UT Austin Villa: An Open-source Gym Environment for Deep Reinforcement Learning on 3DSim

Bo Liu*[1], William Macke*[1], Caroline Wang*[1], Patrick MacAlpine[2], and Peter Stone[1,2]

[1] The University of Texas at Austin, Austin, TX 78712, USA
[2] Sony AI, USA

This presentation discusses a technical environment created by the UT Austin Villa team for using Deep Reinforcement Learning (RL) to learn to walk / run in the RoboCup 3D simulation environment. We have developed a custom OpenAI Gym [2] environment in Python, that connects to the 3DSim platform via network sockets. By wrapping the task in such an environment, we may apply existing deep RL libraries — which are mostly written in Python — to the existing C++ simulation software.

Fig. 1. left: the 3D simulation environment. **right:** the gym-like environment.

The code is organized as follows. We first create a C++ API for interacting with the 3DSim software. This API is used to provide relevant information for the environment, such as the state space and the reward function. We then interface with C++ API through a separate Python process. The Python process creates an instance of a custom Gym environment that provides the states and actions as vectorized values.

We hope that our custom Gym environment will enable the RoboCup community to leverage recent advances in deep RL to improve skills in the 3D simulation environment for which there have already been some promising results [1,3]. We also hope that our Gym environment will encourage the deep RL research community to adopt 3DSim as a benchmark task. We intend to release the Gym environment as an addition to the UT Austin Villa base code release.[1]

References

1. M. Abreu, L. P. Reis, and N. Lau. Learning to run faster in a humanoid robot soccer environment through reinforcement learning. In *Robot World Cup*, pages 3–15. Springer, 2019.
2. G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym. *CoRR*, abs/1606.01540, 2016.
3. L. C. Melo and M. R. O. A. Máximo. Learning humanoid robot running skills through proximal policy optimization. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 37–42. IEEE, 2019.

* indicates equal contribution. Authors are listed in the alphabetical order of their surnames.
[1] `https://github.com/LARG/utaustinvilla3d`