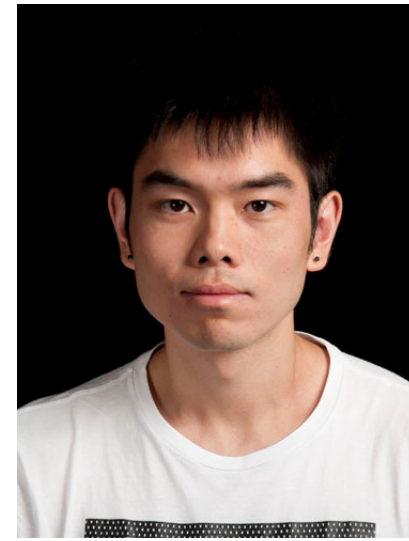


Causally Aligned Curriculum Learning

ICLR 2024, Paper #6200, Halle B #60



Mingxuan Li



Junzhe Zhang



Elias Bareinboim

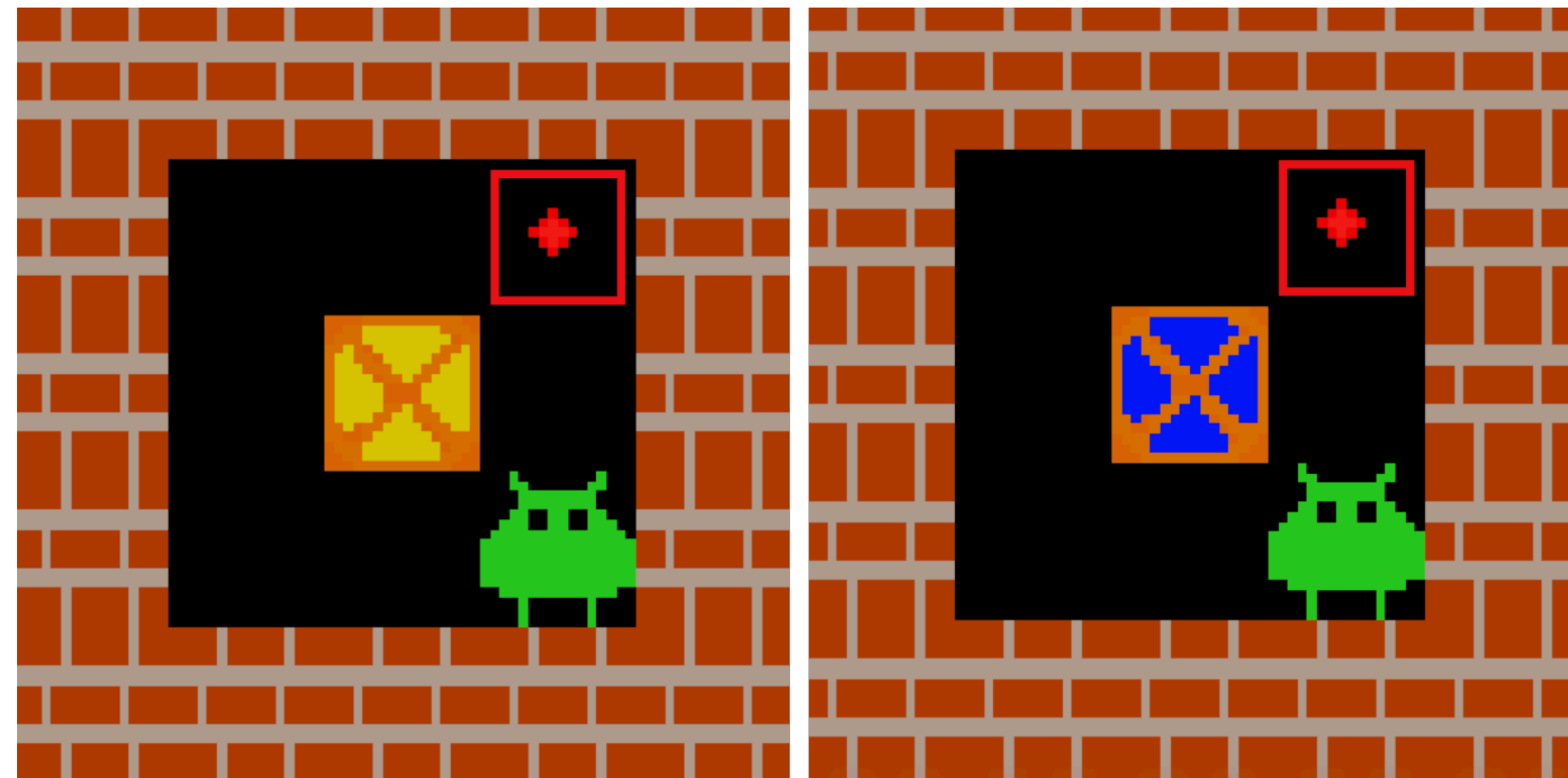
Causal AI Lab @ Columbia University

Motivations

- In complex environments, humans may be guided to learn through a modality called curriculum learning. The expectation in these settings is that when some optimal decision rules are shared across source and target environments, the agent could more quickly pick up the necessary skills to behave optimally in the target environments.
- In practice, however, the critical assumption of an invariance of the optimal policy does not necessarily hold when the underlying environment contains unobserved confounders.
- Our goal in this task is to characterize the conditions under which the invariance of optimal decision rules holds in curriculum learning and to generate such curricula to help agents learn more efficiently.

Colored Sokoban - A Motivating Example

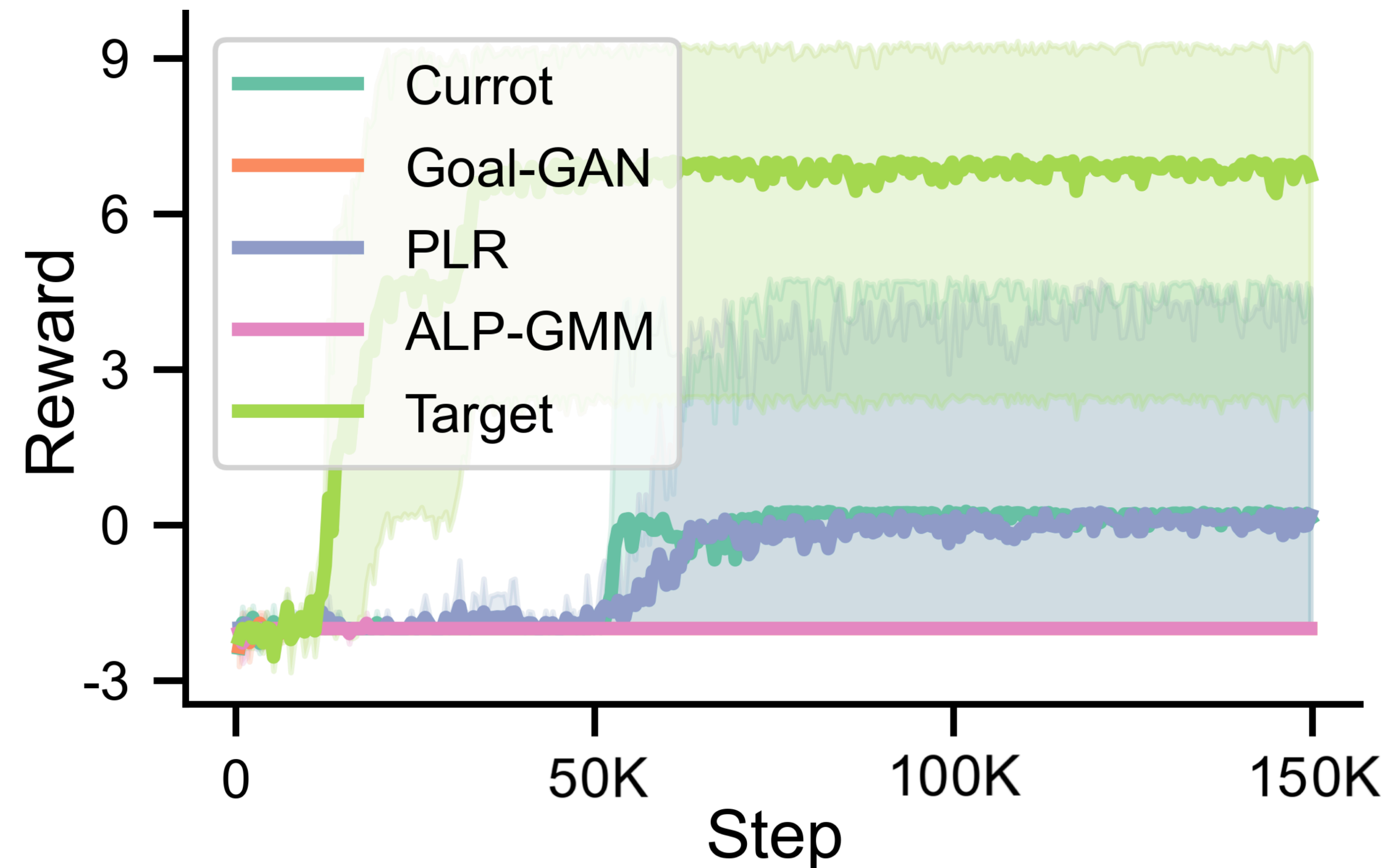
Goal: Find a policy (π) so as to maximize cumulative rewards in the target task.



The box color changes at each time step. The target task is to push a **yellow box** into the red target position. If the agent pushes a **blue box** to the target location, it will receive a large penalty.

Colored Sokoban - A Harmful Source Task

- How do the agents trained by causally unaware curriculum generators and by the target task directly without curriculum perform?

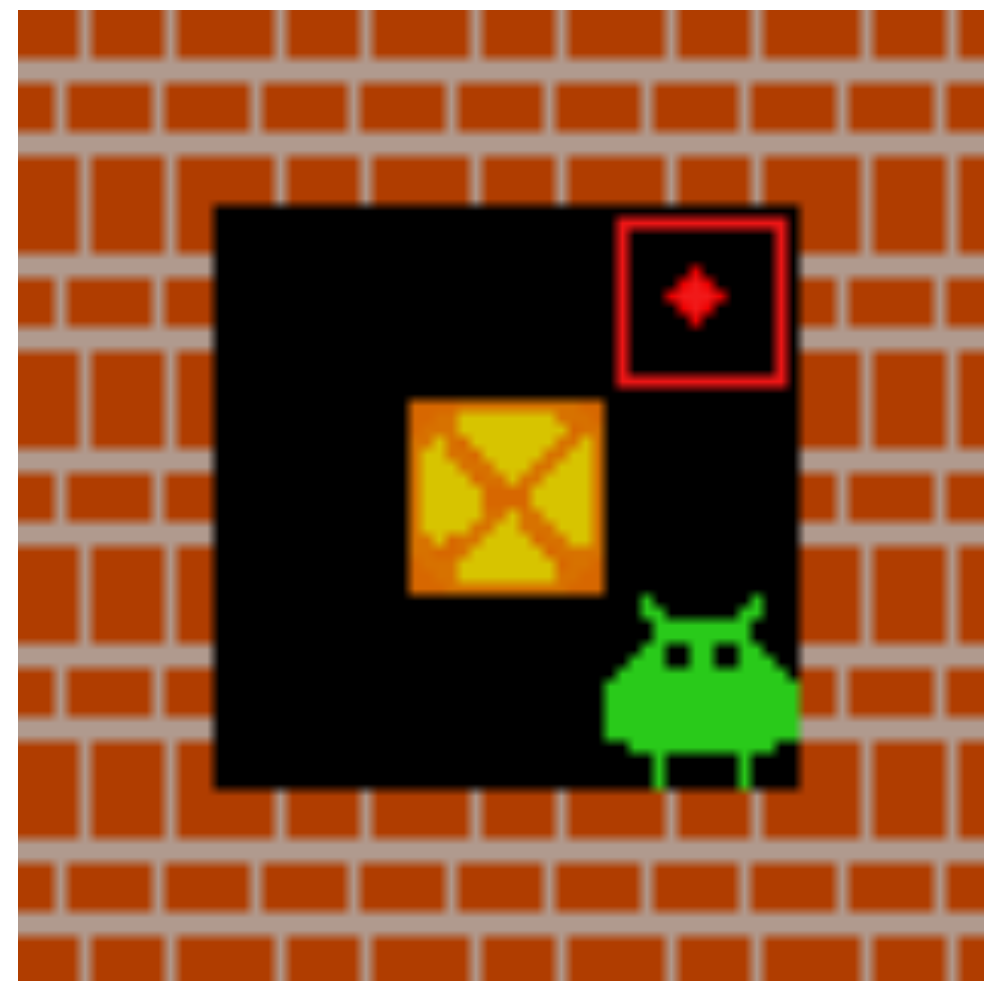


If we were to design a curriculum to help the agent learn, should we modify every possible variable in the causal diagram?

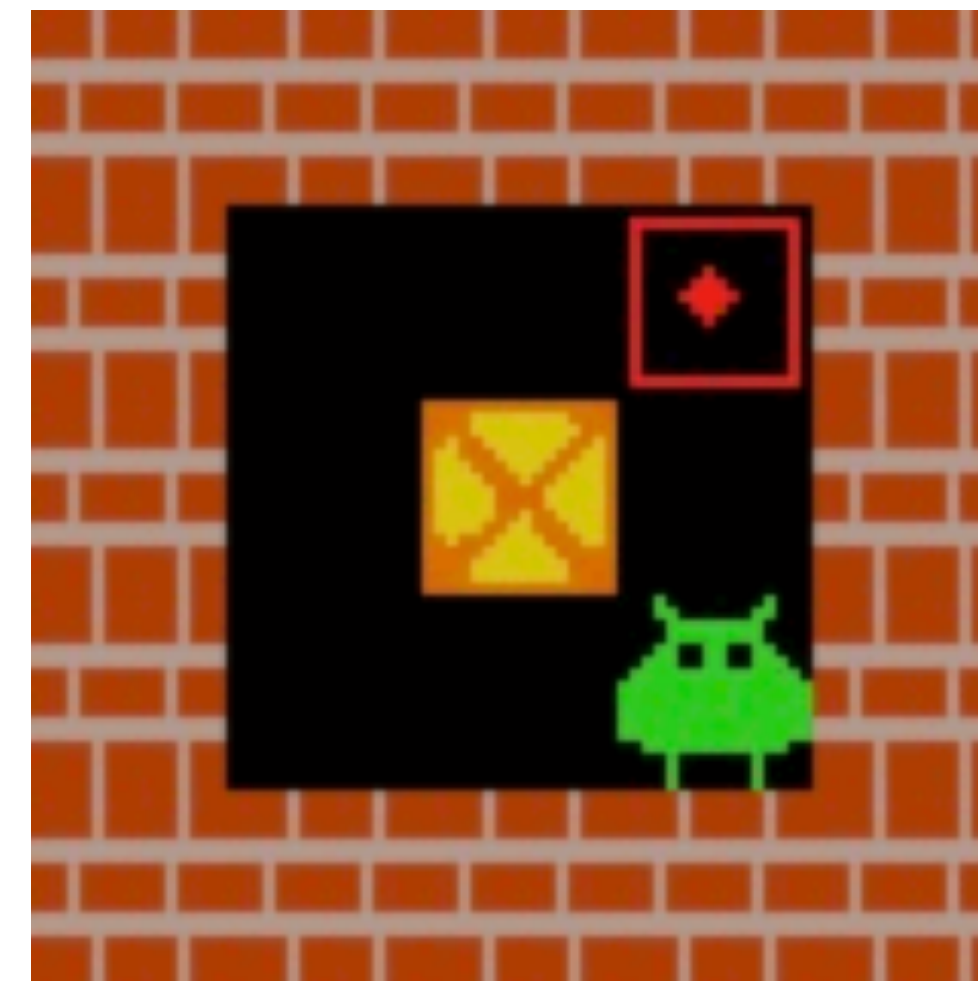
Contributions & Video Demonstration

In this paper, we made the following contributions:

- We derive a sufficient graphical condition determining potentially misaligned source tasks;
- We develop efficient algorithms for detecting misaligned source tasks and constructing source tasks that are guaranteed to align with the target task;
- We introduce a novel augmentation procedure that enables state-of-the-art curriculum learning algorithms to generate aligned curricula to accelerate the agent's learning.



(a) Agent trained by original Currot:
avoid pushing the box



(b) Agent trained by causally augmented Currot:
push the yellow box to the target

Paper Link

Please scan the following for more details on our paper,
“Causally Aligned Curriculum Learning”, ID 6200

