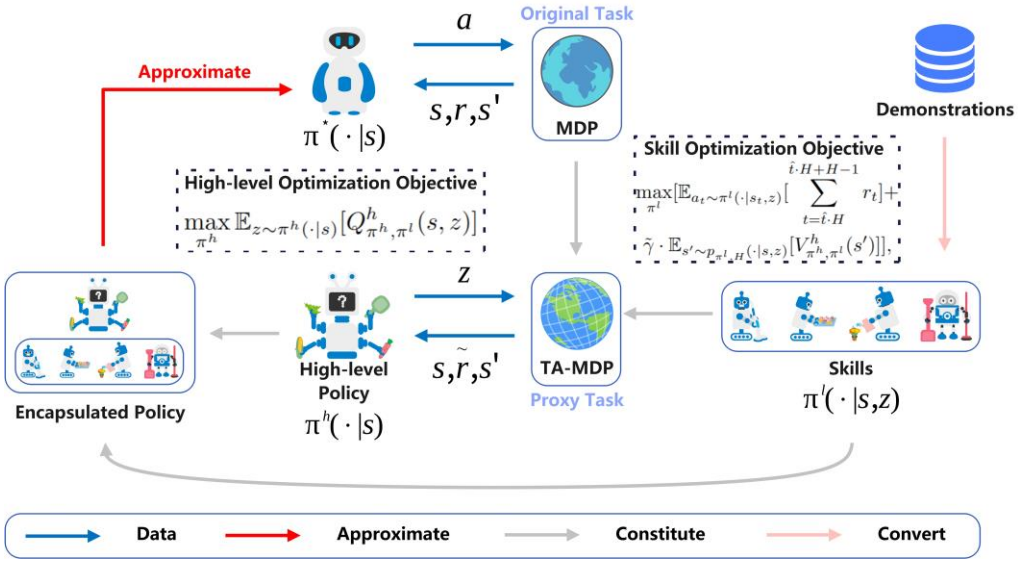
DSR: Optimization of Performance Lower Bound for Hierarchical Policy with Dynamical Skill Refinement

Dongxiang CHEN, Ying WEN

Frontiers of Computer Science, DOI: 10.1007/s11704-025-50561-3

Problems & Ideas

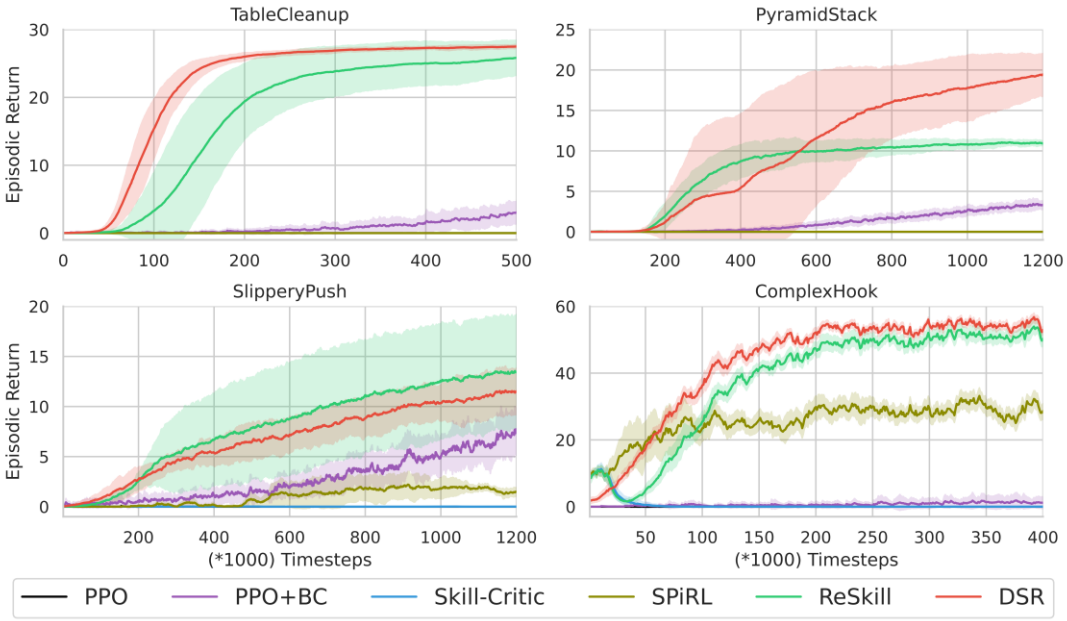
- Problems of Skill-based Reinforcement Learning:
  - Lack of optimization objectives to ensure performance improvement.
  - Skill refinement causes temporal abstraction shift;
  - Refining one skill can lead to uncertain changes in the behavior of other skills, resulting in skill space collapse.
- Ideas: Unified optimization objective to ensure improvement in performance lower bound.Learn skill refinement into a dynamically weighted residual policy.



Skill-based RL can be seen as a proxy task of original RL. Therefore, the essential purpose of skill-based RL is to ensure performance improvement. We propose the unified optimization objective for entire hierarchical policy. We prove that it can achieve performance improvement in TA-MDP and essentially optimize the performance lower bound in original MDP.

Main Contributions

- Contributions:
  - A unified optimization objective for entire hierarchical policy and theoretical foundations for performance improvement;
  - Dynamical skill refinement which prevents skills' behaviors from being disrupted by uncertain changes;
  - An on-policy RL algorithm that simultaneously updates both the high-level policy and skills avoids temporal abstraction shift.



Our method DSR can achieve outstanding performance in all the sparse-reward robotic manipulation tasks. Compare to the method keeping skills fixed (SPiRL), it achieves obviously higher aymptotic performance, which illustrates the necessity of skill refinement.