Application of Neurosymbolic AI to Sequential Decision Making

UNIVERSIDAD DE GRANADA

Carlos Núñez-Molina[1] – PhD supervisors: Juan Fernández-Olivares[2], Pablo Mesejo[3]

[1]ccaarlos@correo.ugr.es, [2]faro@decsai.ugr.es, [3]pmesejo@decsai.ugr.es

Introduction

Sequential Decision Making (SDM) is the problem of solving **Sequential Decision Processes (SDP)**. In an SDP, an agent situated in an environment must make a series of decisions to complete a task or achieve a goal. Two main paradigms for SDM: **Automated Planning (AP)** and **Reinforcement Learning (RL)**.

	Automated Planning	Reinforcement Learning
Method for obtaining the solution	Plan over a model of the environment	Learn a policy directly from data
Knowledge representation	Symbolic, in first-order logic	Subsymbolic, usually as the weights of a neural network

Many works have tried to **bridge the gap between AP and RL**, like model-based RL and methods for learning the structure of the SDP (e.g., planning domains). In recent years, **Neurosymbolic AI** has attracted great attention. These are hybrid models that **combine deep neural networks with symbolic representations.**

Main goal of this PhD: development of neurosymbolic models for both solving and learning the structure of SDPs. We propose **three lines of research**.

Goal Selection with Deep Q-Learning

Deep Q-Planning (DQP): neurosymbolic model that uses **RL (Deep Q-Learning)** to **learn to select goals**, which are then achieved with a **PDDL-based planner**. It outperforms standard Deep Q-Learning and drastically reduces planning times [1, 2].

1. Use **Deep Q-Learning** to predict the length of the total plan for each **goal**, i.e., the length of the plan that first achieves that **goal**, and then, the **final goal**.

2. Select the goal with the shortest predicted length. Then, use the **planner** to find a **plan to the selected goal**.

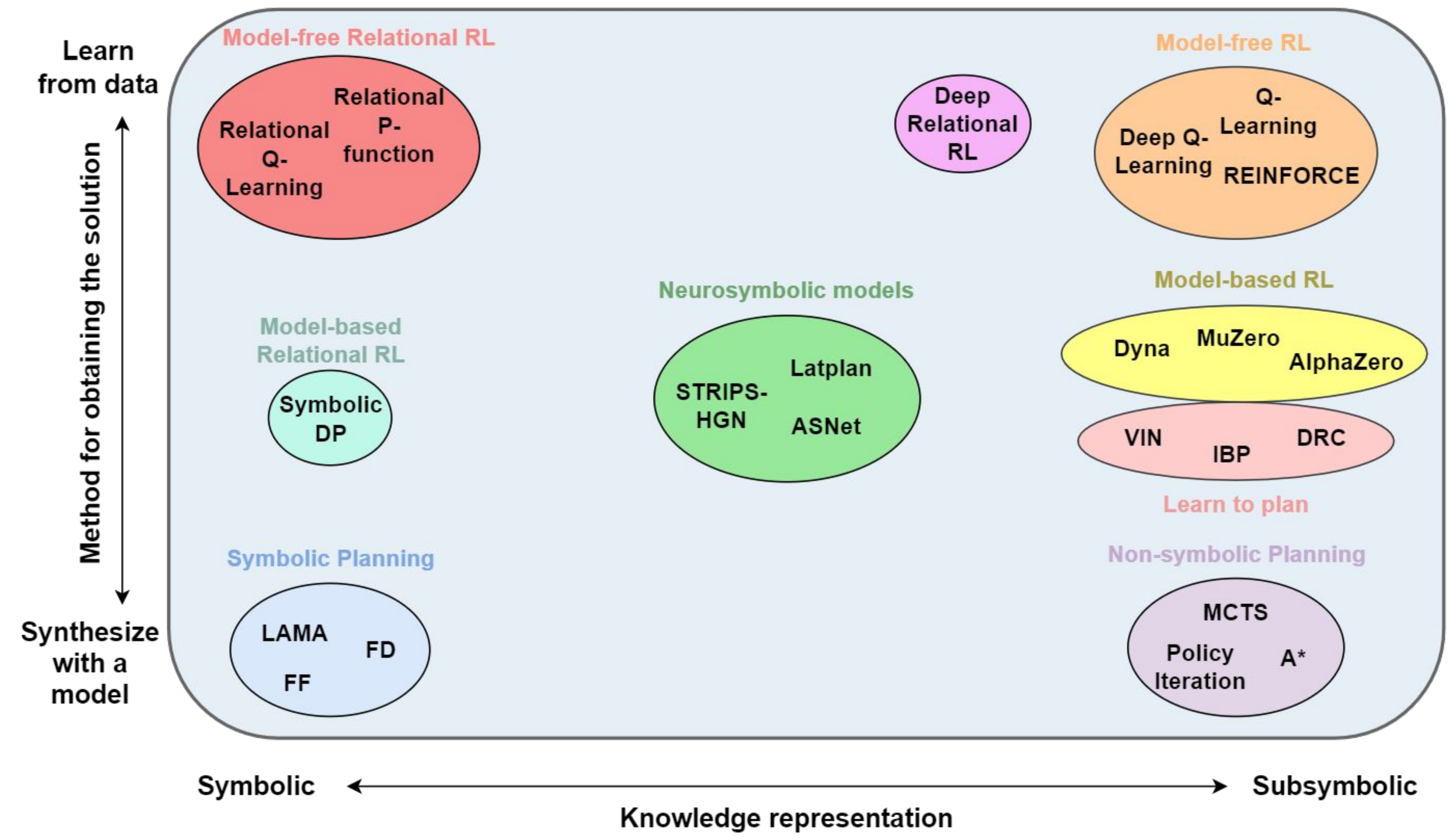
Next steps:

- Enhance DQP to **predict the uncertainty** of each goal and apply it to **stochastic environments**.
- Apply DQP to **manage the logistics** of a package delivery company.

Acknowledgements

This work is being partially funded by the Andalusian Regional Projects B-TIC-668-UGR20 and PYC20-RE-049UGR, and the Spanish National Project RTI2018-098460-B-I00 with FEDER funds.

Review of the State of the Art

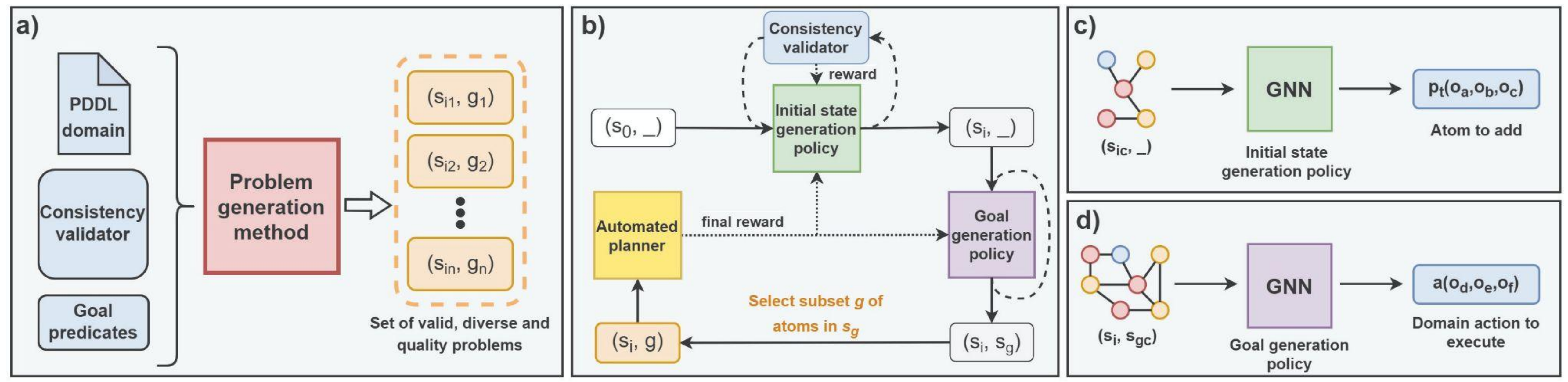


Review about symbolic, subsymbolic and hybrid methods for both solving and learning the structure of SDPs.

Neurosymbolic models pose a promising approach for the **integration of AP and RL**.

Automatic Planning Problem Generation

Automatic method for generating **valid, quality and diverse planning problems for any given domain**. Problem generation is formulated as a **SDP** and a **Graph Neural Network (GNN)** is trained with **RL** to generate problems with the desired qualities [3].



Next steps:

- **PDDL2HTN**: use our problem generation method to obtain the plan traces HTN domain learning techniques need, so that they don't need to be provided by experts.
- **Domain characterization**: given a planning domain, generate problems with our method. Then, apply unsupervised learning techniques (e.g.: clustering) to study their properties.

References

[1] Núñez-Molina, C., Vellido, I., Nikolov-Vasilev, V., Pérez, R., & Fdez-Olivares, J. (2021). A Proposal to Integrate Deep Q-Learning with Automated Planning to Improve the Performance of a Planning-Based Agent. In Conference of the Spanish Association for Artificial Intelligence (pp. 23-32). Springer, Cham.

[2] Núñez-Molina, C., Fernández-Olivares, J., & Pérez, R. (2022). Learning to select goals in Automated Planning with Deep-Q Learning. Expert Systems with Applications, 117265.

[3] Núñez-Molina, C., Mesejo, P., & Fernández-Olivares, J. (2022). A Proposal to Generate Planning Problems with Graph Neural Networks. In ICAPS 2022 Workshop on Bridging the Gap Between AI Planning and Reinforcement Learning. In Press.