Learning a Symbolic Planning Domain through the Interaction with Continuous Environments

Elena Umili*, Emanuele Antonioni*, Francesco Riccio, Roberto Capobianco, Daniele Nardi, Giuseppe De Giacomo

Department of Computer Control and Management Engineering Antonio Ruberti (DIAG) of Sapienza University of Rome, Italy.

E-mail {lastname}@diag.uniroma1.it

Introduction

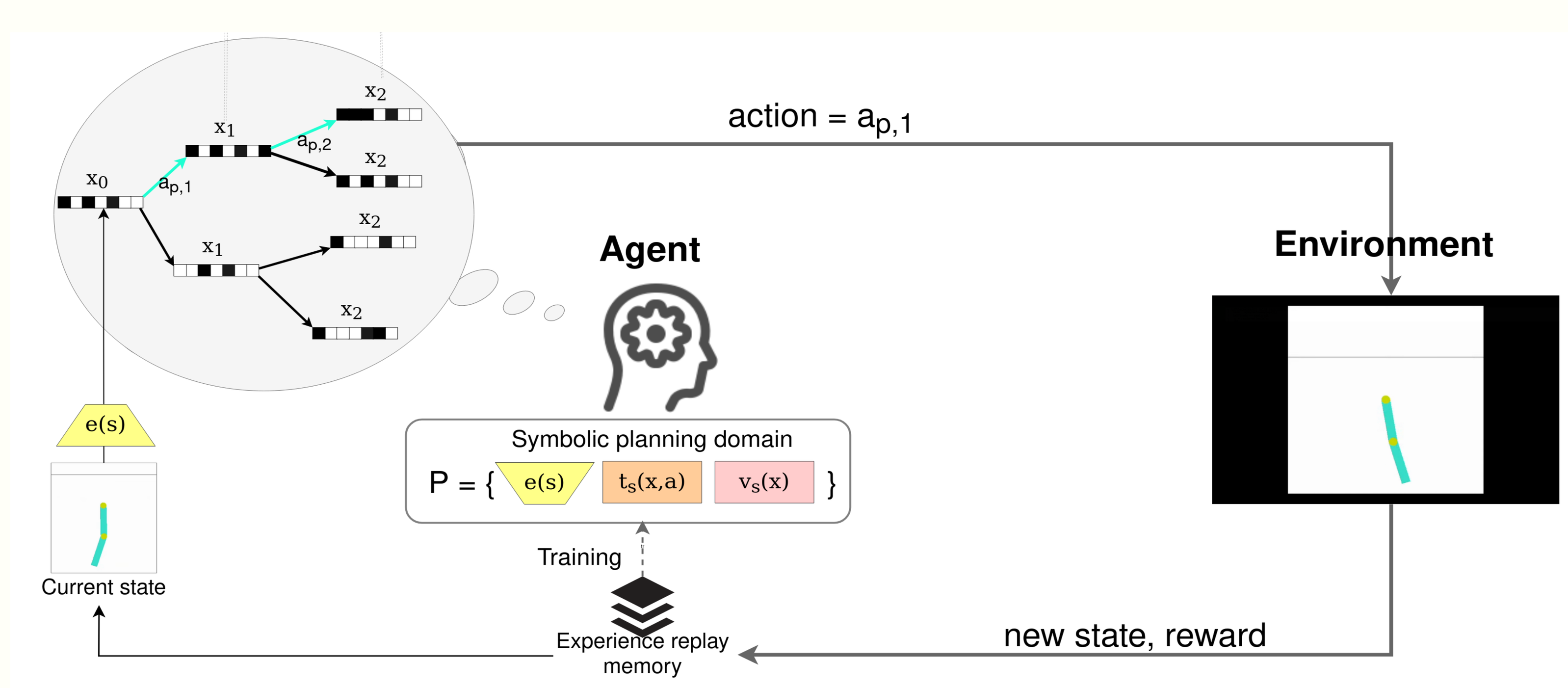
Motivation:

- Predicting environment's **future outcomes** is important to perform **dynamic tasks**.
- State-of-the-art planners can reason effectively **with symbolic representations of the environment**. However, when the environment is continuous and unstructured, **extracting an ad-hoc symbolic model** to perform planning may be **infeasible**.

Contribution:

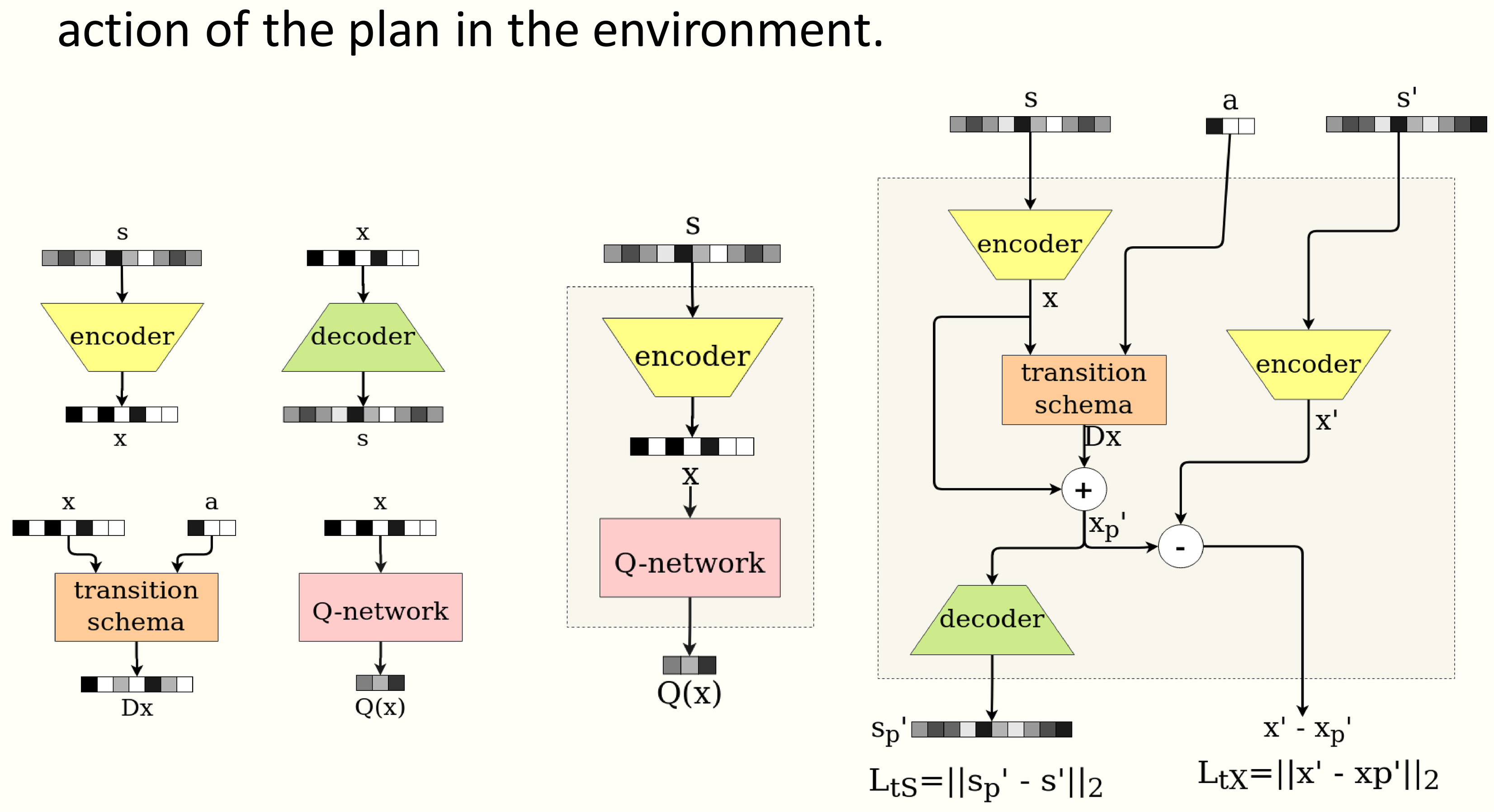
- Interaction with **symbolic representation learning,** and **symbolic online planning**.
- It autonomously learns a **symbolic planning model** composed of: (i) a symbol grounding model to switch from continuous to symbolic space and vice versa; (ii) a symbolic transition model; (iii) a value function for symbolic states.
- We exploit the symbolic model to **plan for the agent actions in the continuous state-space.**

Methodology



Grounding and planning:

- **Symbol grounding:** the encoder network maps the continuous state in a propositional symbolic representation. The latter is used by all the other system components.
- **Symbolic planning:** at each step the agent plans the next T steps in the symbolic space using the learned models and executes the first action of the plan in the environment.



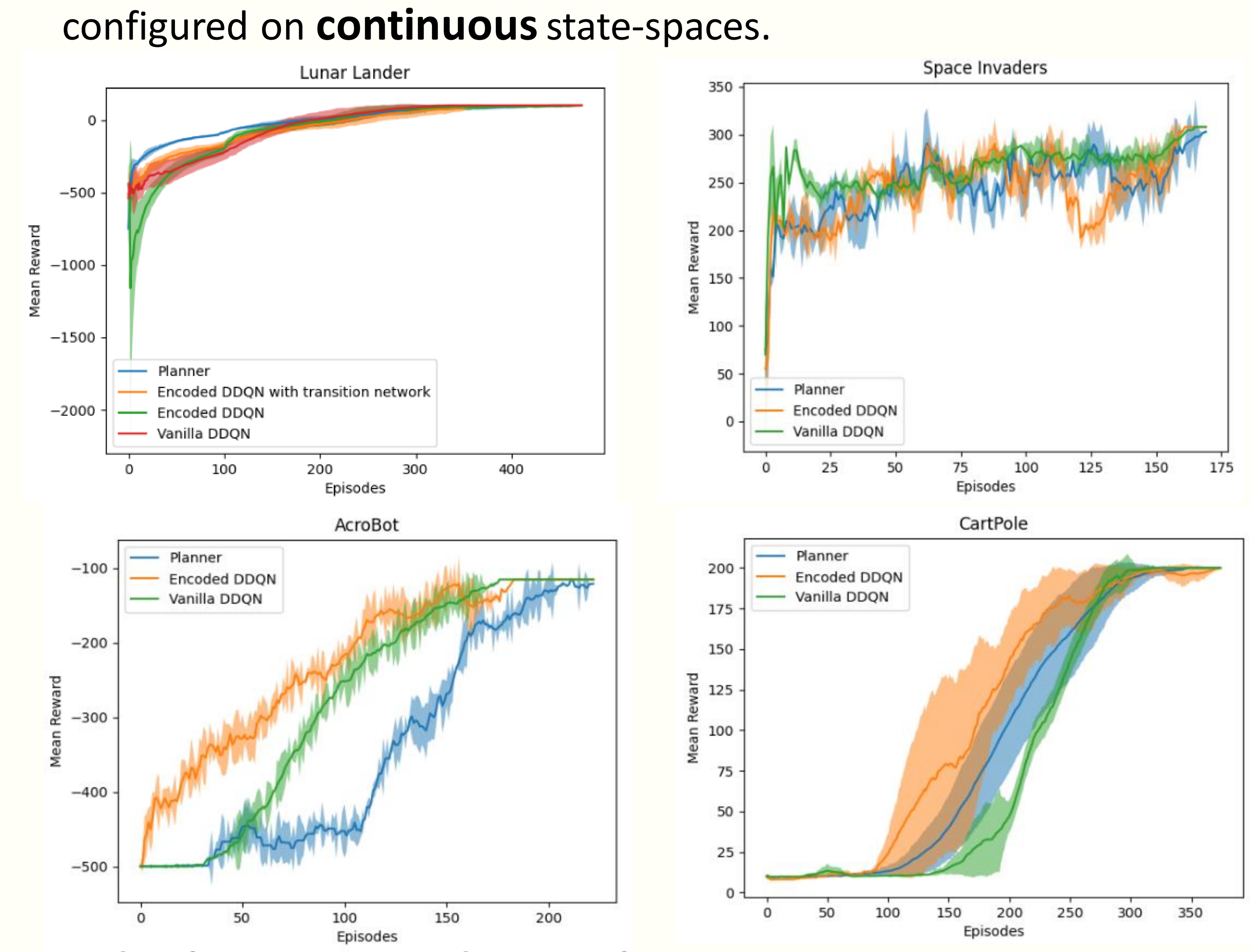
Learning the models:

- The four neural networks are **trained end-to-end** with samples (state, action, next state, rewards) observed in the environment.
- The symbolic **transition network** is trained minimizing the next state prediction error **in the continuous** and **in the symbolic state-space**.

Results

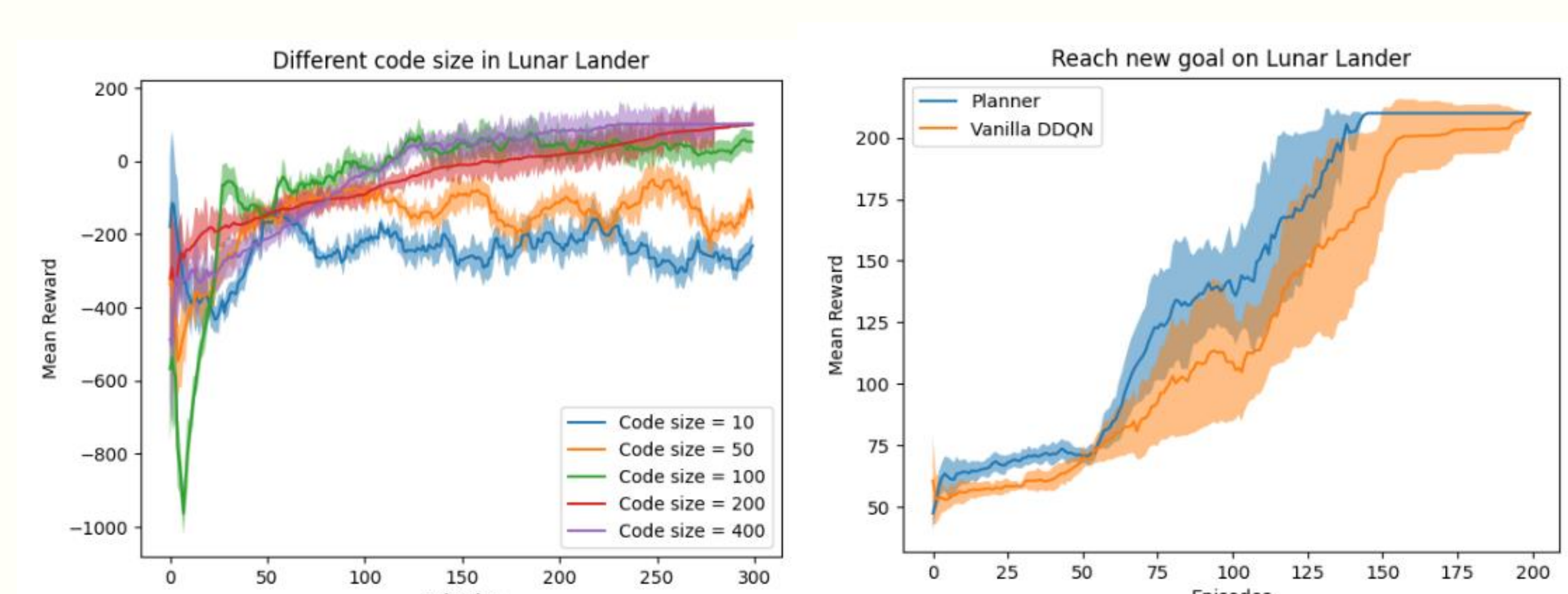
Training rewards:

- Our **discrete** system achieves **competitive results** in different OpenAI gym environments when **compared to Q-learning** configured on **continuous** state-spaces.



Code abstraction and generalization:

- Left) **convergence with different sizes of the symbols set**: the number of symbols is a crucial hyperparameter to make the system converge.
- Right) **Transfer learning**: we can use the distance in the symbolic space to drive the agent to different goals in the same environment.



Conclusions and future works

Conclusions:

- This work proposes an interactive learning algorithm inspired by both **planning** and **Reinforcement Learning**.
- It **automatically learns a symbolic planning domain** from a continuous-state MDP.
- We show that **reasoning in the symbolic space is enough** to effectively guide the agent's action **to achieve the task** in the continuous environment.

Future work:

- Focus on the **domain re-usability** to eliminate the second training necessary for achieving a new goal.
- Learn an **uncertainty model** for the transitions, in order to use the symbolic model to **plan towards little explored states.**