Self-supervised Reinforcement Learning with Independently Controllable Subgoals
Andrii Zadaianchuk1,2, Georg Martius1, Fanny Yang2
1Max Planck Institute for Intelligent Systems 2ETH Zürich

Abstract
Self-supervised agents can learn manipulation skills in multi-object environments. Previous methods do not take the dependencies between objects into account. We propose to estimate relations between objects and use them to independently control different objects. Estimated relations between objects can be used to decompose a complex goal into a compatible sequence of subgoals. Our SRICS agent can efficiently and automatically learn manipulation tasks in multi-object environments with different relations between objects.

SRICS

Object Manipulation in Multi-objects Environments
Multi-object Rearrange and Relational Rearrange environments

Object-centric Representations
Scene is encoded as a set of vectors
All the entities (including agent) are encoded with the same format
Similar representation could be learned fully-unsupervised from high-dimensional observations

SRICS Training
Given: GNN dynamical model D, on sequence from D using the ELBO loss and estimate the interaction graph G.
For n = 1, ..., K episodes do
Sample goal wstd and construct subgoal g∥ using G.
Collect episode data with policy π∥(a∥ | h∥, g∥).
Store transitions (s∥, a∥, h∥, r∥, s∥∥) into replay buffer R.
Sample transitions from replay buffer (s∥, a∥, h∥, g∥, s∥∥) ~ R.
Relabel g∥ goal components to a combination of future-states and goal sampling distribution.
Compute selectivity reward signal R = r∥(s∥, a∥, g∥).
Update policy π∥(a∥ | h∥, g∥) using R with SAC trainer.
end for

SRICS Evaluation
Agent with compositional skills

References