Intrinsically Motivated Model Learning on a Nao

UT Austin Villa
The University of Texas at Austin

Reinforcement learning agents normally learn a specific task based on a pre-defined reward function. However, sometimes an agent may be given the opportunity to gain experience in a domain prior to being given a specific task. In these cases, intrinsic motivation can be used to enable the agent to learn a useful model of the environment.

Previous work by two of our team members presented the TEXPLORE with Variance-And-Novelty-Intrinsic-Rewards algorithm (TEXPLORE-VANIR), an intrinsically motivated model-based reinforcement learning algorithm [1, 2]. The algorithm learns a model of the transition dynamics of a domain and then calculates two different intrinsic motivations from this model. The first intrinsic motivation is to explore where the model is uncertain, while the second is to acquire novel experiences that the model has not yet been trained upon. The combination of these intrinsic rewards enables TEXPLORE-VANIR to learn an accurate model of a domain with no external rewards.

In our Open Challenge demo, we will overview TEXPLORE-VANIR and show videos of using TEXPLORE-VANIR to allow a Nao to learn the transition dynamics of a simple environment (shown in Figure 1). We are currently exploring the use of TEXPLORE-VANIR to learn behavior to approach the ball, so we may discuss this as well.
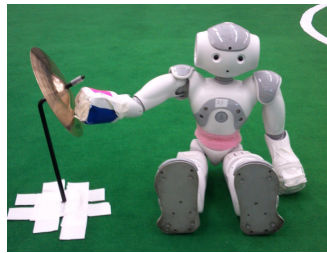
References

Figure 1: Environment for our experiments.

[1] T. Hester and P. Stone. Intrinsically motivated model learning for a developing curious agent. In *The Eleventh International Conference on Development and Learning (ICDL)*, Nov 2012.

[2] T. Hester and P. Stone. TEXPLORE: Real-time sample-efficient reinforcement learning for robots. *Machine Learning*, 90(3), 2013.