The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)

Artificial Intelligence and Machine Learning for Autonomous Agents that Learn to Plan and Operate in Unpredictable Dynamic Environments

Leonardo Lamanna[1, 2]
[1] Fondazione Bruno Kessler (FBK), Povo, Trento, Italy
[2] Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Brescia, Italy
llamanna@fbk.eu, l.lamanna@unibs.it

Abstract

My research activity focuses on the integration of acting, learning and planning. The main objective is to build a system that is capable to learn how to plan and act in an unknown, dynamic and complex environment. The only knowledge the agent has about the environment is provided by a set of sensor observations which returns continuous measures on the environment. On the learning side, I'm interested in developing algorithms that allow an artificial agent to learn an abstract model of the dynamics of the environment (either an explicit model like a deterministic finite state machine or a model described in a language to express planning domains). The type of abstract model is specified by means of discrete state variables rather than continuous variables representing agent observations. In addition to learning the abstract model, I'm interested in learning probabilistic (generative) models that connects the abstract model with the perceptions of the agents. On the acting and planning side, the artificial agent does not rely on a prior set of execution traces, it rather decides online how to act by means of state-of-art planners. With its own experience, it enriches the planner knowledge, as well as the learned model of the environment. On the learning part, the agent applies techniques for dynamic probabilistic clustering of perceptions in a set of abstract states, neural network for learning transition models, and inductive reasoning for learning action model descriptions. Notice that this is different from Reinforcement Learning where the focus is to learn a policy for achieving a goal, we are interested in learning an abstract model of the environment. We do not have a reward function that encodes the goal to be reached. Indeed in this work an agent does not necessarily need to reach a goal.

A large amount of work on learning planning domains focuses on learning action schema from data. Gregory and Cresswell (2016) and Cresswell, McCluskey, and West (2013) propose learning general action schema in a structured language starting from plans containing grounded application instances of actions. Mourão et al. (2012) learn action schemata from noisy and incomplete observations. Each observation is a sequence of alternating actions and set of fluent expressions. Zhuo and Yang (2014) learn an action schema on a target domain by transfer learning from a set of source domains and by observing partial plan traces. Aineto, Jiménez, and Onaindia (2018) propose a method for learning action models from observations of plan executions that compiles the learning task into a classical planning task. In all these approaches, learning is performed at the symbolic level, and mappings to perceptions in a continuous environment are not considered. This is also the case of the work by Bonet and Geffner (2019), which provides a framework for learning first-order symbolic representations from plain graphs. Indeed, plain graphs are state transition systems, and there is no mapping to perceptions in a continuous space, which is what sensors actually provide.

All the above mentioned works do not tackle the problem of finding an abstraction of the continuous environment (with continuous states) into a finite set of states. Building this abstraction (encoded in the perception function) is one of the key contributions of our work. There are however a set of approaches that learn a discrete planning domain from a continous environment. Causal InfoGAN learns discrete or continuous models from high dimensional sequential observations (Kurutach et al. 2018). This approach fixes a priori the size of the discrete domain model. Differently from our approach, their goal is to generate an execution trace in the high dimensional space. LatPlan takes in input pairs of high dimensional raw data (e.g., images) corresponding to transitions (Asai 2019). LatPlan is an offline approach, while our approach is online and works also in dynamic environments. Konidaris, Kaelbling, and Lozano-Pérez (2018) construct a STRIPS model by learning the boolean atoms of the preconditions and effects of actions. However, their basic assumption is that a continuous model of the world is available, and that it is possible to know a fixed mapping from the continuous model to the deterministic planning domain. We do not rely on such assumptions. Moreover, in our approach, the mapping through perception functions is learned dynamically. Finally, in the work by Konidaris, Kaelbling, and Lozano-Pérez (2018), the mapping is set-theoretic, while we allow for a probabilistic mapping through a probability density function. Most of the work on learning and planning in Partially Observable Markov Decision Processes (POMDP) – see, e.g., (Ross et al. 2011; Katt et al. 2017) – focuses on learning transitions and policies by assuming a fixed and given set of states and a given reward function. Some of them drop the assumption of a bounded state space, see, e.g., (Doshi-Velez 2009). However, none of these works uses an intensional representation to guide the search for learning an extensional representation of the planning domain.

The proposed research plan consists in:

- Learn an extensional representation of a discrete deterministic planning domain from observations in a continuous space navigated by the agent actions. This is achieved through the use of a perception function providing the likelihood of a real-value observation being in a given state of the planning domain after executing an action. The agent learns an extensional representation of the domain (the set of states, the transitions from states to states caused by actions) and the perception function on-line, while it acts. I have integrated and tested the framework presented in (Serafini and Traverso 2019) with three planning domains (Logistics, Grid and Rovers).
- Guide agent exploration of the environment and provide a practical approach that can scale up to large state spaces through the exploitation of a draft intensional (PDDL-based) model of the planning domain. The intensional model is draft in the sense that the action preconditions can be incomplete. When an action execution fails, the intensional model is updated in order to prevent the action from being performed again in the future. I have integrated the framework proposed in (Serafini and Traverso 2019) with the use of a draft intensional model and tested the method to update the intensional model suggested by A.E.Gerevini.
- Learn an intensional representation of a discrete deterministic planning domain, i.e., learn an action model. Particularly, the focus is on learning the action preconditions by online execution. Together with L.Serafini, I developed a method to learn action preconditions by online execution and tested it in three planning domains.

We wrote an article summarizing the work done so far, it has been presented in the ICAPS-KEPS 2020 workshop.

The future works of the proposed research plan are:

- Perception clustering in abstract states: the objective is to learn the perception function which can effectively abstract away irrelevant details of the environment and clustering continuous perceptions into a finite set of significantly different states. Possible approaches are: incremental clustering, Gaussian clustering and autoencoders.
- Learning abstract state structure: the objective is to learn the right set of state variables and the set of values of their domains, that allow to represent the set of abstract states provided by the perception clustering.
- Action Model Learning: the objective is to learn a generative model which is capable to predict the perceptions after the execution of the action (without executing it) and use this model to learn an explicit representation of the abstract transition model. Some possible approaches are: inductive learning, abstraction.

This thesis will contribute to the integration of acting, learning and planning in dynamic and uncertain environments. Particularly, it will propose a method to map continuous perception variables into discrete state variables. This will enable the reasoning at an abstract level. An abstract and discrete representation of the agent state will enable the learning of an intensional representation of the environment. Finally the intensional representation will be used by the agent to accomplish its task.

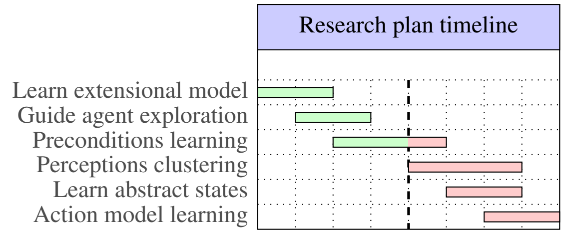

Figure 1: Brief timeline of the proposed research plan

References

Aineto, D.; Jiménez, S.; and Onaindia, E. 2018. Learning STRIPS action models with classical planning. In *ICAPS*.

Asai, M. 2019. Unsupervised Grounding of Plannable First-Order Logic Representation from Images. In *ICAPS*.

Bonet, B.; and Geffner, H. 2019. Learning First-Order Symbolic Planning Representations from Plain Graphs. *CoRR* abs/1909.05546.

Cresswell, S.; McCluskey, T. L.; and West, M. M. 2013. Acquiring planning domain models using *LOCM*. *Knowledge Eng. Review* 28(2): 195–213.

Doshi-Velez, F. 2009. The Infinite Partially Observable Markov Decision Process. In *NIPS*, 477–485.

Gregory, P.; and Cresswell, S. 2016. Domain Model Acquisition in the Presence of Static Relations in the LOP System. In *IJCAI*.

Katt, S.; Oliehoek, F. A.; Amato, C.; and s. 2017. Learning in POMDPs with Monte Carlo Tree Search. In *ICML*.

Konidaris, G.; Kaelbling, L. P.; and Lozano-Pérez, T. 2018. From Skills to Symbols: Learning Symbolic Representations for Abstract High-Level Planning. *J. Artif. Intell. Res.* 61: 215–289.

Kurutach, H.; Tamar, A.; Yang, G.; Russell, S.; and Abbeel, P. 2018. Learning Plannable Representations with Causal InfoGAN. In *NIPS*.

Mourão, K.; Zettlemoyer, L. S.; Petrick, R. P. A.; and Steedman, M. 2012. Learning STRIPS Operators from Noisy and Incomplete Observations. In *UAI*.

Ross, S.; Pineau, J.; Chaib-draa, B.; and Kreitmann, P. 2011. A Bayesian Approach for Learning and Planning in Partially Observable Markov Decision Processes. *Journal of Machine Learning Research* 12: 1729–1770.

Serafini, L.; and Traverso, P. 2019. Incremental Learning of Discrete Planning Domains from Continuous Perceptions. *CoRR* abs/1903.05937.

Zhuo, H. H.; and Yang, Q. 2014. Action-model acquisition for planning via transfer learning. *Artif. Intell.* 212: 80–103.