



# Προτεινόμενα Θέματα Διπλωματικής Εργασίας

Από

Νίκο Κατζούρη

## 1. Towards Neuro-Symbolic World Model Learning

World models are compact representations of how an environment evolves over time, so that an intelligent system can anticipate future states, support counterfactual reasoning, and improve decision making. They are increasingly central in agentic AI, where situational awareness depends on modeling temporal dynamics from rich sensory input. Current state-of-the-art world modeling remains largely neural-based and often suffers from limited interpretability, weak use of prior knowledge, and difficulty in capturing high-level structured regularities.

Neuro-symbolic AI combines neural-based learning with symbolic structure and knowledge reasoning and is, therefore, a promising direction for overcoming some of these limitations. The aim of this thesis is to study modern world-modeling techniques, such as autoregressive predictive models, contrastive representation learning, and diffusion-based approaches, and explore how they can be combined with neuro-symbolic methods for richer, more interpretable, and more robust temporal modeling. The work will focus on multimodal settings, including video and multivariate time series.

### References:

- [1] Ha, D. and Schmidhuber, J., *World Models*, arXiv preprint arXiv:1803.10122, 2018
- [2] Cano, L.H. et al., *Neurosymbolic World Models for Sequential Decision Making*, ICML 2025
- [3] van den Oord, A., Li, Y. and Vinyals, O., *Representation Learning with Contrastive Predictive Coding*, arXiv preprint arXiv:1807.03748, 2018.
- [4] Marra, G. et al., *From Statistical Relational to Neurosymbolic Artificial Intelligence: A Survey*, Artificial Intelligence, 2024
- [5] De Smet, L., & De Raedt, L., *Defining Neurosymbolic AI*. arXiv preprint arXiv:2507.11127, 2025

## 2. Learning Symbolic World Models for Planning

Symbolic world models - typically expressed as STRIPS/PDDL action theories - provide a compact, interpretable representation of an environment's dynamics and enable sound long-horizon planning with off-the-shelf classical planners. A line of recent work has shown that such models can be learned from a handful of demonstrations by inventing predicates and operators directly from raw sensory input, with vision-language models (VLMs) acting as a perceptual front-end that proposes and grounds candidate predicates from images. The state of the art (e.g. pix2pred, VisualPredicator) couples this VLM front-end with hill-climbing procedures that select predicates and learn PDDL operators, but these procedures are slow, sensitive to hyperparameters, restricted to a propositional STRIPS-like fragment, and brittle to perception noise.

The purpose of this thesis is to explore more efficient and robust symbolic world model learning techniques for planning. Concretely, the student will keep the VLM-based predicate proposal and labelling layer of a pix2pred-like system, but reformulate operator learning as an Answer Set Programming (ASP)-based inductive task. The expected benefits include (i) principled noise tolerance through learners' built-in penalties, replacing ad hoc thresholds; (ii) a richer hypothesis language with negation, defaults, integrity constraints and choice rules; (iii) a unified, ASP-based framework for hypothesis induction and plan construction via Answer Set Planning; (iv) automatic predicate subselection through Minimum Description Length-style optimisation; and (v) the possibility of abductive planning at deployment time to recover from VLM mislabelling - the dominant failure mode reported in the literature. Evaluation will be on standard simulated benchmarks and against pix2pred-style baselines, measuring success rate, planning time, model compactness, and robustness to perception noise.

### References:

- [1] Athalye, A. et al., From Pixels to Predicates: Learning Symbolic World Models via Pretrained Vision-Language Models, IEEE Robotics and Automation Letters, 2026.
- [2] Silver, T. et al., Predicate Invention for Bilevel Planning, AAAI 2023.
- [3] Tran, S. C., et al. Answer Set Planning: a Survey. Theory and Practice of Logic Programming, 2023
- [4] Konidaris, G., Kaelbling, L.P. & Lozano-Pérez, T., From Skills to Symbols: Learning Symbolic Representations for Abstract High-Level Planning, Journal of Artificial Intelligence Research, 2018.

## 3. Supervision Augmentation Techniques in Weakly Supervised Neuro-Symbolic Training

Techniques for weakly supervised neuro-symbolic (NeSy) learning allow us to train hybrid neural-symbolic models with supervision that is limited, indirect, or available only at a coarse level. This is typically achieved by using the knowledge in the symbolic component of a NeSy system as a “teacher”, that produces weak labels for the neural component during training, so as to align the neural predictions with what the knowledge expects to reason correctly. This setting is especially relevant in

temporal domains, where labels often concern whole sequences, complex events, or trajectories, rather than the latent states, symbols, or relations that drive the neural predictions. Such settings arise naturally in event-based modeling from video, multivariate time series and other multimodal streams.

Although useful, the indirect training paradigm is susceptible to so-called “*reasoning shortcuts*”, where the neural component learns misaligned symbolic mappings with unintended semantics, which suffice for high predictive performance in the training domain, but hurt cross-domain generalization. Moreover, indirect NeSy training is insufficient in *joint training settings*, where the neural and the symbolic components are trained simultaneously. The aim of this thesis is to explore methods for augmenting the supervision signal with minimal manual annotation. In particular, the thesis will study techniques such as Active Learning and Data Programming, and assess how they can support weakly supervised neuro-symbolic training in temporal and multimodal settings. The goal is to improve sample efficiency, reduce annotation cost, and preserve interpretability while maintaining strong predictive performance.

### **References:**

- [1] Roy, N. and McCallum, A., 2001. *Toward Optimal Active Learning through Monte Carlo Estimation of Error Reduction*, ICML 2001
- [2] Ratner, A. et al.. *Data Programming: Creating Large Training Sets, Quickly*, NeurIPS, 2016.
- [3] Wang, Y. et al., 2023. *Rapid Image Labeling via Neuro-Symbolic Learning*, KDD 2023.
- [4] Barnaby, C. et al., 2025. *Active Learning for Neurosymbolic Program Synthesis*, Proceedings of the ACM Conf. on Programming Languages, 2025
- [5] Marconato, E. et al., *Not All Neuro-Symbolic Concepts are Created Equal: Analysis and Mitigation of Reasoning Shortcuts*, NeurIPS 2023

## **4. Neuro-Symbolic Complex Event Forecasting**

Event forecasting concerns the prediction of complex events before they occur, based on partial observations of an evolving process. It is important in domains such as activity recognition, robotics, maritime monitoring, critical infrastructures, finance, and intelligent decision support, where early warnings and situational awareness are crucial. Existing forecasting methods are often either predominantly statistical, offering limited interpretability and weak integration of domain knowledge, or predominantly symbolic, requiring manually specified abstractions and rules and struggling with noisy, high-dimensional inputs.

The aim of this thesis is to study neuro-symbolic methods for event forecasting from temporal and multimodal data streams, such as video, sensor data, trajectories, or multivariate time series. The thesis will investigate how neural models can extract meaningful representations from low-level observations, and how symbolic temporal models - such as automata-based event patterns, either hand-crafted, or possibly learned - can be used to reason over these representations and forecast future events. Emphasis will be placed on predictive accuracy, earliness of forecasts, interpretability, robustness, and the ability to incorporate prior knowledge.

## References:

- Alevizos, E., Artikis, A. and Paliouras, G., 2022. Complex Event Forecasting with Prediction Suffix Trees.
- Giatrakos, N., Alevizos, E., Artikis, A., Deligiannakis, A. and Garofalakis, M., 2020. Complex Event Recognition in the Big Data Era: A Survey.
- Marra, G., Dumančić, S., Manhaeve, R. and De Raedt, L., 2024. From Statistical Relational to Neurosymbolic Artificial Intelligence: A Survey.
- Katzouris, N. et al., 2023. Answer Set Automata: A Learnable Pattern Specification Framework for Complex Event Recognition.

## 5. Learning Interpretable Temporal Models from Video and Time Series

Interpretable temporal models are important in applications where one must not only predict or detect temporal phenomena, but also explain how they arise over time. Symbolic and rule-based temporal models are attractive in this respect because they can provide compact, human-readable descriptions of temporal structure while remaining suitable for analysis and debugging. Recent work has also shown that differentiable methods can help learn thresholds, predicates, and rules directly from data, reducing the need for manual discretization and hand-crafted symbolic abstractions.

The aim of this thesis is to investigate methods for learning interpretable temporal models from multimodal sequential data, with emphasis on video and multivariate time series. The focus will be on models that combine learned low-level representations with symbolic temporal structure, so as to obtain predictors that are both effective and understandable. Particular emphasis will be placed on rule- and automata-based representations that support transparent reasoning over temporal patterns.

## References:

- [1] Xu, S., et. al., *Neural Rule Lists: Learning Discretizations, Rules, and Order in One Go*, NeurIPS 2025
- [2] Katzouris, N. & Paliouras G., *Answer Set Automata: A Learnable Pattern Specification Framework for Complex Event Recognition*, ECAI 2023.
- [3] Feldstein, J et. al., *Efficiently Learning Probabilistic Logical Models by Cheaply Ranking Mined Rules*, arXiv preprint arXiv:2409.16238, 2024
- [4] Umili E. & Capobianco R., *Automata Learning through Neural Probabilistic Relaxations*, ECAI 2024

## 6. Comparing Differentiable and Purely Symbolic Techniques for Symbolic Automata Learning

Symbolic automata extend classical automata by allowing transitions to be labeled with predicates over rich input domains rather than simple symbols from a finite alphabet. This makes them particularly useful in settings with structured or infinite alphabets, and has made them crucial in areas such as systems' modelling and formal verification, model checking, and event recognition/forecasting. In complex event

processing in particular, symbolic automata provide an expressive and efficient representation of temporal patterns over complex events with attributes. Learning such automata from data is an important but difficult problem, since it involves discovering both structure and logical conditions under combinatorial constraints. The aim of this thesis is to compare purely symbolic techniques - especially methods based on Inductive Logic Programming and Answer Set Programming - with differentiable symbolic-automata learning methods and hybrid approaches. The comparison will consider efficiency, model compression, interpretability, and predictive performance.

### **References:**

- [1] D'Antoni, L. and Veanes, M., *The Power of Symbolic Automata and Transducers*, Int. Conf. on Computer Aided Verification, 2017.
- [2] Katzouris, N. & Paliouras G., *Answer Set Automata: A Learnable Pattern Specification Framework for Complex Event Recognition*, ECAI 2023.
- [3] Umili E. & Capobianco R., *Automata Learning through Neural Probabilistic Relaxations*, ECAI 2024.
- [4] Fisman, D., Frenkel, H., & Zilles, S., *Inferring symbolic automata*, Logical Methods in Computer Science, 2023
- [5] Furelos-Blanco, D. et al., *Induction and Exploitation of Subgoal Automata for Reinforcement Learning*, Journal of Artificial Intelligence Research, 2021.