The Role of Schemas in Reinforcement Learning: Insights and Implications for Generalization

Mandana Samiei[13*], Doina Precup[123], Blake A. Richards[13]

[1] McGill University [2] Google DeepMind [3] Mila - Quebec AI Institute

Overview

Humans are remarkably good at learning **new tasks** by building on what they already know. This ability relies on **schemas**, structured mental representations that help us to understand the most likely **relationship between objects and sequence of events** in an environment.

Contribution

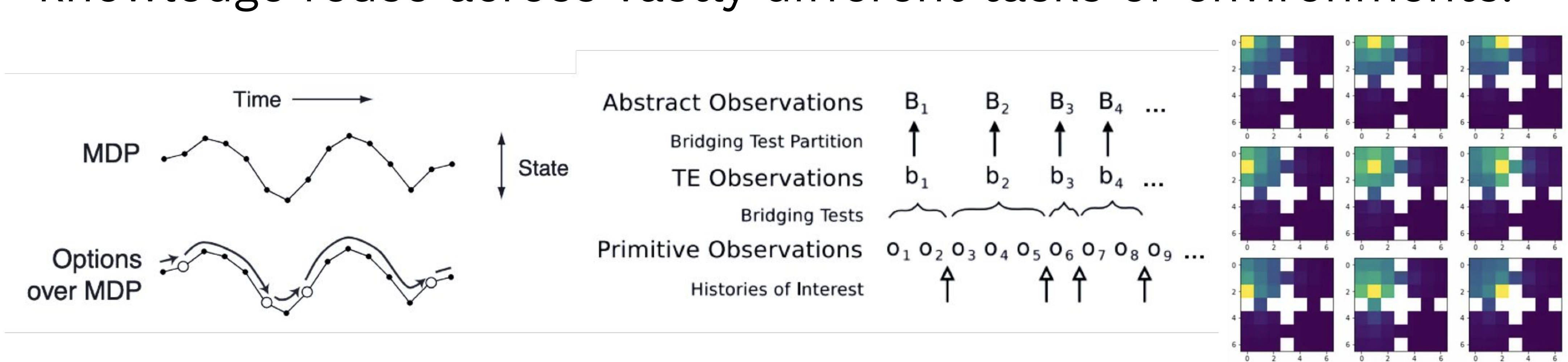
We **propose a method for learning schemas in RL.** The importance of learning schemas in RL lies in their potential to improve decision-making by enabling agents to perform **hierarchical planning** and r**eusing previously learned sub-tasks** for novel tasks. First, we focus on the identification and learning of **bottleneck features**—key components of a task that significantly impact the agent's ability to succeed. Second, how these features **can be organized into schemas that structure** knowledge in a hierarchical manner.

Related Work

Options: Lack the flexibility to adapt to new environments.

Partial Models: Capture limited knowledge of the environment, often modelling only specific aspects.

Successor Features: They do not inherently support modular knowledge reuse across vastly different tasks or environments.

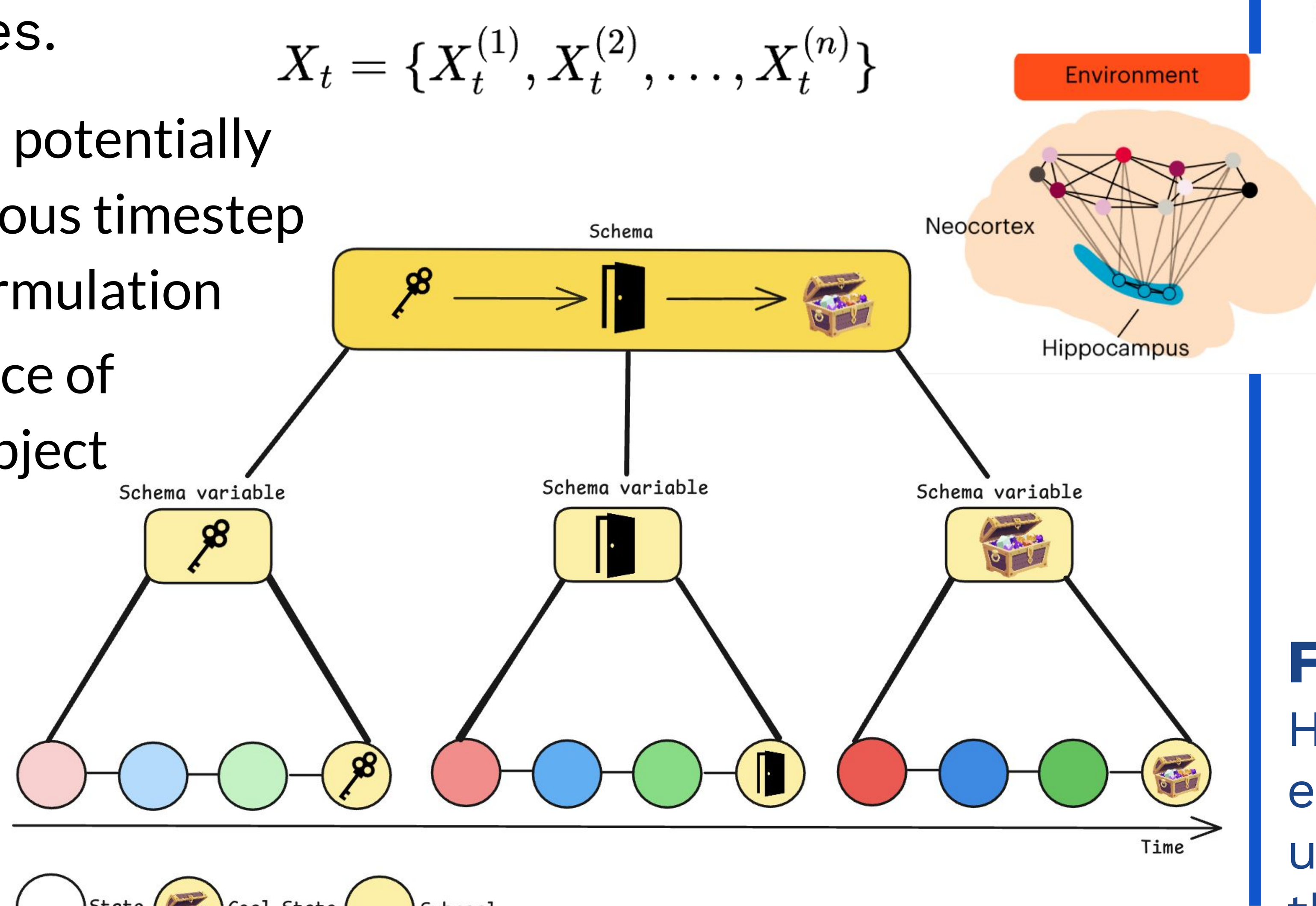


Methodology

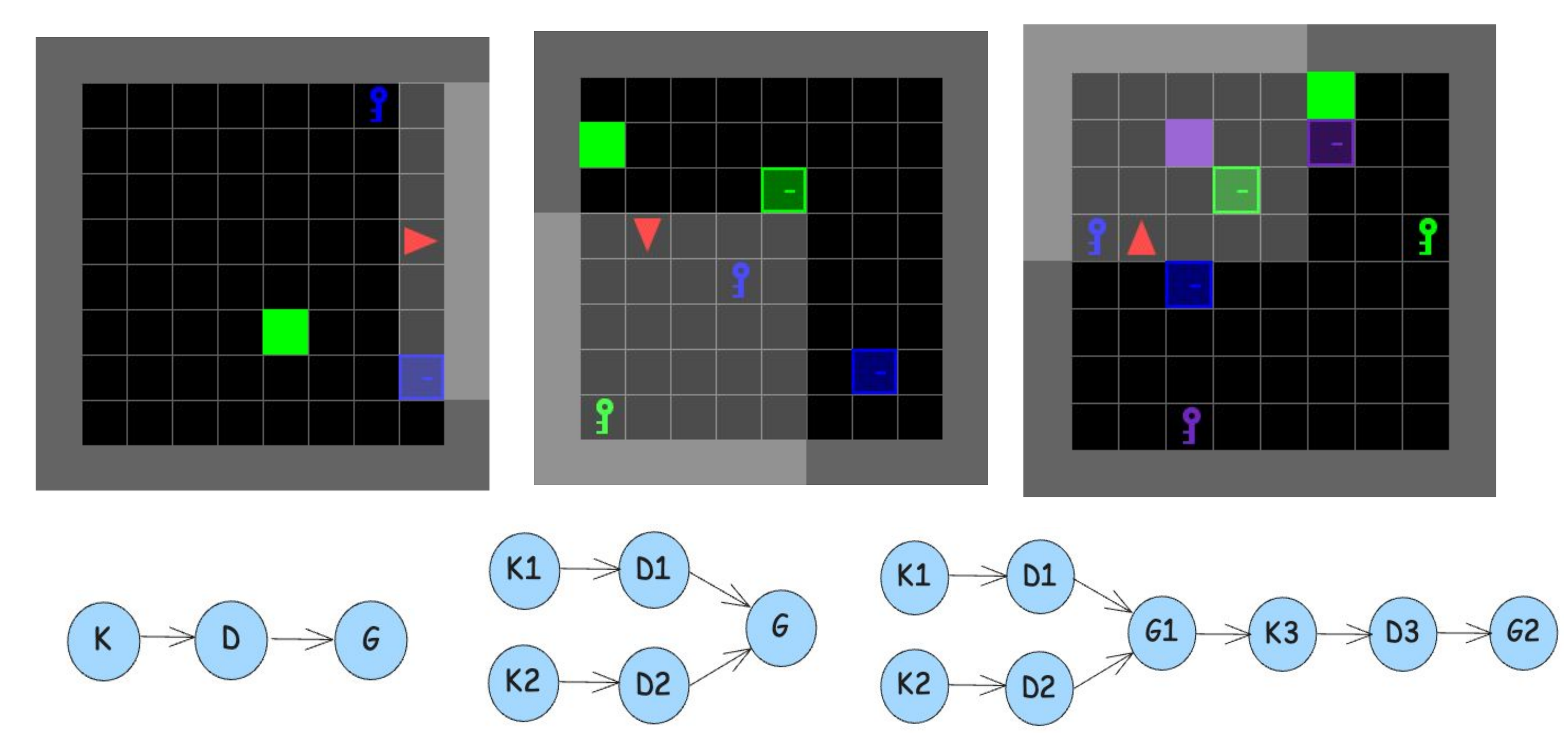
We construct a framework where the abstract level consists of a Structural Causal Model (SCM). Each abstract state consists of a set of abstract variables.

$$X_t = \{X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(n)}\}$$

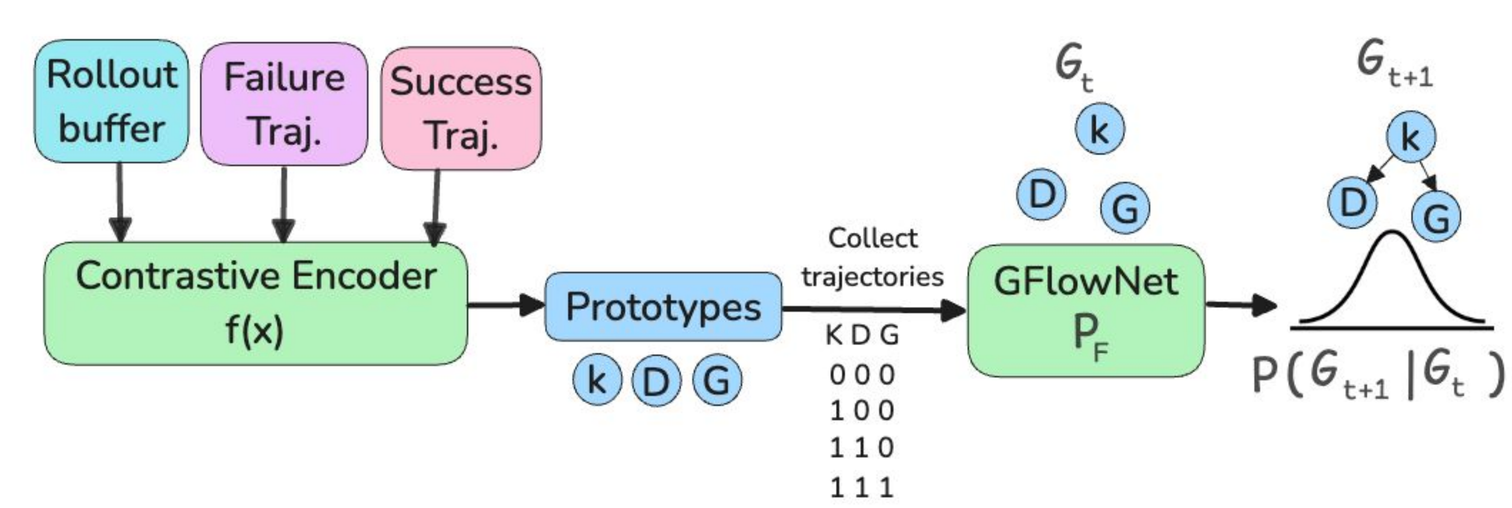
Each abstract variable can potentially depend on the set of previous timestep abstract variables. This formulation naturally give rise to a space of schemas about possible object relationships. We can test them by training / executing *conditioned policies* at the low level, π (a|s, X).



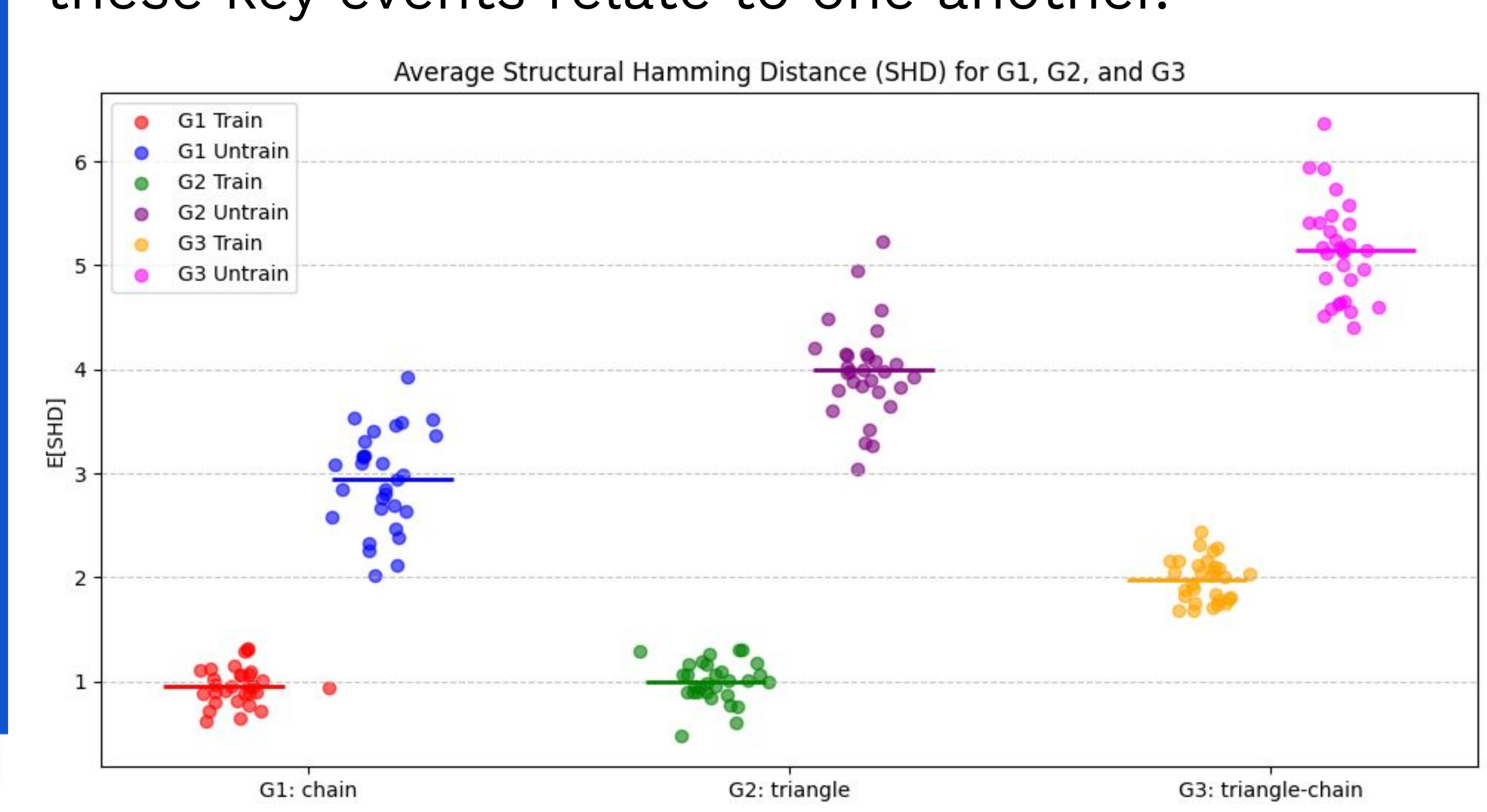
Experimental Results



Consider a grid world environment: a schema consists of discovering the relationship that 'goal' can be reached, only if 'key' is picked, and 'door' is open.



Our model contains two critical steps. First, we train agents to identify key events that are crucial for completing a task, using a contrastive learning technique. Next, we train a generative model over graphs structures that describe how these key events relate to one another.



For G1 (chain), the trained agent show very low SHD (around 1), indicating that the agent has recovered the underlying structure with high accuracy. In G2 (triangle), the SHD for the trained structure also remains low (close to 1), which confirms that the agent effectively captures the dependencies in a more complex task.

Future Work: Schemas are key components for Hierarchical planning, and chunking prior experiences into subgoals. Schemas are particularly useful in environments where the critical states stay the same, but the sensory observations change.