Scaling-Up Reinforcement Learning (SURL)

Trends in Reinforcement Learning Research

Ruben Glatt[1], Felipe Leno da Silva[2], Denis Steckelmacher[3], and Patrick MacAlpine[4]

[1] Lawrence Livermore National Laboratory, USA
glatt1@llnl.gov
[2] Advanced Institute for AI, Brazil
f.leno@usp.br
[3] Vrije Universiteit Brussel, Belgium
[4] Microsoft Research, USA

Reinforcement Learning (RL) has achieved many successes in training autonomous agents to perform simple tasks over the last years. However, there are a number of major remaining challenges in RL, such as scaling it to high-dimensional, real-world applications, dealing with sample complexity, or reusing already acquired knowledge. Although many works have already focused on strategies to scale-up RL techniques and to find solutions for more complex problems with reasonable successes, many issues still exist. The second *Scaling-Up Reinforcement Learning (SURL)* workshop encouraged the discussion of diverse approaches to accelerate and generalize RL, such as the use of approximations, abstractions, hierarchical approaches, Transfer Learning, and Meta-Learning. Scaling-up RL methods has major implications on the research and practice of complex learning problems and will eventually lead to successful implementations in real-world applications. *SURL* is an effort towards bridging the gap between conventional and scalable RL approaches, providing a platform for community interaction and discussion. This workshop aims at bringing together researchers working on different approaches to scale-up RL to solve more complex or larger scale problems.

This edition of the *SURL* workshop was the second after the first one during the European Conference on Machine Learning (ECML) 2017 in Skopje. It sparked the interest of many RL researchers around the globe. With 2 invited speakers and 12 accepted high-quality contributed talks, *SURL* was the stage for discussions about opportunities and future research directions for RL research. Presenters from 3 continents shared their newest research efforts with a sizable crowd of RL enthusiasts in one of the best visited workshops of the conference.

The 2 invited speakers *Peter Stone*, from UT Austin, and *Balaraman Ravindran*, from IIT Madras, are well-known and reputed researchers who shared interesting insights on their current research.

Balaraman Ravindran talked about extending RL beyond the classical "reward-based" modeling approaches. He discussed how his group managed to enable agents to

http://surl.tirl.info/

solve increasingly difficult tasks through a two-level *curriculum* learning technique, and how agents can discover hierarchical structure in problems by exploiting the properties of successor representations.

Peter Stone presented his latest research on scaling-up RL to enable efficient robot skill learning. Due to the inherent restriction on performing exploration on robotic platforms, his group has been studying how to reuse knowledge gathered from simulations and human demonstrations to speed up and improve learning in this challenging domain.

Amongst the outstanding contributions (published on the workshop website), *Extending Sliding-step Importance Weighting from Supervised Learning to Reinforcement Learning* by *Tian Tian* and *Richard Sutton* was selected to represent our workshop in this special volume. Their work extends a class of elegant algorithms for importance weighting from supervised learning to the RL framework. These algorithms are much more robust in the face of highly variable importance weights in supervised learning but had not been used in RL before where importance weighting can be particularly variable due to the sampling involved in off-policy learning algorithms. The workshop chairs believe this work was well presented and covers a challenging problem in contemporary RL research. It is well therefore suited to give a taste of the insightful and creative solutions that were presented and discussed during *SURL*.

The workshop also featured many other interesting approaches from subfields including hierarchical approaches and curriculum learning, multiagent settings, multitask and subgoal challenges, and reward shaping. Experimental results were reported in applications like robotics, physics simulator, Multiplayer Online Battle Arena games, and Atari games. We highly recommend taking a look at the workshop homepage and get inspired by the great contributions.

Both the contributed and invited talks have shown that RL research has evolved far beyond the classical framework and toy problems. Varied strategies for knowledge reuse, sample-efficient learning algorithms, and smart exploration are necessary to solve the challenging domains in which RL is now applied.

Acknowledgments

We would like to thank the Artificial Intelligence Journal (AIJ) editorial board for their funding which enabled us to provide student travel grants to support the promotion and dissemination of AI research.

We also thank all unmentioned authors and presenters for their contributions: A. Remonda, S. Krebs, E. Veas, G. Luzhnica, R. Kern, Y. Deng, K. Yu, D. Lin, X. Tang, C. C. Loy, A. Ray, R. Verma, H. Khadilkar, S. Han, Y. Sung, H. Wei, K. Decker, Z. Zhang, H. Li, L. Zhang, T. Zheng, T. Zhang, X. Hao, X. Chen, M. Chen, F. Xiao, W. Zhou, A. Bassich, D. Kudenko, D. Chen, Q. Yan, S. Guo, Z. Yang, X. Su, F. Chen, H. Itaya, T. Hirakawa, T. Yamashita, H. Fujiyoshi, L. Zhang, Z. Zhang, Z. Pan, Y. Chen, J. Zhu, Z. Wang, M. Wang, C. Fan, R. Saphal, B. Ravindran, D. Mudigere, S. Avancha, and B. Kaul.

Ruben Glatt was supported by a postdoctoral fellowship at the Lawrence Livermore National Laboratory. His portion of the work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC. LLNL-PROC-794262. Felipe L. Silva thanks the São Paulo Research Foundation (FAPESP), grant 2015/16310-4.