CURIous: Intrinsically Motivated Modular Multi-Goal RL

Cédric Colas
Phd student @ Flowers team, INRIA

Co-authors:
Pierre Fournier, Olivier Sigaud, Mohamed Chetouani, Pierre-Yves Oudeyer
Problem: Intrinsically Motivated Modular Multi-Goal RL

Which type of goal should I target?

Reach, Push, Pick & Place, Stack .. ?

Modular Multi-Goal Fetch Arm environment
Problem: Intrinsically Motivated Modular Multi-Goal RL

Modular Multi-Goal Fetch Arm environment

Which goal exactly?
Pick & Place at (x,y,z)!
Problem: Intrinsically Motivated Modular Multi-Goal RL

Modular Multi-Goal Fetch Arm environment

Controllable objects (learnable goals)

Distracting objects (unlearnable goals)
The Curious Algorithm

Modular goal encoding for UVFA:¹
e.g. of modular goals:
- Move gripper to \((x,y,z)\)
- Pick & Place cube2 at \((x,y,z)\)
- Push cube1 at \((x,y)\)

Sampling of modules and goals using \textit{absolute learning progress}²
(using Bandit algorithm)

External world

Modular replay buffer:
with hindsight learning³, ⁴
(module and goal substitutions)

External world

1: UVFA, Schaul et al., 2015
2: IMGEP, Forestier, 2017
3: HER, Andrychowicz et al., 2017
4: Unicorn, Mankowitz et al., 2018
Modular goal encoding vs Multi-Goal Module Experts

Impact of the policy and value function architecture. Average success rates over the set of tasks (mean +/- std, 10 seeds).

Curious: Intrinsically Motivated Modular Multi-Goal RL
Automatic Curriculum with Absolute Learning Progress

Forgetting due to interferences among modules/goals

Mitigated thanks to fast LP-based refocus

Using a bandit for module selection and replay

Competence

Absolute Learning Progress

Selection Probabilities

Reach
Push
Pick & Place
Stack

Curious: Intrinsically Motivated Modular Multi-Goal RL
Resilience to distracting goals: 0, 4 or 7 distracting modules. CURIOUS (intrinsically motivated) and Random (random module). Mean +/- sem, 10 seeds.
Resilience to Forgetting and Sensory Failures

Resilience to sensory failure: Recovery following a sensory failure. Mean +/- std, 10 seeds.

CURIOSUS recovers 95% of its original performance twice as fast as Random.

Curious: Intrinsically Motivated Modular Multi-Goal RL
Curious: Intrinsically Motivated Modular Multi-Goal RL

Poster: Pacific Ballroom #42