Data-driven Approaches for Formal Synthesis of Dynamical Systems

Doctoral Consortium

Milad Kazemi
Newcastle University
Newcastle upon Tyne, United Kingdom
m.kazemi2@newcastle.ac.uk

ABSTRACT

My research lies at the intersection of control theory, machine learning and formal methods. This paper presents part of the work developed so far within the scope of my PhD and suggests possible future research directions. Towards trustworthy computing, my research has focused on simplifying the designing pipeline of safe and reliable AI systems. I have worked on data-driven controller synthesis, i.e., the automated generation of control systems from a given high-level specification with theoretical guarantees of correctness. In this way, I analyze the satisfaction of properties in both episodic and continual settings. Moreover, in my research I provide correctness for satisfying specifications using different approaches including abstraction-based techniques, game-theoretic techniques, and model-free reinforcement learning.

KEYWORDS

Reinforcement Learning, Formal Synthesis, Average Reward, Omega-Regular, Linear Temporal Logic, Reward Machine

ACM Reference Format:
Milad Kazemi. 2022. Data-driven Approaches for Formal Synthesis of Dynamical Systems: Doctoral Consortium. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 2 pages.

1 INTRODUCTION

The traditional view in control theory can be summarized into how we connect sensing, actuation, and computation in a feedback loop to provide stability, performance, and robustness. However, in recent years, the control community has started to look at controlling complex systems such as autonomous vehicles, cells, and network systems. As a result, the question has been changed into how to think about dynamic, interconnection, and computing in a unified scalable framework. This is especially so for safety-critical systems.

Safety plays an important role in applications such as autonomous vehicles, air traffic control systems, chemical reactors, and industrial robots. However, the existence of model uncertainty, measurement noises, and disturbances poses a great challenge to design safe controllers. In the traditional view, tools such as robust control and adaptive control have been developed to deal with these issues, but these methods are expensive in terms of computation for large-scale systems or nonlinear systems. Such systems also must interact with complex environments that are ever-changing and difficult to model. These challenges motivates the use of data-driven decision making and control.

To overcome these difficulties, one approach is to use machine learning techniques, which can improve the system performance due to their data-driven nature and the ability to infer an unknown model from data. However, despite the ubiquitous use of these techniques in controller design, they hardly come with safety guarantees. This thesis is an attempt to develop and extend some approaches for formal synthesis when we do not have access to the model of the system. This thesis focuses on automated generation of controllers from a given high-level specification. This paradigm generalizes the task of designing reward functions to large class of tasks. It worth noting that there exists a large variety of synthesis algorithms and each one of them they have their own strength and weaknesses. In the following I explain in more detail the contribution of each work.

2 DATA-DRIVEN APPROACHES FOR FORMAL SYNTHESIS OF DYNAMICAL SYSTEMS

Formal abstraction-based synthesis schemes rely on a precise mathematical model of the system to build a finite abstract model, which is then used to design a controller. The abstraction-based schemes are not applicable when the dynamics of the system are unknown. In [1], I study formal synthesis of controllers for continuous-space systems with unknown dynamics to satisfy requirements expressed as linear temporal logic (LTL) formulas. I proposed a data-driven approach that computes the growth bound of the system using a finite number of trajectories. The growth bound together with the sampled trajectories are then used to construct the abstraction and synthesise a controller. The approach casts the computation of the growth bound as a robust convex optimisation program (RCP). Since the unknown dynamics appear in the optimisation, I formulate a scenario convex program (SCP) corresponding to the RCP using a finite number of sampled trajectories. I established a sample complexity result that gives a lower bound for the number of sampled trajectories to guarantee the correctness of the growth bound computed from the SCP with a given confidence. I also provided a sample complexity result for the satisfaction of the specification on the system in closed loop with the designed controller for a given confidence. I showed that our data-driven approach can be readily used as a model-free abstraction refinement scheme by modifying the formulation of the growth bound and providing similar sample complexity results.

Even though we can use SCP for control synthesis, in practice these approaches are not scalable and it needs large sample size for convergence guarantees. This motivates the use of model-free

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

approaches. In [2], I studied satisfaction of temporal properties on unknown stochastic processes that have continuous state spaces. We show how reinforcement learning (RL) can be applied for computing policies that are finite-memory and deterministic using only the paths of the stochastic process. We address properties expressed in linear temporal logic (LTL) and use their automaton representation to give a path-dependent reward function maximised via the RL algorithm. I developed the required assumptions and theories for the convergence of the learned policy to the optimal policy in the continuous state space. To improve the performance of the learning on the constructed sparse reward function, I proposed a sequential learning procedure based on a sequence of labelling functions obtained from the positive normal form of the LTL specification. I used this procedure to guide the RL algorithm towards a policy that converges to an optimal policy under suitable assumptions on the process.

As we know in most of real world application the environment is changing and we need to design controllers in presence of other independent agents. This leads to introducing compositional synthesis approaches to enhance the scalability. In [3], I introduced a novel reinforcement learning (RL) scheme to synthesize policies for *networks* of continuous-space stochastic control systems with unknown dynamics. The proposed *compositional* framework applies model-free *two-player* RL in an assume-guarantee fashion and *compositionally* compute strategies for continuous-space interconnected systems without explicitly constructing their finite-state abstractions. This approach gives a guaranteed lower bound for probability of property satisfaction by the interconnected system based on those of individual controllers over subsystems.

In [4], we shift our attention to average objectives. I restricted my attention to the omega-regular languages which correspond to *absolute liveness* specifications. These specifications cannot be invalidated by any finite prefix of agent behavior, in accordance with the spirit of a continuing problem. I proposed a translation from absolute liveness omega-regular languages to an average reward objective for RL. This reduction can be done on-the-fly, without full knowledge of the environment, thereby enabling the use of model-free RL algorithms. Additionally, I proposed a reward structure that enables RL without episodic resetting in communicating MDPs, unlike previous approaches. I demonstrated empirically with various benchmarks that this proposed method of using average reward RL for continuing tasks defined by omega-regular specifications is more effective than competing approaches that leverage discounted RL.

3 FUTURE WORK

I this paper, I presented the research I am doing as part of my study. I believe certification of engineered systems that have learning-enabled components is of importance and it is essential for safety critical applications. During my PhD, I have explored abstraction-based and model-free approaches, studied single agent and multi-agent settings, and worked on discounted and average reward perspectives. Some possible future research directions include developing assume-guarantee model-free reinforcement learning algorithms, applying the achieved results to privacy preserved reinforcement learning algorithms, and using logic for transfer learning approaches.

REFERENCES

[1] M. Kazemi, R. Majumdar, M. Salamati, S. Soudjani, and B. Wooding. Data-driven abstraction-based control synthesis. *IEEE Transactions on Automatic control*, 2022(Under submission).

[2] Milad Kazemi and Sadegh Soudjani. Formal policy synthesis for continuous-state systems via reinforcement learning. In *International Conference on Integrated Formal Methods*, pages 3–21. Springer, 2020.

[3] A. Lavaei, M. Perez, M. Kazemi, F. Somenzi, S. Soudjani, A. Trivedi, and M. Zamani. Model-free reinforcement learning for formal control of stochastic systems. *IEEE Transactions on Automatic control*, 2022(Under review).

[4] M. Kazemi, M. Perez, F. Somenzi, S. Soudjani, A. Trivedi, and A. Velasquez. Translating omega-regular specifications to average objectives for model-free reinforcement learning. In *In Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, 2022(Under publication).