Neuro-Symbolic Decision Making for Autonomous Agents

Doctoral Consortium

Celeste Veronese
University of Verona
Verona, Italy
celeste.veronese@univr.it

ABSTRACT

Planning and sequential decision-making in complex environments are central challenges in Artificial Intelligence, particularly when action spaces are large, rewards are sparse, and long-horizon reasoning is required. Deep Reinforcement Learning (DRL) has shown strong performance in many domains but suffers from limited interpretability, sample inefficiency, and poor generalization in these settings. Symbolic reasoning and logic-based methods offer complementary strengths by providing structured, human-interpretable representations that can guide decision-making and enhance transparency. This extended abstract presents the research topic of my PhD, which investigates how symbolic knowledge can be integrated with planning and DRL within a Neuro-Symbolic (NeSy) AI framework to overcome the limitations of purely black-box approaches. By leveraging symbolic heuristics to guide exploration and decision-making, while neural components provide scalability and noise robustness, my research aims to support effective, explainable, and generalizable decision-making in complex domains.

KEYWORDS

Neuro-Symbolic Decision Making; Knowledge Transfer; Symbolic Knowledge; Knowledge Representation

ACM Reference Format:

Celeste Veronese. 2026. Neuro-Symbolic Decision Making for Autonomous Agents: Doctoral Consortium. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/PVRI5073

1 RESEARCH PROBLEM AND MOTIVATION

Efficient planning and sequential decision-making in complex domains remain a fundamental challenge in Artificial Intelligence (AI). Tasks with large action spaces, long planning horizons, sparse rewards, or complex relational structures require agents to explore efficiently and reason over extended action sequences. Traditional approaches rely on heuristics that can be effective but are typically either handcrafted, requiring significant domain expertise, or learned via neural methods, which can be opaque, data-hungry, and limited in their ability to generalize [1, 8, 16].

Deep Reinforcement Learning (DRL) has emerged as a powerful tool for addressing these challenges, demonstrating success in domains ranging from robotics [5] to sustainability [24]. Nonetheless, DRL also brings several limitations: policies are black-box and difficult to interpret, learning is often sample-inefficient, and generalization to unseen or long-horizon environments remains challenging [3, 22]. Heuristic-guided DRL can partially mitigate these issues, but existing approaches, such as reward shaping or reward machines, are sensitive to heuristic quality and often computationally costly [2, 17].

On the other hand, symbolic reasoning and logic-based methods offer complementary advantages in planning and sequential decision-making. Such approaches allow agents to represent relational and structured knowledge in a compact, interpretable form, supporting reasoning over complex tasks and long sequences of actions [9, 11, 15]. By explicitly encoding domain knowledge, symbolic methods can provide guidance, improve efficiency, and enhance transparency, which is particularly valuable in data-scarce or safety-critical scenarios. While these approaches traditionally rely on expert knowledge, recent advances have explored ways to extract or learn symbolic knowledge from experience, leveraging Inductive Logic Programming (ILP) approaches [4, 10, 12].

In this context, my research focuses on integrating symbolic reasoning and learning-based methods within the Neuro-Symbolic (NeSy) AI framework. The goal is to combine structured, human-interpretable knowledge with data-driven learning in order to support effective decision-making in complex domains. In particular, I study how symbolic policy heuristics, mainly learned through ILP, can be integrated into planning and DRL systems. These heuristics are used to guide exploration and decision-making, while neural components provide scalability and robustness to noise. Overall, this line of research aims to overcome the limitations of purely black-box approaches and to develop agents that can operate reliably in long-horizon tasks, producing policies that are interpretable, trustworthy, and able to generalize from limited experience.

2 MAIN RESEARCH CONTRIBUTIONS

My research contributes to the state of the art in NeSy decision making by studying how symbolic knowledge can be learned and integrated into planning and RL systems to improve interpretability, efficiency, and generalization. The main contributions I developed are summarized below.

Learning human-interpretable symbolic policies. As an initial contribution, I investigated the use of ILP to learn human-interpretable representations of policies generated by RL agents. By observing execution traces of a multi-objective RL agent in the autonomous driving domain [13], I showed that ILP can successfully extract symbolic specifications capturing socially acceptable decision-making criteria. These learned representations improve

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/PVRI5073

policy interpretability and can also be exploited to implement an ASP planner achieving performance comparable to the original RL agent. This work demonstrated the feasibility of using ILP to bridge black-box RL policies and symbolic reasoning [18].

Learning temporal symbolic heuristics for planning. Building on these results, I extended prior work on logic-based heuristics to handle richer temporal representations by learning domain-dependent theories in Event Calculus (EC) [6]. While previous approaches relied on atemporal symbolic abstractions, this work applies ILP-based heuristic learning to explicitly model the temporal dimension of planning domains through temporal logic. The proposed method learns persistent, time-extended macro-actions from a small number of execution traces using ILP, enabling reasoning over temporally extended dependencies that cannot be captured by time-independent representations. These learned temporal heuristics are then used to guide Monte Carlo Tree Search (MCTS)-based planners, such as POMCP [14] and DESPOT [23], improving both computational efficiency and solution quality.

Empirical evaluations in benchmark domains, including Pocman and Rocksample, show that the learned temporal macro-actions are more expressive and general than time-independent heuristics, while requiring minimal prior knowledge beyond the domain model. This demonstrates that automatically learned temporal symbolic knowledge can effectively guide online planning under uncertainty. This work was presented at the Conference on Neuro-Symbolic Learning and Reasoning (NeSy 2025) [19].

Online learning of symbolic heuristics for RL exploration. As a next step, I developed an online neuro-symbolic methodology for learning and refining policy heuristics during tabular RL training. In contrast to previous neuro-symbolic approaches, where symbolic knowledge is learned offline from fixed datasets or pre-collected traces, this work performs the induction and continuous refinement of symbolic heuristics online, directly during agent-environment interaction. The approach leverages a scalable ILP system, FastLAS [7], to induce ASP heuristics from batches of experience collected by the RL agent. State–action trajectories are mapped to high-level symbolic concepts, enabling the learning of an interpretable logical approximation of the agent's evolving policy.

Rather than shaping the reward function, the learned heuristics are used to softly bias the exploration process through probabilistic reasoning, preserving the asymptotic convergence guarantees of the underlying RL algorithm while allowing the symbolic knowledge to adapt dynamically as training progresses. The integration results in a significant improvement in discounted return and faster convergence, with limited computational overhead. Moreover, the learned symbolic rules provide a coherent and compact explanation of the underlying black-box policy. This contribution was presented at the PRL workshop (ICAPS'24) and at the HYDRA workshop (ECAI'24) [20].

Integrating symbolic heuristics into Deep Reinforcement Learning. Since previous work was limited to tabular RL algorithms, I further investigated how symbolic knowledge can be integrated directly into the learning dynamics of Deep Reinforcement Learning (DRL). This work employs previously acquired symbolic heuristics to guide the DRL learning process at algoritmic level, without modifying the reward function and embedding symbolic reasoning directly into the action selection process. The proposed framework extracts symbolic heuristics from simpler problem instances and transfers them to guide training in more complex environments. Logical specifications approximating learned policies are used to identify promising actions, supporting both exploration and exploitation in ϵ-greedy DRL algorithms.

By influencing the learning mechanism itself rather than indirectly shaping behavior through rewards, the approach improves sample efficiency in domains with long planning horizons, sparse rewards, and multiple sub-goals. The experimental evaluation demonstrates improved learning efficiency and performance compared to reward-shaping neuro-symbolic baselines. This work has been accepted for publication as an Extended Abstract at AAMAS 2026. [21]

3 OPEN DIRECTIONS

The results I obtained so far clearly highlighted the potential of NeSy approaches in improving planning and RL systems in terms of interpretability, efficiency, and robustness. Building on this, my future work will aim to further broaden and consolidate the integration between symbolic reasoning and learning-based methods across a wider range of decision-making settings.

The first open direction concerns the study of how symbolic knowledge can be exploited at different stages of the learning and decision-making process, beyond guiding exploration alone. This includes investigating alternative forms of interaction between symbolic reasoning and learned policies, such as adaptive guidance during policy refinement, abstraction over skills or behaviors, and the reuse of symbolic knowledge to enhance generalisation across related tasks.

Another line of research I'm investigating involves the development and use of richer symbolic representations to capture higher-level structure in complex environments. Exploring temporal, relational, and hierarchical abstractions may enable agents to reason more effectively over long horizons, support transfer across tasks, and improve robustness to sparse or delayed rewards.

In addition, I am investigating the definition of quantitative explainability metrics to formally measure the level of adherence between symbolic policy representations and the underlying DRL policies. While symbolic abstractions are often assumed to improve interpretability, there is currently a lack of principled methods to evaluate how faithfully symbolic rules approximate the actual decision-making behaviour of neural policies. My goal is to develop metrics that capture aspects such as behavioral agreement, coverage, and consistency across states, enabling a systematic assessment of explanation fidelity. This would support both the comparison of different neuro-symbolic methods and the principled design of symbolic models that provide faithful and reliable explanations of learned policies.

Finally, I'm considering how to exploit symbolic learning in planning-oriented applications more broadly. I believe that learning symbolic abstractions, constraints, or heuristics from experience has the potential to reduce reliance on handcrafted domain knowledge while maintaining interpretability, enabling more scalable and adaptive planning and decision-making systems.

REFERENCES

[1] Panpan Cai and David Hsu. 2022. Closing the Planning–Learning Loop With Application to Autonomous Driving. *IEEE Transactions on Robotics* 39, 2 (2022), 998–1011.

[2] Ching-An Cheng, Andrey Kolobov, and Adith Swaminathan. 2021. Heuristic-guided reinforcement learning. *Advances in Neural Information Processing Systems* 34 (2021), 13550–13563.

[3] Gabriel Dulac-Arnold, Nir Levine, Daniel J Mankowitz, Jerry Li, Cosmin Paduraru, Sven Gowal, and Todd Hester. 2021. Challenges of real-world reinforcement learning: definitions, benchmarks and analysis. *Machine Learning* 110, 9 (2021), 2419–2468.

[4] Daniel Furelos-Blanco, Mark Law, Anders Jonsson, Krysia Broda, and Alessandra Russo. 2021. Induction and exploitation of subgoal automata for reinforcement learning. *Journal of Artificial Intelligence Research* 70 (2021), 1031–1116.

[5] Julian Ibarz, Jie Tan, Chelsea Finn, Mrinal Kalakrishnan, Peter Pastor, and Sergey Levine. 2021. How to train your robot with deep reinforcement learning: lessons we have learned. *The International Journal of Robotics Research* 40, 4-5 (2021), 698–721.

[6] Robert Kowalski and Marek Sergot. 1989. A logic-based calculus of events. In *Foundations of Knowledge Base Management*. Springer, 23–55.

[7] Mark Law, Alessandra Russo, Elisa Bertino, Krysia Broda, and Jorge Lobo. 2020. FastLAS: Scalable Inductive Logic Programming Incorporating Domain-Specific Optimisation Criteria. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 03 (2020), 2877–2885. https://doi.org/10.1609/aaai.v34i03.5678

[8] Yiyuan Lee, Panpan Cai, and David Hsu. 2020. MAGIC: Learning macro-actions for online POMDP planning. *arXiv preprint arXiv:2011.03813* (2020).

[9] M. Leonetti, L. Iocchi, and P. Stone. 2016. A synthesis of automated planning and reinforcement learning for efficient, robust decision-making. *Artificial Intelligence* 241 (2016), 103–130.

[10] Daniele Meli, Alberto Castellini, and Alessandro Farinelli. 2024. Learning logic specifications for policy guidance in pomdps: an inductive logic programming approach. *Journal of Artificial Intelligence Research* 79 (2024), 725–776.

[11] Daniele Meli, Hirenkumar Nakawala, and Paolo Fiorini. 2023. Logic programming for deliberative robotic task planning. *Artificial Intelligence Review* 56, 9 (2023), 9011–9049.

[12] Stephen Muggleton. 1991. Inductive logic programming. *New generation computing* 8, 4 (1991), 295–318.

[13] Manel Rodriguez-Soto, Roxana Rădulescu, Juan A Rodriguez-Aguilar, Maite Lopez-Sanchez, and Ann Nowé. 2023. Multi-objective reinforcement learning for guaranteeing alignment with multiple values. In *2023 Adaptive and Learning Agents Workshop at AAMAS*.

[14] David Silver and Joel Veness. 2010. Monte-Carlo planning in large POMDPs. *Advances in neural information processing systems* 23 (2010).

[15] M. Sridharan, M. Gelfond, S. Zhang, and J. Wyatt. 2019. REBA: A refinement-based architecture for knowledge representation and reasoning in robotics. *Journal of Artificial Intelligence Research* 65 (2019), 87–180.

[16] Jayakumar Subramanian, Amit Sinha, Raihan Seraj, and Aditya Mahajan. 2022. Approximate information state for approximate planning and reinforcement learning in partially observed systems. *The Journal of Machine Learning Research* 23, 1 (2022), 483–565.

[17] R. Toro Icarte, T. Q. Klassen, R. A. Valenzano, and S. A. McIlraith. 2018. Using Reward Machines for High-Level Task Specification and Decomposition in Reinforcement Learning. In *Proceedings of the International Conference on Machine Learning (ICML)*. 2112–2121.

[18] C. Veronese, D. Meli, F. Bistaffa, M. Rodríguez-Soto, A. Farinelli, and J. Rodríguez-Aguilar. 2023. Inductive Logic Programming For Transparent Alignment With Multiple Moral Values. In *BEWARE-23: 2nd International Workshop on Emerging Ethical Aspects of AI @ AIxIA*.

[19] Celeste Veronese, Daniele Meli, and Alessandro Farinelli. 2025. Learning Symbolic Persistent Macro-Actions for POMDP Solving Over Time. In *Proceedings of The 19th International Conference on Neurosymbolic Learning and Reasoning (Proceedings of Machine Learning Research, Vol. 284)*, Leilani H. Gilpin, Eleonora Giunchiglia, Pascal Hitzler, and Emile van Krieken (Eds.). PMLR, 1026–1040. https://proceedings.mlr.press/v284/veronese25a.html

[20] Celeste Veronese, Daniele Meli, and Alessandro Farinelli. 2025. Online Inductive Learning from Answer Sets for Efficient Reinforcement Learning Exploration. In *Hybrid Models for Coupling Deductive and Inductive Reasoning*, Pierangela Bruno, Francesco Calimeri, Francesco Cauteruccio, and Giorgio Terracina (Eds.). Springer Nature Switzerland, Cham, 93–106.

[21] Celeste Veronese, Daniele Meli, and Alessandro Farinelli. 2026. Sample-Efficient Neurosymbolic Deep Reinforcement Learning. *arXiv preprint arXiv:2601.02850* (2026).

[22] George A Vouros. 2022. Explainable deep reinforcement learning: state of the art and challenges. *Comput. Surveys* 55, 5 (2022), 1–39.

[23] Nan Ye, Adhiraj Somani, David Hsu, and Wee Sun Lee. 2017. DESPOT: Online POMDP planning with regularization. *Journal of Artificial Intelligence Research* 58 (2017), 231–266.

[24] Maddalena Zuccotto, Alberto Castellini, Davide La Torre, Lapo Mola, and Alessandro Farinelli. 2024. Reinforcement learning applications in environmental sustainability: a review. *Artificial Intelligence Review* 57, 4 (2024), 88.