Graph Embedding Priors for Multi-task Deep RL

N. Parikh[*], Z. Horvitz[*], N. Srinivasan[*], A. Shah, G. Konidaris
Brown University

Overview

Humans efficiently generalize behavior across different tasks. We exploit graph-based priors with a flexible, modular architecture to obtain such sample efficiency over multiple RL tasks. We introduce *Symbolic Procgen*, a procedurally-generated, symbolic, multi-task environment and demonstrate improved sample complexity over competing methods.

Symbolic Procgen

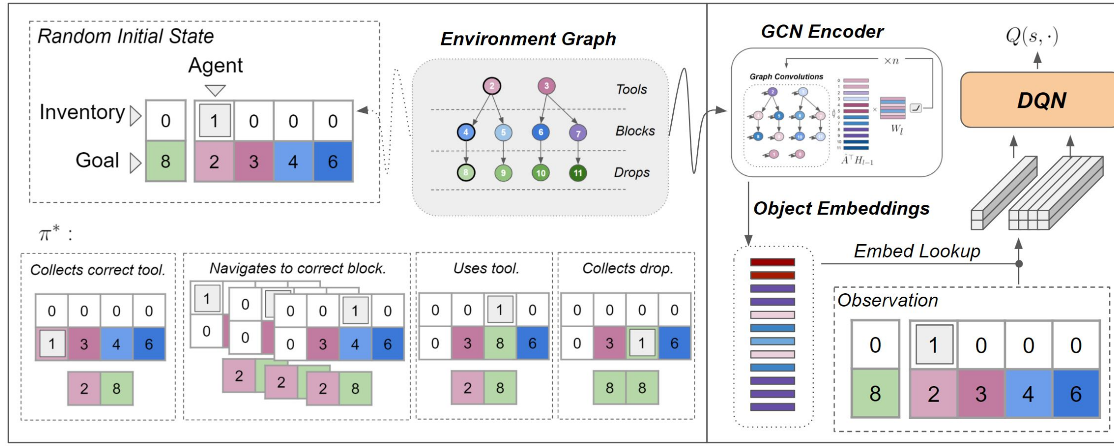
When dealing with multiple tasks, agents can learn efficiently by reusing information. *Symbolic Procgen* effectively tests this by providing a scalable domain with many tasks that share structure but involve different sets of objects.

GEM-RL

An effective way of encoding prior knowledge about the world is via a knowledge graph. Our architecture's **GCN encoder** uses this graph to reason about objects in the state *and* across tasks.

$$a_v^{(l)} = \mathrm{mean}\left(\left\{h_u^{(l)}, u \in \mathcal{N}(v) \cup \{v\}\right\}\right) \qquad h_v^{(l+1)} = \mathrm{ReLU}\left(W^{(l)} a_v^{(l)}\right)$$

For any object, the encoder combines its embedding with those of its neighbors, allowing these embeddings to include relational information across tasks. These are used to augment the state, which is passed into the DQN. Our method's flexibility allows this to sit atop of any RL algorithm. When evaluated on *Symbolic Procgen*, GEM-RL scales to hundreds of tasks with hundreds of different objects and relations.

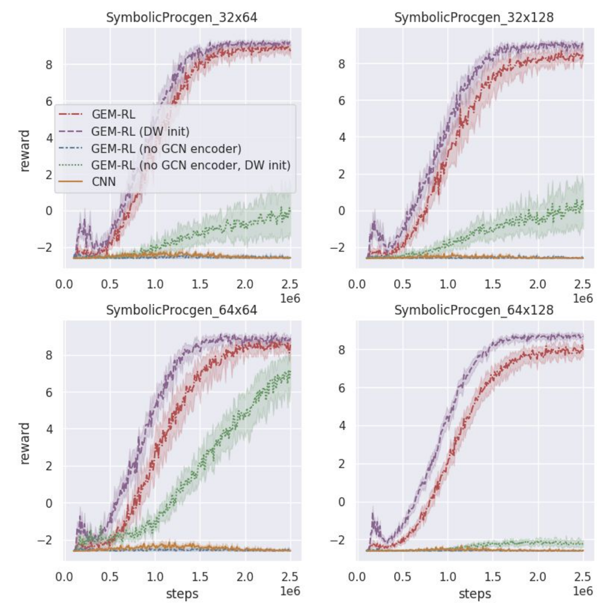


Left: An example Symbolic Procgen environment initialized with ten objects: two tools (objects 2,3), four blocks (objects 4-7), and four drops (objects 8-11). Following an optimal policy (π^*), the agent will collect the correct tool (tool 2), navigate to the correct block (block 4), use the tool on the block, and then collect the resulting drop. Right: Model architecture.

Perspective

We demonstrate that graph-based priors are effective at improving sample complexity for multi-task environments. We show that state agnostic, object representations perform very well in complex multi-task settings. We also find that performance gracefully degrades as graph-based priors become less reliable.

We plan to explore the effectiveness of GEM-RL with knowledge graphs derived from structured databases and unstructured sources such as language. Additionally, we hope to leverage computer-vision techniques to obtain sample efficient learning in domains with pixel-based observations.



Results on *Symbolic Procgen*. For further results, see our paper.

neev_parikh@brown.edu