns a joint distribution of the input, predicted class output and the explanation provided by the model. The second layer uses a symbolic model such as graph Markov network to learn spatial and temporal relationships between the objects for an object detection task.

In our initial results, this architecture is shown to exhibit several characteristics suitable for IoBT:

1) **Robustness**: Layer 1 detects OOD, novel and adversarial inputs. For example, for a WideResNet model trained on CIFAR10 and SVNH as OODs, our approach detects 86.8% of OODs, compared to the state-of-the-art 57.2%. OOD detection has been also extended to other modalities such as audio, time-series data, and video [17] (for e.g., drifting car in Figure 2 where OODness is temporal and not in a single frame).

Fig. 2. Temporal Novelty corresponds to novelty in the time-series data where individual states (such as frames of a video) are in distribution, but the sequence of states is an OOD.

2) **Context-awareness**: Layer 2 makes our approach context-aware, and it can also detect out-of-context inputs [1] where the individual components (such as objects in a scene) are not new, but they appear in a context that is novel and is likely to confuse even a context-based ML model. Figure 3 shows examples of such out-of-context inputs for an object detection problem.

Fig. 3. Example of out-of-context inputs where the objects are in-distribution but they are composed in a manner that creates out-of-context scenes, such as an airplane in a room or a keyboard where we expect a traffic sign. TrinityAI [1] can detect such anomalies.

3) **Data-efficiency:** The context-aware ML model exhibits better generalization and is able to correctly identify objects even with small data. We evaluated this using NuScenes dataset with traffic images from Boston and Singapore. The dataset is unbalanced with less than 1% examples of bicycles, some of which are occluded (see Figure 2c). Our approach improves detection of bicycles from 2.4% to 31.4% and 12.5% to 66.6% for 50% and 30% occlusions while retaining a high overall accuracy of 95.5%.

Fig. 4. Classes such as bicycles are underrepresented in the NuScenes dataset. Real-world applications such as IoBT will often have unbalanced datasets

4) **Adaptive communication, coordination, and collaborative decision-making**: For an active, resilient, communication-efficient and fighting IoBT, we envision a fundamental shift in how sensor data is communication, collected and used for decision-making. Instead of all sensors streaming their observations to processing nodes, we will investigate the use of the principle of predictive processing in designing the IoBT architecture. The individual decision-making nodes running machine learning models based on predictive processing can query sensors to provide information needs to confirm their current hypothesis and in the event of surprise, query additional sensors or other learned models to provide further observations and evidence. This would reduce the required communication in the IoBT and make it an active network that builds model of the world and queries sensors when needed to update its model and make decisions. Thus, the sensing and communication in IoBTs is guided by the uncertainty quantification of the machine learning models.

V. FUTURE RESEARCH DIRECTIONS

In this section, we identify directions of future research that builds on our technical approach and addresses the research challenges in developing a collaborative context-aware neuro-symbolic robust machine learning and inference in IoBTs. We categorize these future directions into three thrusts.

- **Integration of deep learning with a richer symbolic reasoning and inference framework**: Building on our recent work on DeepProbCEP [27], we argue for extending our Layer 2 symbolic models from graphical models to a richer logical reasoning framework. This is critical to data efficiency in IoBTs by incorporating knowledge. It is also vital to support reasoning to derive context and enable context-sensitive adaptation that considers the resource constraints of the IoBT nodes. This will also go beyond the initial generation of two-tier symbolic-after-neural neuro-symbolic architectures with hub-spoke organization to develop more general neuro-symbolic architectures suited for IoBT settings. Specifically, leveraging advances in neuro-symbolic architecture search, translation across neural and logic domains, and efficient tensorized implementation of logic models, this integration will develop neuro-symbolic architectures that harness richer information flows among collaborating IoBT edge devices to meet resource constrained. Lastly, we argue for developing methods by which the neuro-symbolic architectures can be rapidly fine-tuned to environment characteristics and tactics, techniques, and procedures of a new IoBT deployment via an approach that combines injection of expert knowledge into symbolic components with transfer learning-based adaptation of neural components. This integration of symbolic reasoning will incorporate background knowledge and symbolic reasoning for data-efficiency, and generalization to hub-spoke neuro-symbolic architecture. It will also enable adapting neuro-symbolic framework to resource constraints of the IoBT nodes where inference needs to be performed.
- **Principled uncertainty-quantification of neuro-symbolic models:** Prior work has mostly been focused on identifying promising test statistics and corresponding thresholds, motivated primarily by empirical observations of the values taken by these statistics. In our recent work [17], [21], we have attempted to provide guarantees on false alarm detection for such methods using methods from conformal prediction and multiple testing. Continuing this line of investigation, we suggest the need to develop principled approaches to combine test statistics that reduce the computational burden in learning distributions using complex models such as normalizing flows. There is a need to focus on methods for uncertainty propagation in neuro-symbolic models with iterative cyclic inference. In particular, we can use prediction sets at the perception layers, satisfying explicit finite-sample guarantees on uncertainty, which are refined through iterations with information from the symbolic layers. Uncertainty quantification in neuro-symbolic learning is a challenging and relatively explored area in machine learning requiring the development of basic sciences and fundamental research.
- **Neuro-symbolic-architecture-inspired algorithms and foundations of edge coordination**: A remarkable analogy exists between neuro-symbolic architecture and what's known as System-1 and System-2 in human psychology, where the former refers to fast instinctive processing, whereas the latter to more deliberate cognitive processing (often at a slower time-scale). Often the fast intuitive processing occurs below the threshold of conscious awareness, whereas anomalies and unusual events (besides invoking a local System-1 reaction) are reported at some higher level of abstraction to invoke a more deliberative slower cognitive response. This analogy offers a starting point for designing the strategy for allocating neural (System-1) and cognitive/symbolic (System-2) components in the distributed IoBT system and for orchestrating the communication among them. In such an architecture, common-case processing is delegated to the local components, whereas uncommon events are processed more globally at a higher-level of (symbolic) abstraction, thus minimizing network use in the common case, and offering some autonomy to distributed local components. We need to build on anomaly/quickest-change detection (to detect conditions for a bottom-up information push) as well as decision-driven communication (that produce a symbolic top-down pull).

While the context-aware collaborative neuro-symbolic inference approach advocated here is driven by building a neuro-symbolic model of the world, quantifying uncertainty in perception and risk-aware decision-making, we emphasize that our approach does not require any explicit manual building of the environment. Complex and rapidly evolving environments such as IoBTs cannot be modeled manually. Instead, our approach relies on maintaining a continuous machine learned model of the environment. Using a neuro-symbolic representation of this model makes our approach:

- interpretable that is necessary for IoBT application which often requires symbiotic human-in-the-loop or human-on-the-loop decision-making
- data-efficient because the amount of supervision available to quickly learn in a rapidly evolving environment is sparse and the use of background symbolic knowledge is crucial
- robust because the purely neural representations that do not capture the full context are known to be fragile to small perturbations

The development of this context-aware neuro-symbolic approach needs a validation approach that can be used to drive research across these different thrusts. Such a validation can use a combination of simulators such as AirSim, CARLA, and Gazebo. CARLA simulator provides data from a variety of modalities including RGB/depth cameras, LiDAR, Radar, IMU and GNSS. In addition to multimodal simulation, we can use a plethora of multimodal datasets such as the Ford AV Dataset with camera/3LIDAR modalities recorded in Michigan, nuScenes dataset comprising visual, 3D LiDAR and Radar modalities and KAIST multispectral pedestrian dataset comprising visual and thermal camera.

VI. Conclusion

The complexity and the tempo of contested and conflict areas such as battlefields or even purely cyberspaces such as computer networks, have created demands for rapid and precise information dissemination, and created unique challenges in Command, Control, Communication, Computers, Cyber, Intelligence, Surveillance and Reconnaissance (C5ISR). While a large volume of sensed data gathered through a plethora of sensors across different modalities are readily available, the time available to interpret and understand this information correctly and robustly is becoming prohibitively smaller. This necessitates the development of cognitive processing capability that can integrate background knowledge with learning from data.

The adoption of AI capabilities in warfighting will make the battlefield even more dynamic and rapidly evolving beyond human capability of comprehension and reaction. ML models are crucial to facilitate holistic perception of conflict that integrates information from different heterogeneous sensors and ensures timely autonomous decision-making to implement command-by-intent. This requires ML models to be robust and resilient to change in battlefield environments and adversarial perturbations. Our neuro-symbolic approach to addressing this challenge focuses on IoBT-specific requirements of robustness to natural and adversarial perturbations, data-sparsity, low supervision, addressing rapid change in environment from the training distribution, and ensure compliance with rules of engagement and safety requirements.

References

[1] Manoj Acharya, Anirban Roy, Kaushik Koneripalli, Susmit Jha, and Christopher Kanan. Detecting out-of-context objects using graph contextual reasoning network. *31st International Joint Conference on Artificial Intelligence (IJCAI)*, 2022.

[2] Margaret A Boden. Is equilibration important?—a view from artificial intelligence. *British Journal of Psychology*, 73(2):165–173, 1982.

[3] Choksi, Mozafari, Biggs O'May, and others. Predify: Augmenting deep neural networks with brain-inspired predictive coding dynamics. *Processing*, 2021.

[4] Cunningham, Adam Cobb, and Susmit Jha. Principal manifold flows. February 2022.

[5] Luc De Raedt, Angelika Kimmig, and Hannu Toivonen. ProbLog: A probabilistic prolog and its application in link discovery. In *IJCAI*, volume 7, pages 2462–2467, 2007.

[6] Ellis, K., Morales, L., Sablé-Meyer, M., Solar-Lezama, A. and Tenenbaum, J. Learning libraries of subroutines for neurally–guided bayesian program induction. *NeurIPS*, 2018.

[7] Jonathan St BT Evans and Jodie Curtis-Holmes. Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, 11(4):382–389, 2005.

[8] Richard Evans and Edward Grefenstette. Learning explanatory rules from noisy data. *J. Artif. Intell. Res.*, 61:1–64, January 2018.

[9] Kun Gao, Katsumi Inoue, Yongzhi Cao, and Hanpin Wang. Learning First-Order rules with differentiable logic program semantics. April 2022.

[10] A S D'avila Garcez, A S d'Avila Garcez, G Zaverucha, and V N A L da Silva. Applying the connectionist inductive learning and logic programming system to power system diagnosis.

[11] Shalini Ghosh, Amaury Mercier, Dheeraj Pichapati, Susmit Jha, Vinod Yegneswaran, and Patrick Lincoln. Trusted neural networks for Safety-Constrained autonomous control. In *Machine Learning for Safety-Critical Applications in Engineering at DISE/ICML*, 2018.

[12] Dimitrios C Gklezakos and Rajesh P N Rao. Active predictive coding networks: A neural solution to the problem of learning reference frames and Part-Whole hierarchies. January 2022.

[13] Uyeong Jang, Susmit Jha, and Somesh Jha. On the need for Topology-Aware generative models for Manifold-Based defenses. In *8th International Conference on Learning Representations (ICLR)*, 2020.

[14] Susmit Jha, Brian Jalaian, Anirban Roy, and Gunjan Verma. Trinity: Trust, resilience and interpretability of machine learning models. In *Game Theory and Machine Learning for Cyber Security*, pages 317–333. Wiley, September 2021.

[15] Susmit Jha and John Rushby. Inferring and conveying intentionality: Beyond numerical rewards to logical intentions. *AAAI Symposium on Towards Conscious AI Systems*, 2019.

[16] Susmit Jha, John Rushby, and Natarajan Shankar. Model-Centered assurance for autonomous systems. In *Computer Safety, Reliability, and Security*, pages 228–243. Springer International Publishing, 2020.

[17] R Kaur, Susmit Jha, Anirban Roy, Park, Dobriban, and others. iDECODe: In-distribution equivariance for conformal out-of-distribution detection. *36th AAAI Conference on Artificial Intelligence (AAAI)*, 2022.

[18] Kersting, Natarajan, and Poole. Statistical relational AI: Logic, probability and computation. In *International Conference on Logic Programming and Nonmonotonic Reasoning*.

[19] Yann Lecun. A path towards autonomous machine intelligence version 0.9.2, 2022-06-27. https://openreview.net/pdf?id=BZ5a1r-kVsf, 2022. Accessed: 2022-7-25.

[20] Ruoyu Li, Sheng Wang, Feiyun Zhu, and Junzhou Huang. Adaptive graph convolutional neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[21] Akshayaa Magesh, Venugopal V Veeravalli, Anirban Roy, and Susmit Jha. Multiple testing framework for out-of-distribution detection. *arXiv preprint arXiv:2206.09522*, 2022.

[22] Alun D Preece, Dave Braines, Federico Cerutti, Jack Furby, Liam Hiley, Lance Kaplan, Mark Law, Alessandra Russo, Mani Srivastava, Marc Roig Vilamala, and Tianwei Xing. Coalition situational understanding via explainable neuro-symbolic reasoning and learning. In Tien Pham, Latasha Solomon, and Myron E Hohil, editors, *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III*. SPIE, April 2021.

[23] Anirban Roy, Adam Cobb, Nathaniel D Bastian, Brian Jalaian, and Susmit Jha. Runtime monitoring of deep neural networks using top-down context models inspired by predictive processing and dual process theory. In *AAAI Spring Symposium*.

[24] Luciano Serafini and Artur D'avila Garcez. Logic tensor networks: Deep learning and logical reasoning from data and knowledge. June 2016.

[25] Pietro Totis, Angelika Kimmig, and Luc De Raedt. SMProbLog: Stable model semantics in ProbLog and its applications in argumentation. October 2021.

[26] M R Vilamala, T Xing, H Taylor, L Garcia, and others. Using DeepProbLog to perform complex event processing on an audio stream. *arXiv preprint arXiv*, 2021.

[27] Vilamala, M.R., Xing, T., Taylor, H., Garcia, L., Srivastava, M., Kaplan, L., Preece, A., Kimmig, A. and Cerutti, F. Using DeepProbLog to perform complex event processing on an audio stream. 2021.

[28] B Wagner and A S d'Avila Garcez. Neural-Symbolic reasoning under Open-World and Closed-World assumptions. In *CEUR Workshop Proceedings*, volume 3121. CEUR, March 2022.

[29] Walmer, Sikka, Sur, and others. Dual-Key multimodal backdoors for visual question answering. *Proc. Estonian Acad. Sci. Biol. Ecol.*

[30] Tianwei Xing, Luis Garcia, Marc Roig Vilamala, Federico Cerutti, Lance Kaplan, Alun Preece, and Mani Srivastava. Neuroplex: learning to detect complex events in sensor networks through knowledge injection. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 489–502. Association for Computing Machinery, New York, NY, USA, November 2020.

[31] Alan Yuille and Daniel Kersten. Vision as bayesian inference: analysis by synthesis? *Trends in cognitive sciences*, 10(7):301–308, 2006.