KHIPU
LATIN AMERICAN MEETING IN ARTIFICIAL INTELLIGENCE
11-15 November, 2019 - Montevideo, Uruguay

A Hierarchical Approach to Hyper-parameter Optimization in Reinforcement Learning

Juan Cruz Barsce[1], Jorge Palombarini[1,2], Ernesto Martínez [3]

[1] Dpto. de Ingeniería en Sistemas de Información, Facultad Regional Villa María, UTN, Argentina.
[2] CIT Villa María - CONICET-UNVM, Argentina
[3] Instituto de Desarrollo y Diseño CONICET-UTN, Argentina

Abstract

Optimization of hyper-parameters in reinforcement learning (RL) algorithms is a key task, because they determine how the agent will learn its policy by interacting with its environment, and thus what data is gathered. In this work, an approach that uses Bayesian optimization to perform a two-step optimization is proposed: first, categorical RL structure hyper-parameters are taken as binary variables and optimized with an acquisition function tailored for such variables. Then, at a lower level of abstraction, solution-level hyper-parameters are optimized by resorting to the expected improvement acquisition function, while using the best categorical hyper-parameters found in the optimization at the upper-level of abstraction.

Reinfocement learning (RL) hyper-parameter optimization

Problem: **How to approach hyper-parameter optimization in RL** [3], considering that:

- RL differs from supervised learning in that data must be generated in order to speed up the learning curve, so hyper-params must be optimized while the agent is learning.
- There is not correct examples to learn from when learning a way of behaving (policy).
- In addition, $V(s \mid \Theta_1)$ cannot be compared with $V(s \mid \Theta_2)$ if $\Theta_1 \neq \Theta_2$ because they were generated under different hyper-parameters.

Having that in mind, **how to measure the optimization of RL hyper-params?**

- Proposed approach: Optimization of hyper-parameters measured as an objective function $f(\Theta)$, where Θ are the hyper-parameters. A *meta-episode* comprises several episodes with the same set of hyper-parameters Θ.

How to approach optimization of RL hyper-params?

- Assumption: RL hyper-parameters as a hierarchy, as shown in Fig. 1.

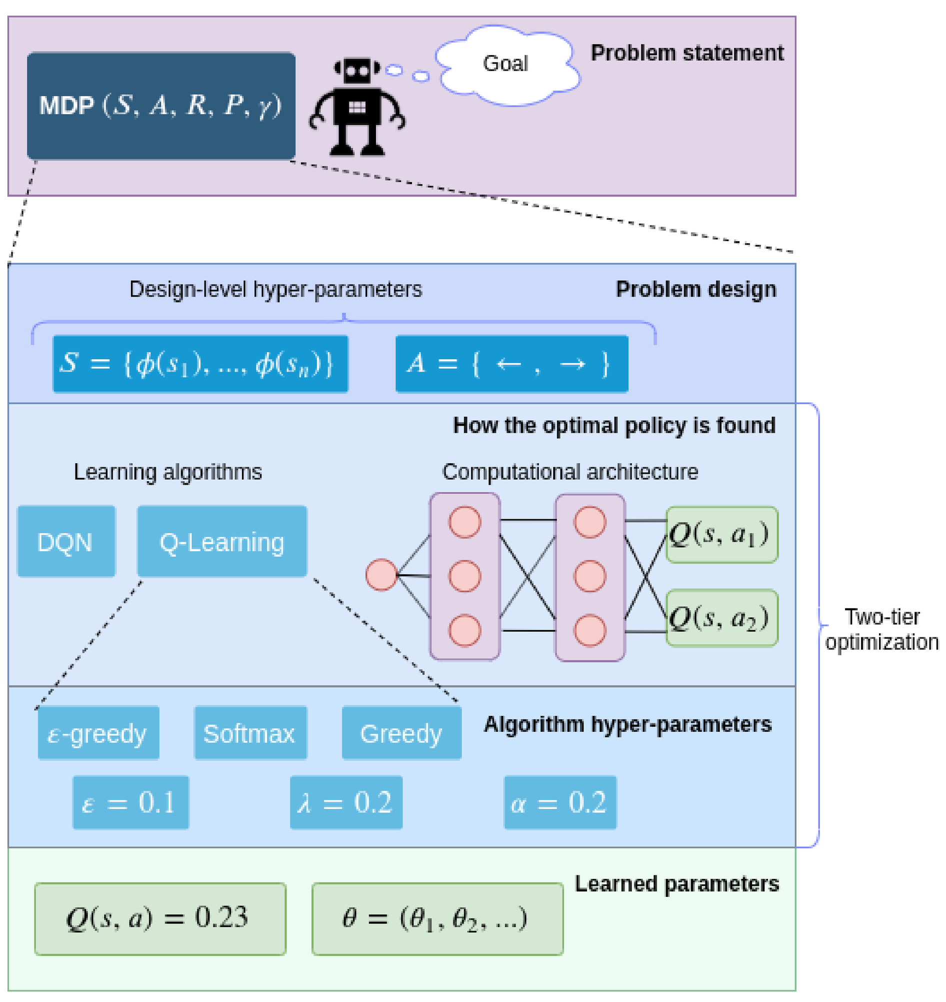

Fig. 1: *Proposed RL hiper-param hierarchy*

How to optimize hyper-parameters of different hierarchy levels?

Proposed approach

In the proposed algorithm (depicted in Fig. 2):

- Each time the agent is instantiated with a different set of $\Theta = \Theta_s \cup \Theta_a$, its learning is reset (e.g. the weights in a NN or tabular Q-values).
- Structural and categorical/discrete hyper-parameters are optimized with BOCS [1].
- Then, algorithm hyper-parameters are optimized with Bayesian optimization [2] and expected improvement acquisition function.
- A repository of experience is shared among different meta-episodes, and is sampled to perform experience replay.

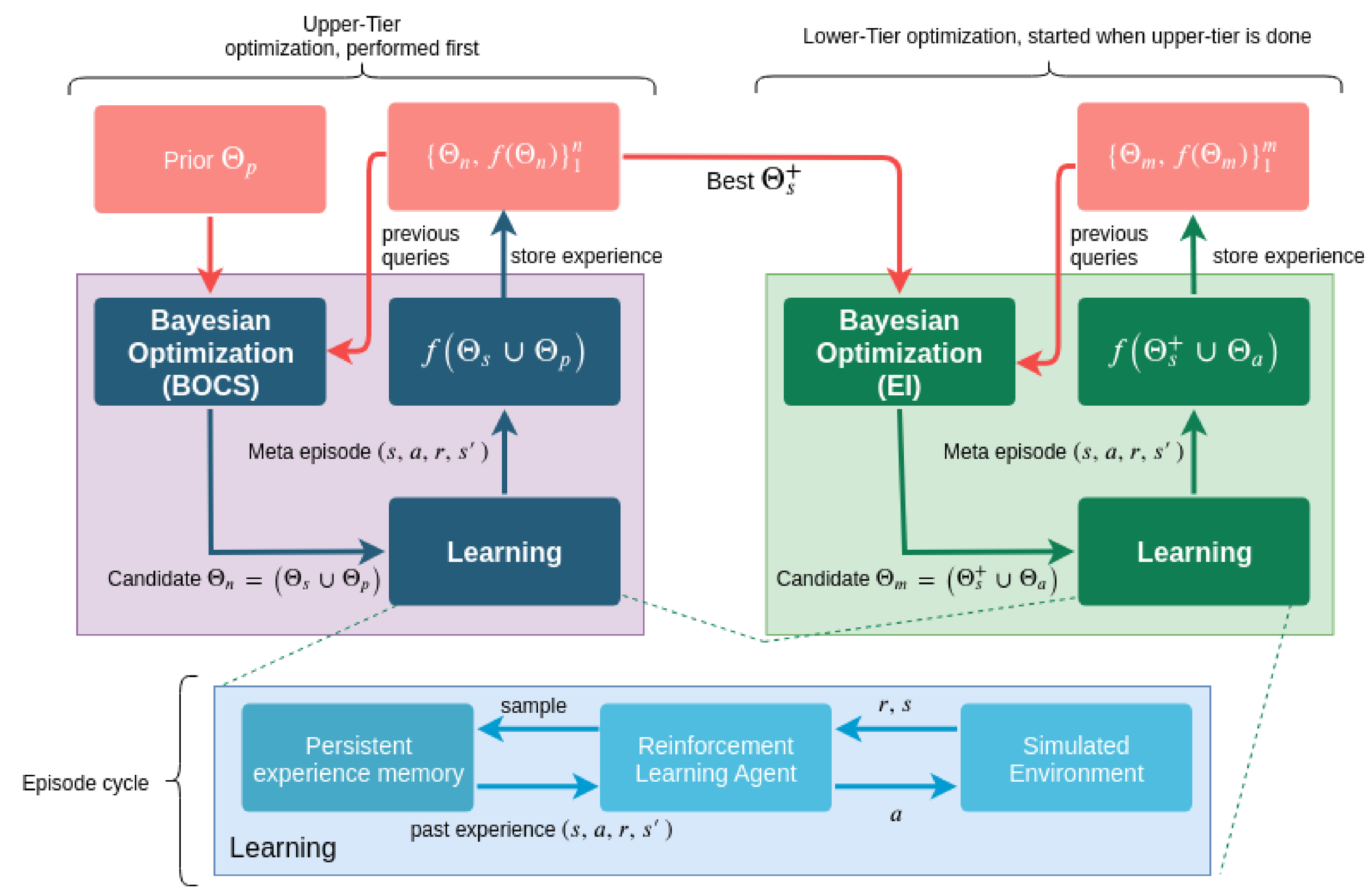

Fig. 2: *Two-tier RL hyper-parameter optimization approach.*

Results

Initial results with classic control problem Cartpole can be seen in Fig. 3, where the thick lines and their nearby curves correspond to the average and the 95% confidence interval for ten simulations with different random seeds. The green curve represents the proposed optimizer, whereas the blue curve uses Bayesian optimization without assuming a hierarchical relationship, and the orange curve represents a random search optimizer.

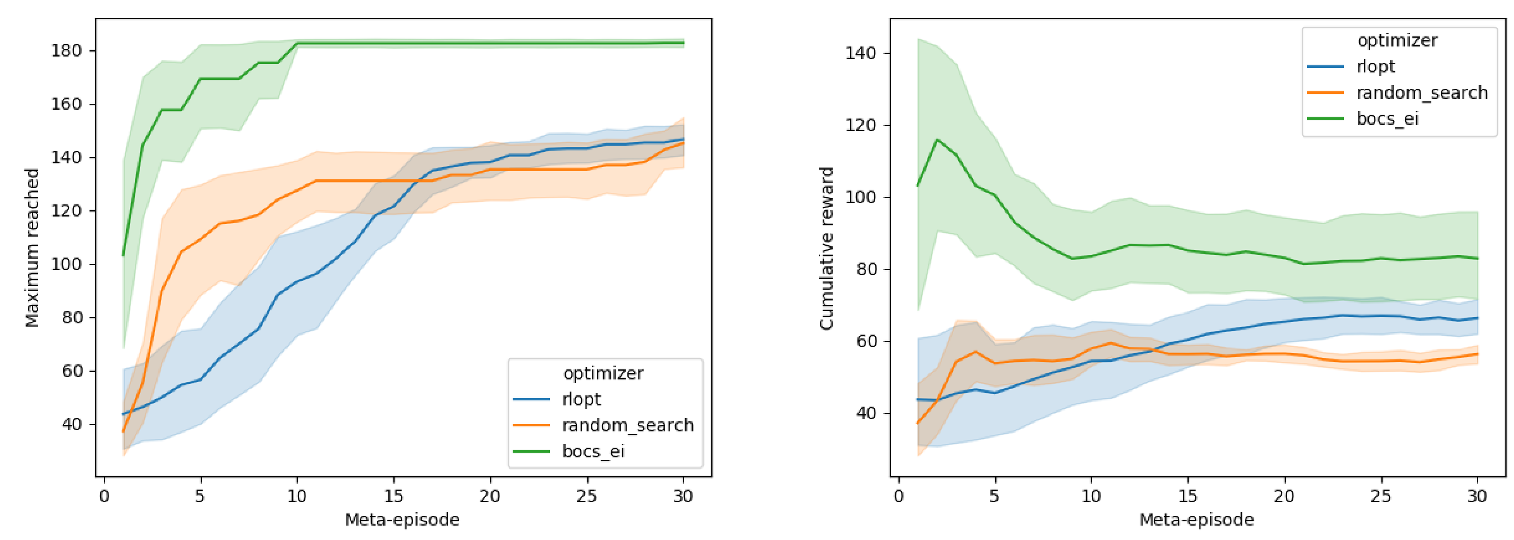

Fig. 3: *Avg. maximum reached by each optimizer (left), and cumulative reward per meta-episode (right) per each meta-episode*

Conclusion

An approach that involved the optimization of both categorical and real-valued RL hyper-parameters, assuming a hierarchical relationship between them was presented. Current research efforts are focused on including the extension of the concept of a hierarchical relationship among many hyper-parameters, the optimization of complex computational structures such as deep neural networks, among others.

References

[1] Ricardo Baptista and Matthias Poloczek. Bayesian Optimization of Combinatorial Structures. *arXiv:1806.08838 [cs, math, stat]*, June 2018.

[2] B. Shahriari, K. Swersky, Ziyu Wang, R.P. Adams, and N. de Freitas. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175, January 2016.

[3] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, Mass, 2nd edition, 2018.