Contrastive Representations for Temporal Reasoning

Alicja Ziarko, Michał Bortkiewicz, Michał Zawalski, Benjamin Eysenbach*, Piotr Miłoś*

Motivation

- Many practical problems are inherently combinatorial in nature—for example, **traffic control**, **job scheduling**, and the **traveling salesman problem.**
- They are usually solved using complicated heuristic/neurosymbolic methods - we are taking a step toward solving them with purely neural approaches.
- Contrastive learning can result in representations which make planning easy [1].

Standard CRL Fails

CRL uses InfoNCE [2, 3] loss to train the critic:

$$\mathbb{E}_{\substack{(x_j,x_{j+})\sim\mathcal{P}(X,X_+),\\ x^k_{j-}\sim\mathcal{P}(X)}}\left[\frac{1}{N}\sum_{j=1}^{N}\frac{e^{f(x_j,x_{j+})}}{e^{f(x_j,x_{j+})}+\sum_{k=1}^{N-1}e^{f(x_j,x^k_{j-})}}\right]$$

- CRL is unable to the temporal structure in problems such as Sokoban.
- This is due to reliance on non-temporal features, specifically walls in the Sokoban example.
- The usual InfoNCE loss minimizes: $I(X;X_+)$
- To remove the context from representations, we instead minimize: $I(X;X_+) - I(X_+;C)$
- We propose a different objective:

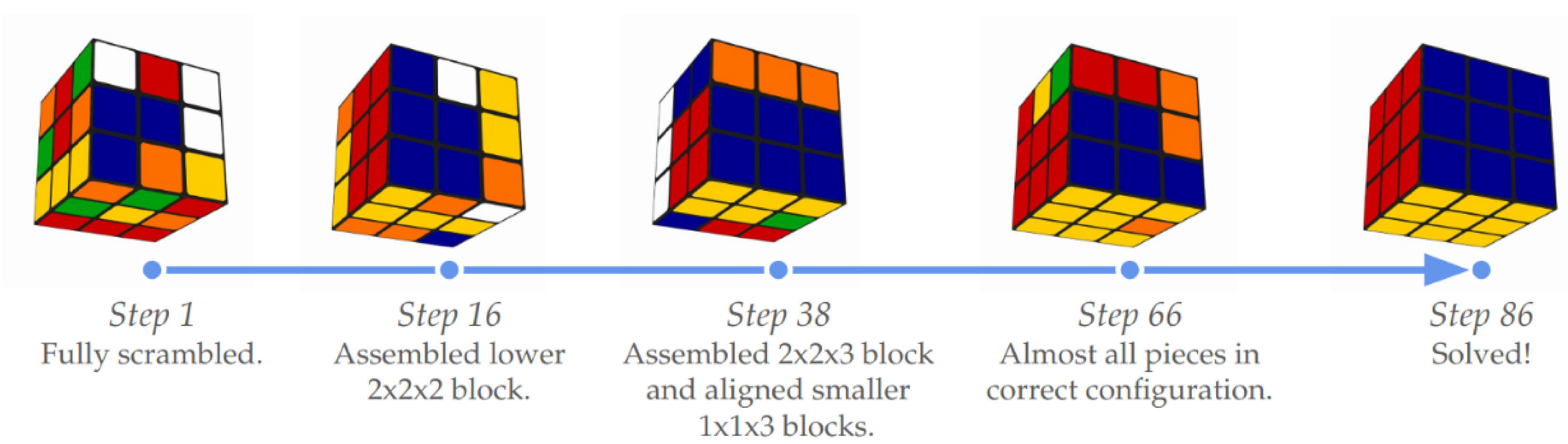
$$\mathbb{E}_{\substack{c\sim\mathcal{P}(C),(x_j,x_{j+})\sim\mathcal{P}(X,X_+|C),\\ x^k_{j-}\sim\mathcal{P}(X|c)}}\left[\frac{1}{N}\sum_{j=1}^{N}\frac{e^{f(x_j,x_{j+})}}{e^{f(x_j,x_{j+})}+\sum_{k=1}^{N-1}e^{f(x_j,x^k_{j-})}}\right]$$

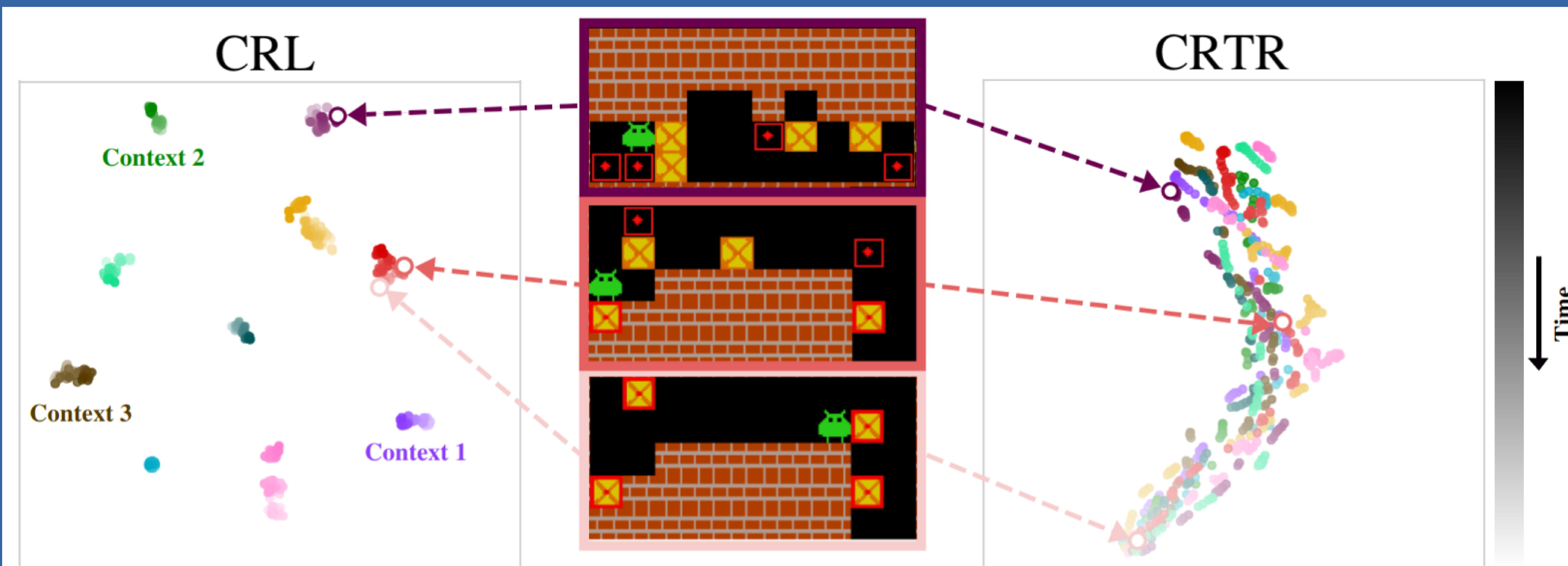
Practical Method

- In practice we might not know what the context is, and the context might not even be constant (Rubik's Cube).
- We propose a simple method which does not require that and can be added to standard CRL in a few lines of code:

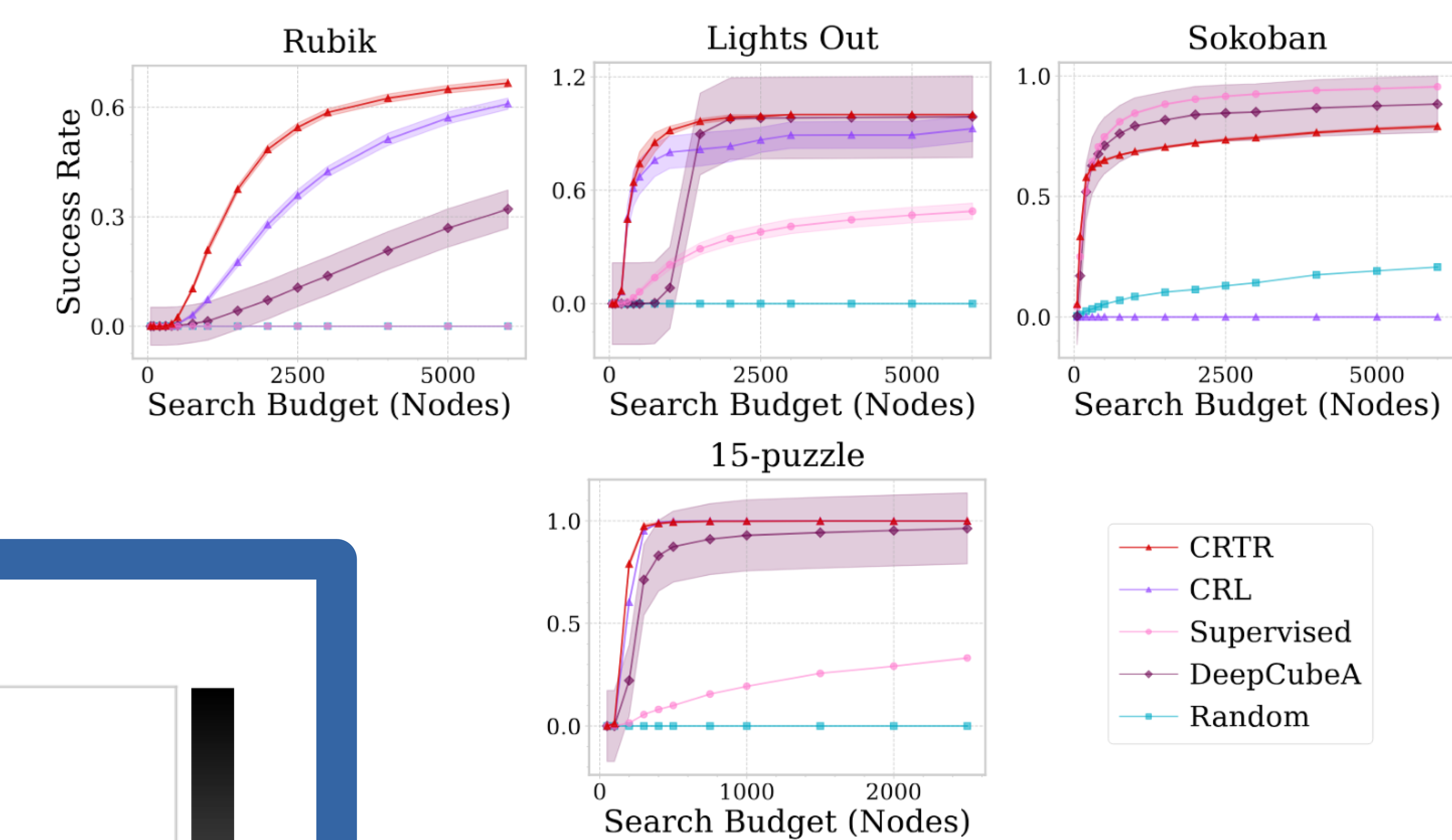
```
# dataset.shape == [num_traj, traj_len, obs_dim]
t0 = np.random.choice(dataset.shape[1], batch_size)
t1 = t0 + np.random.geometric(1 - discount, batch_size)
traj_id = np.random.choice(dataset.shape[0], batch_size)
# 1 new line of code for CRTR (our approach):
traj_id = np.repeat(traj_id[:batch_size // repetition_factor],
                    repetition_factor, axis=0)
batch = (dataset[traj_id, t0], dataset[traj_id, t1])
# further batch processing, the same for CRL and CRTR
```

We Can Solve Problems without Explicit Search

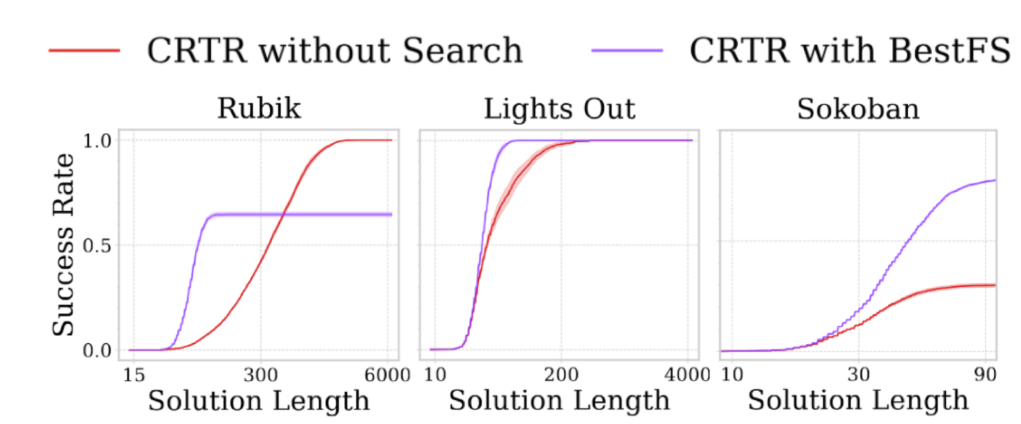




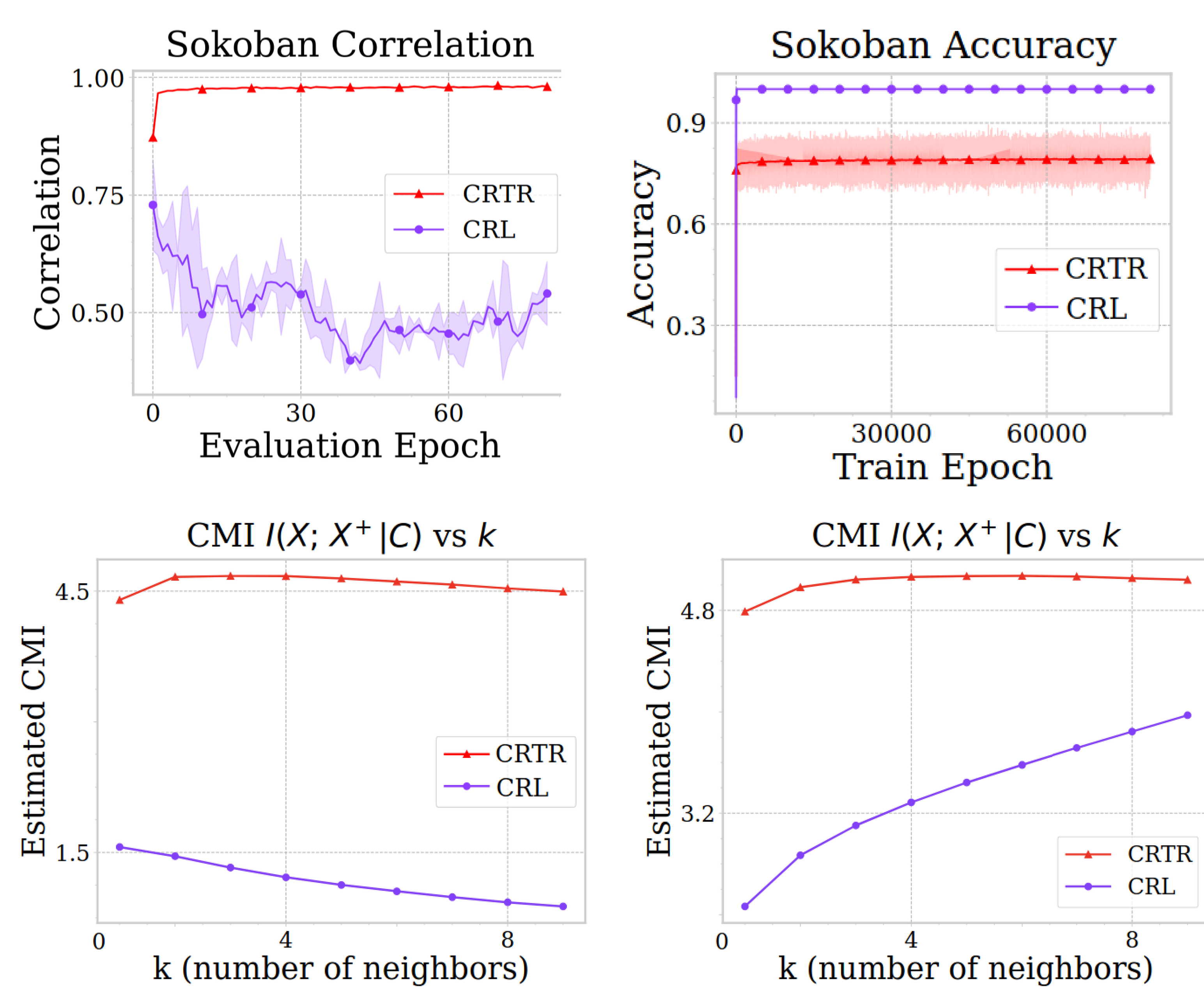
CRTR Representations Facilitate Planning



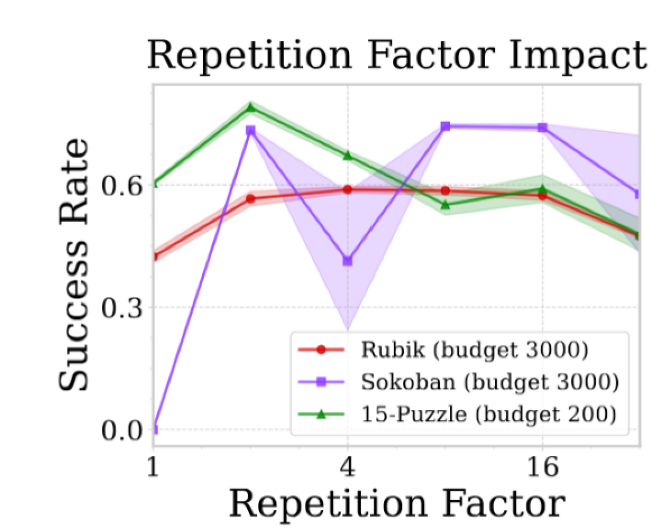
No Explicit Search

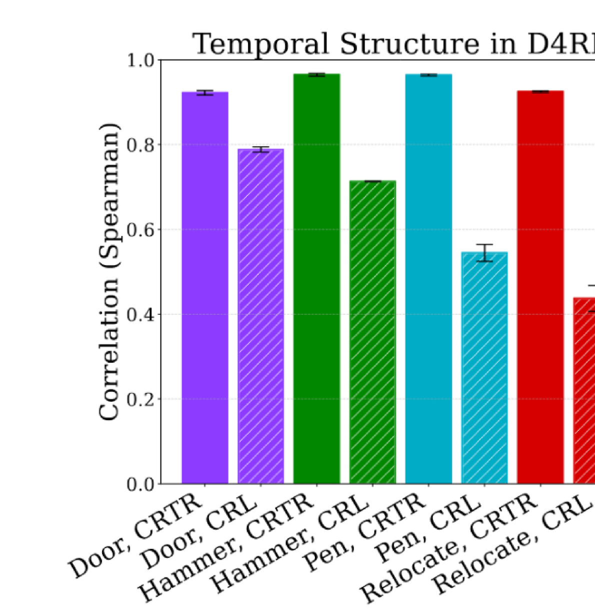


What Happens in Sokoban?



Additional Results





References:

[1] Eysenbach et al.: Inference via Interpolation: Contrastive Representations Provably Enable Planning and Inference, NeurIPS 2024

[2] van den Oord et al.: Representation Learning with Contrastive Predictive Coding, CoRR abs/1807.03748

[3] Eysenbach et al.: Contrastive Learning as Goal-Conditioned Reinforcement Learning, NeurIPS 2022

[4] Agostinelli et al.: Solving the Rubik's cube with deep reinforcement learning and search. Nat. Mach. Intell. 2019

[5] Czechowski et al.: Subgoal Search For Complex Reasoning Tasks. NeurIPS 2024

Project Page

Paper