A Surprise Guided Causal Structure and Transfer Learning in Robotic Manipulation Tasks

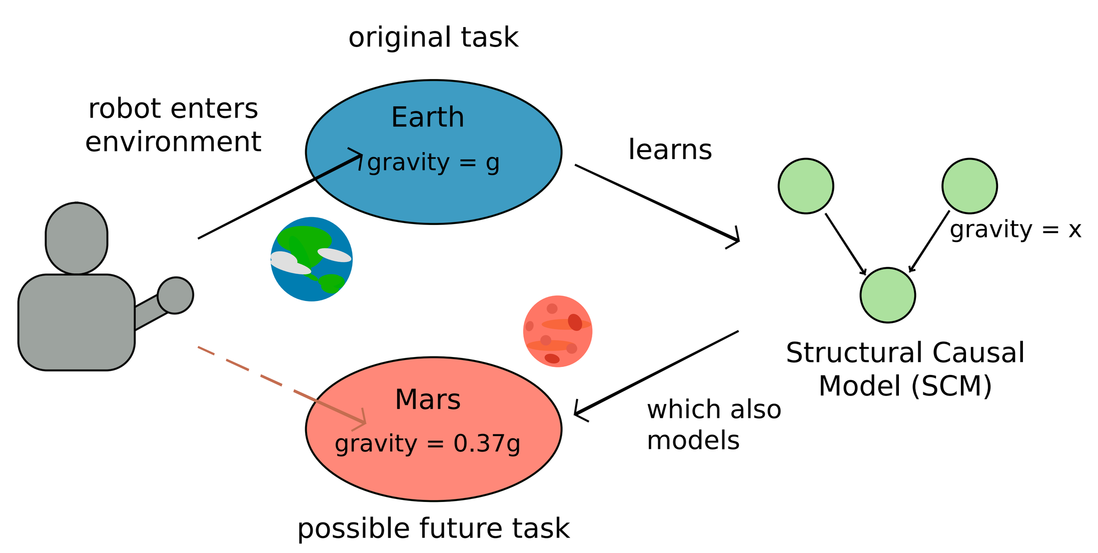
Madelyn Khoury, Dr. Pooyan Jamshidi, & Dr. Forest Agostinelli

University of Virginia

University of South Carolina

Introduction

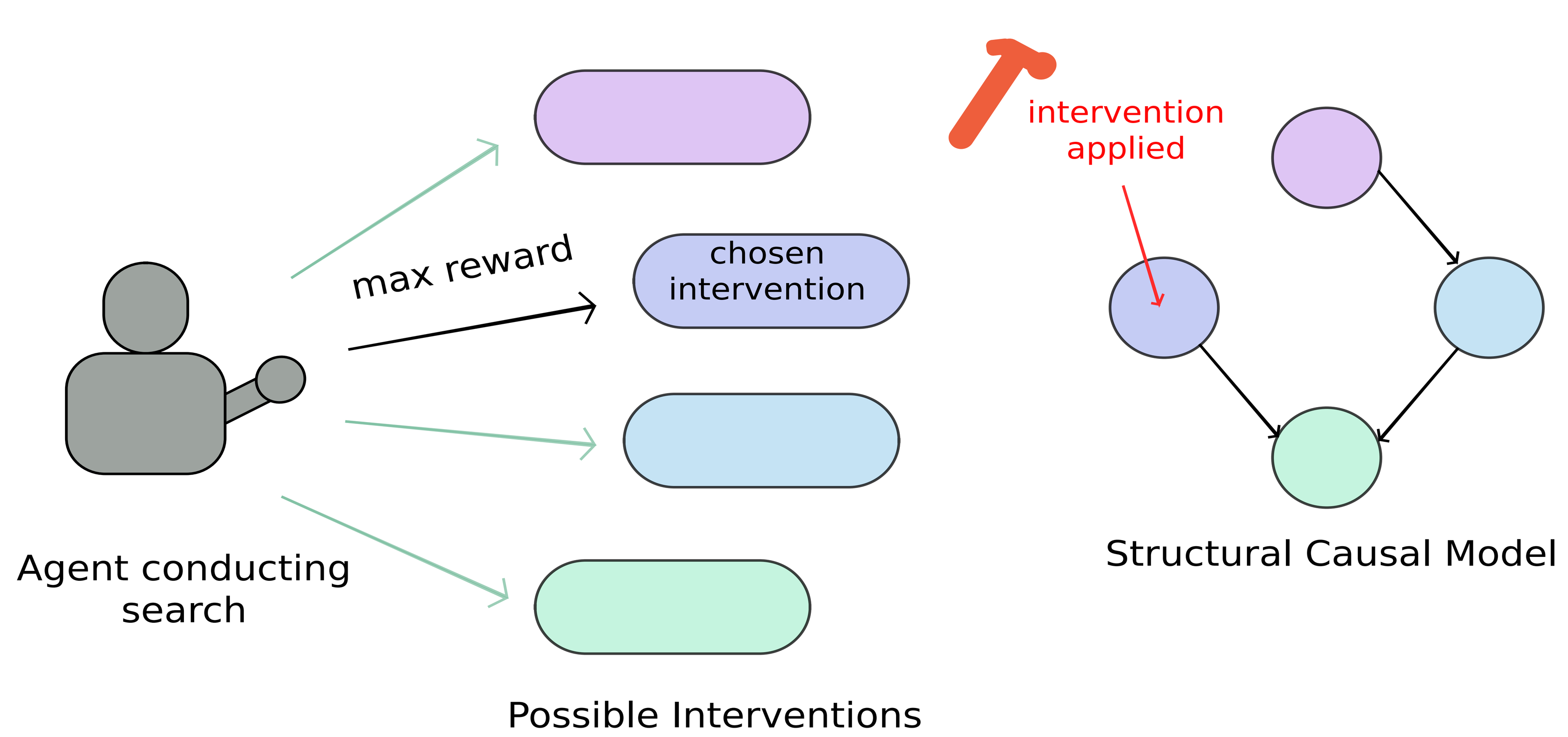
- Unlike humans, robots struggle to transfer the skills they learn while doing one task to other similar tasks
- Learning the causal structure (i.e. cause and effect relationships) of the situation helps with transferring knowledge because most of the causal structure remains the same when moving between similar scenarios, as shown in the example below
- One way to model the causal structure of an environment is with a Structural Causal Model (SCM)



- We want our robot to learn the SCM of its environment as efficiently as possible

Methods

- The agent learns via staging interventions — by changing the value of one variable and observing the effect of the change.
- We implemented a surprise-guided search to help the agent choose the best interventions to take.
- Every time the agent staged an intervention, we recorded the change in entropy of the SCM that occurred due to the intervention — the "surprise" — as the reward of taking that intervention.
- When the agent needed to stage another intervention, it prioritized interventions that had been taken fewer times and that had higher rewards.



Experimental Setup

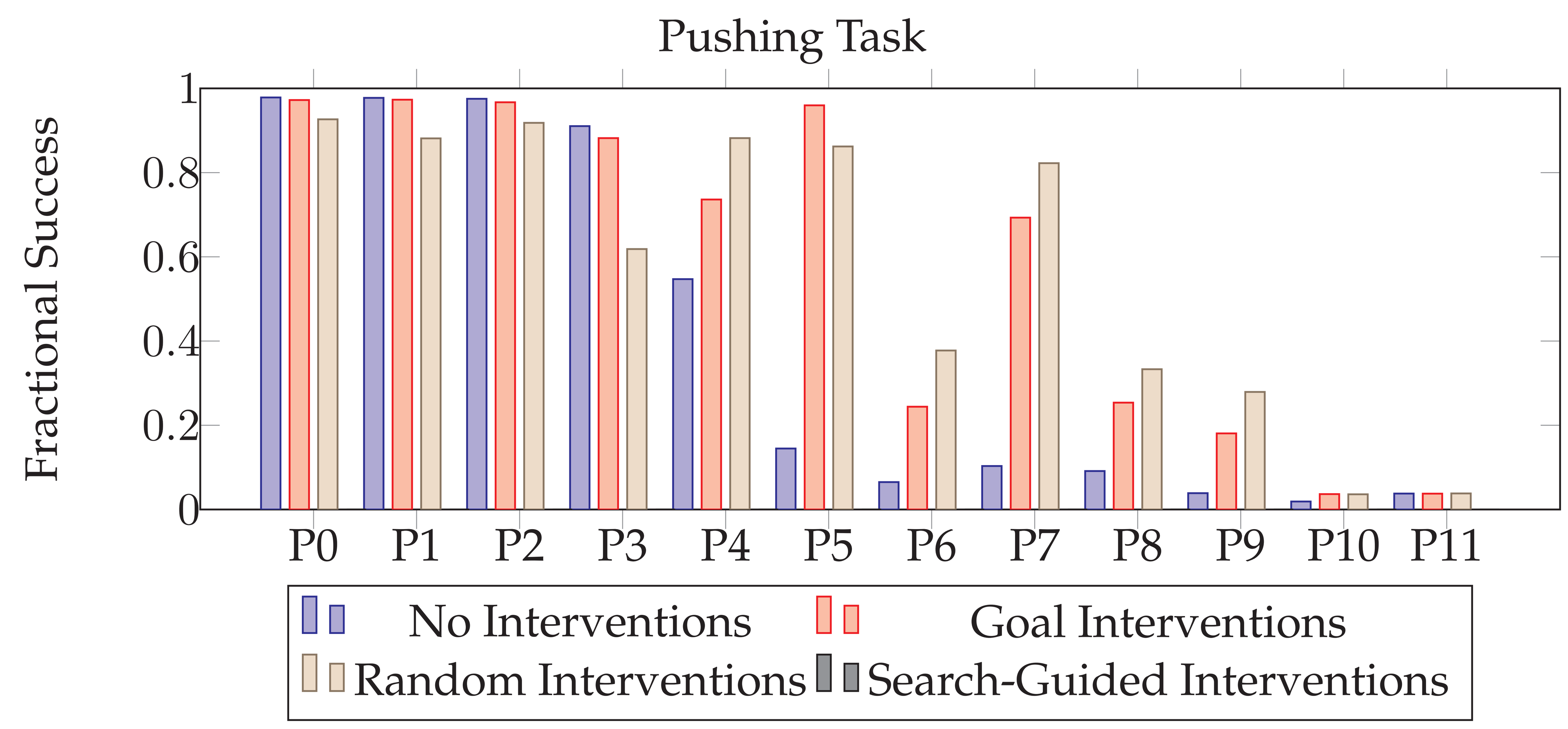
- We modified Ahmed et al.'s CausalWorld environment to allow for interventions on gravity and wind.
- The robot was trained using four curricula and evaluated using two tasks: pushing a block or picking and placing a block into the goal.

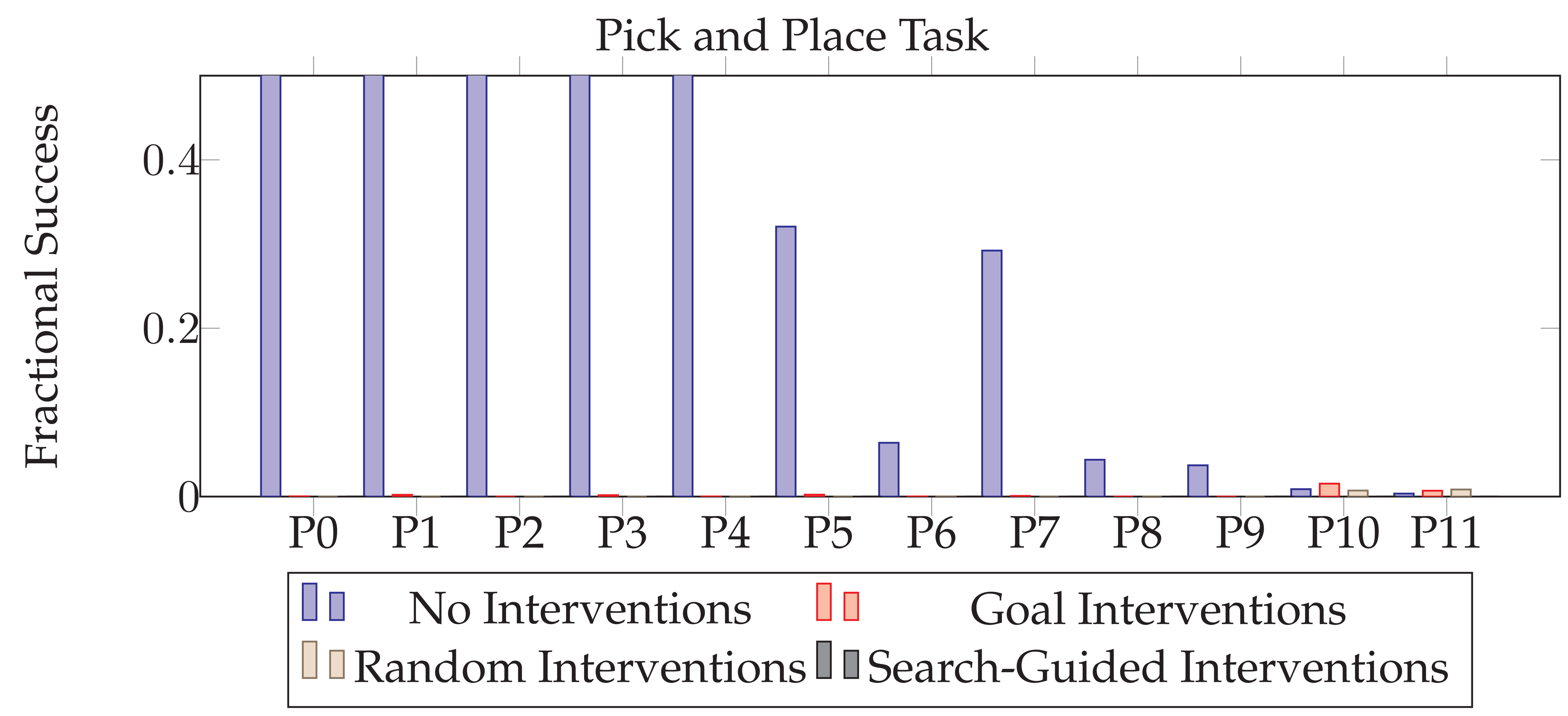
References

[1] Ossama Ahmed, Frederik Träuble, Anirudh Goyal, Alexander Neitz, Manuel Wüthrich, Yoshua Bengio, Bernhard Schölkopf, and Stefan Bauer. Causalworld: A robotic manipulation benchmark for causal structure and transfer learning, 2020.

Results

- Robot success was evaluated in 12 situations, "protocols" P0-P11
- Protocol P0 (which tested the robot on the exact same task as during training) required the least amount of knowledge transfer, while P11 (which tested the robot under different values of every environmental variable) required the most transfer.





- The search intervention actor took longer to complete one intervention than the other actors
- The robot is most successful when completing the easier pushing task in situations where not much has changed from the training task
- When trained on curricula involving interventions, the robot transfers its learning better in the pushing task, but not the pick and place task

Conclusion

- We expect that agents trained on the Search Curriculum will have higher fractional success on both tasks and approach high success rates when trained for fewer time steps.
- However, more tests are needed to draw significant results, as we have not yet observed quantitative data for the search intervention actor.
- We plan to continue this project by predicting rewards before staging interventions; since the benefit of taking an intervention decreases after it is staged, a predicted future reward is more accurate than the mean reward already observed for an intervention. This will ensure that the best interventions are staged.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Nos. 2007202 and 2050896. Additionally, thank you to Austin Loudermilk, Rohit Rajagopalan, and Abir Hossen.