Evento	Salão UFRGS 2019: SIC - XXXI SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2019
Local	Campus do Vale - UFRGS
Título	Identifying Reusable Early-Life Options
Autor	ALINE WEBER
Orientador	BRUNO CASTRO DA SILVA

Identifying Reusable Early-Life Options

Aline Weber (Prof. Bruno Castro da Silva)
Federal University of Rio Grande do Sul

When designing robots that autonomously learn to solve tasks, Reinforcement Learning (RL) methods are commonly used. RL methods allow agents to improve their behaviors by interacting with an environment and observing the results of their actions. One important challenge in RL is that during the initial moments of an agent's lifetime, its behavior is essentially random, which often results in catastrophically poor performance. A common way of dealing with this issue is to equip the agent with motor priors known as *options*. These are behaviors that can be reused in many situations during the agent's lifetime; consider, e.g., options for grasping different types of objects or to regain its balance if pushed.

Most of the existing state-of-the-art methods for learning options focus on identifying options that help an agent to perform well *throughout its entire lifetime*. **We, by contrast, wish to identify options that allow robots to perform well in the *very early stages* of their lives**. We refer to options of this type as *early-life options*. They are important whenever agents need to operate in environments where the use of poor-performing policies (such as the random policies with which they are typically initialized) may be catastrophic. <u>Early-life options are akin to primitive reflexes in infant mammals—agents born with them, even if acting randomly, are capable of producing rudimentary behaviors comparable to those acquired by agents that actively optimize a policy for hundreds of thousands of steps.</u>

We introduce a method[1] that augments the action set of an agent with specially-constructed reusable early-life behaviors that maximize performance over a possibly infinite family of related motor tasks. We also introduce three novel metrics for evaluating early-life options: **(1)** Maximum-Mean metric, which selects options with high expected performance; **(2)** Negative-Tail Averse metric, which favors options with both high expected performances and with a low probability of producing poor (possibly catastrophic) results; and **(3)** Positively-Skewed metric, which favors options with high expected performance and that maximize the probability of behaviors with above-average quality. We show that these metrics allow for the discovery of early-life behaviors that maximize expected return while minimizing the risk of producing catastrophically poor performances.

We evaluated our technique on three challenging high-dimensional simulated robots. These robots were tasked with learning to walk efficiently under different battery consumption constraints. **Our experiments show that our method is capable of automatically learning behaviors akin to primitive reflexes in infant mammals: our "infant agents", born with early-life options, are capable of directly producing rudimentary behaviors with performances comparable to those acquired by learning agents trained for hundreds of thousands of steps.**

---

[1]Aline Weber, Charles P. Martin, Jim Torresen, Bruno C. da Silva. **Identifying Reusable Early-Life Options**. Proceedings of the 9th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (**ICDL 2019; Oslo, Norway**).