Quantitative Model Checking for Analysis and Repair of Stochastic Control Policies

Armando Tacchella

Dipartimento. di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi (DIBRIS)
Università degli Studi di Genova, Via Opera Pia, 13 – 16145 Genova – Italy
Armando.Tacchella@unige.it

Abstract

Given a system operating in an environment modeled as a Markov decision process, it is possible to synthesize (optimal) stochastic control policies in a variety of ways. In particular, if the model of the environment is not available, a policy can be obtained using Reinforcement Learning (RL) techniques, wherein trial-and-error interaction between the system and the environment is leveraged. RL methods have shown robust and efficient learning on a variety of robot-control problems — see, e.g., [1]. However, while policies learned by reinforcement are satisfactory according to utility-based measures, they may fail to meet other requirements, e.g., safety. In this direction RL methods *per se* cannot provide adequate guarantees. In the words of [2]: "The asymptotic nature of guarantees about RL performances makes it difficult to bound the probability of damaging the controlled robot and/or the environment". How to guarantee that, given a control policy synthesized by RL, such policy will have a very low probability of yielding undesirable behaviors? Our answer leverages Probabilistic Model Checking techniques — see, e.g., [3] – by describing robot-environment interactions using Markov chains, and the related safety properties using probabilistic logic. Both the encoding of the interaction models and their verification can be fully automated, and only properties have to be manually specified. Our research goes beyond automating verification, to consider the problem of automating repair, i.e., if the policy is found unsatisfactory, how to fix it without manual intervention. In this talk, we detail how to automate the analysis of policies using probabilistic model checking techniques. Our methodology includes algorithms to repair control policies until they satisfy a set of safety requirements. We describe theoretical and empirical evidence about the effectiveness of our methodology. Alternative approaches and similar results available in the literature are also discussed.

References

1. J.A. Bagnell and S. Schaal. Special issue on Machine Learning in Robotics (Editorial). *The International Journal of Robotics Research*, 27(2):155–156, 2008.
2. J.H. Gillula and C.J. Tomlin. Guaranteed Safe Online Learning via Reachability: tracking a ground target using a quadrotor. In *ICRA*, pages 2723–2730, 2012.
3. M. Kwiatkowska, G. Norman, and D. Parker. Stochastic model checking. *Formal methods for performance evaluation*, pages 220–270, 2007.