



# Identifying Reusable Early-Life Options

Aline Weber; Prof. Dr. Bruno C. da Silva

## Motivation

- **Options**
  - behaviors a robot can reuse in many situations
  - e.g., grasping different objects; opening doors
- Options discovered by existing methods
  - help an agent during its entire lifetime
- We propose to identify **early-life options**:
  - **short-duration** evolved behaviors
  - **reusable** across many related tasks
  - akin to **primitive reflexes** in mammals
  - allow robots to perform well
    - in the **very early stages** of their lives

## Method

(a) Augment an agent's action set with early-life behaviors:

### 1. Generate Candidate Options

- Sample and solve a few sample tasks
- Observe/record short chunks of the behaviors

### 2. Estimate Performance Distribution of Options

- Evaluate behaviors over wide range of situations
- Estimate behaviors' performance distribution

### 3. Select Top Candidate Early-Life Options

- Select top candidate behaviors according to a given metric

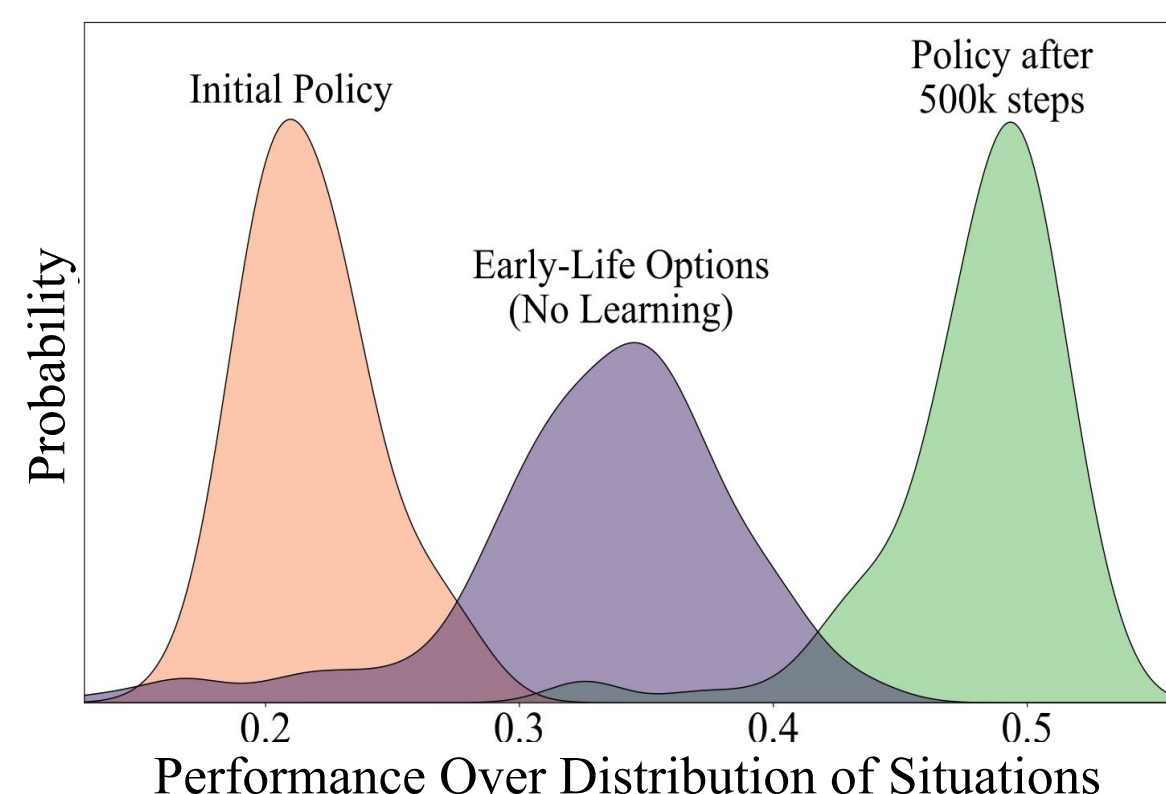
## (b) Evaluate Early-Life Options (ELOs)

We introduce three novel metrics:

- *Maximum-Mean metric*:
  - $\psi_{\mu}(\pi_o) \doteq \mathbb{E}[V(\pi_o)]$
  - creates options with **high expected return**
- *Negative Tail-Averse metric*:
  - $\psi_{-}(\pi_o) = \mathbb{E}[V(\pi_o)] - k\Pr(V(\pi_o) < (\mu^* - \alpha\sigma^*))$
  - options with **high expected return** and **low probability of producing poor performances**
- *Positively-Skewed metric*:
  - $\psi_{+}(\pi_o) = \mathbb{E}[V(\pi_o)] + k\Pr(V(\pi_o) > \mu^*)$
  - options with **high expected return** and **with maximum probability of above-average quality**

## Experiments

- Three high-dimensional simulated robots
  - walking efficiently under *different battery costs*



### Agents' performances

- Typical newborn agent
- Trained for 500k steps
- Our newborn agent
  - untrained
  - equipped with 5 ELOs

## Results

- Our method is capable of learning behaviors
  - **akin to primitive reflexes in infant mammals**
- Our newborn, untrained agents equipped with ELOs
  - **performance similar to agents trained for ~200k+ steps!**