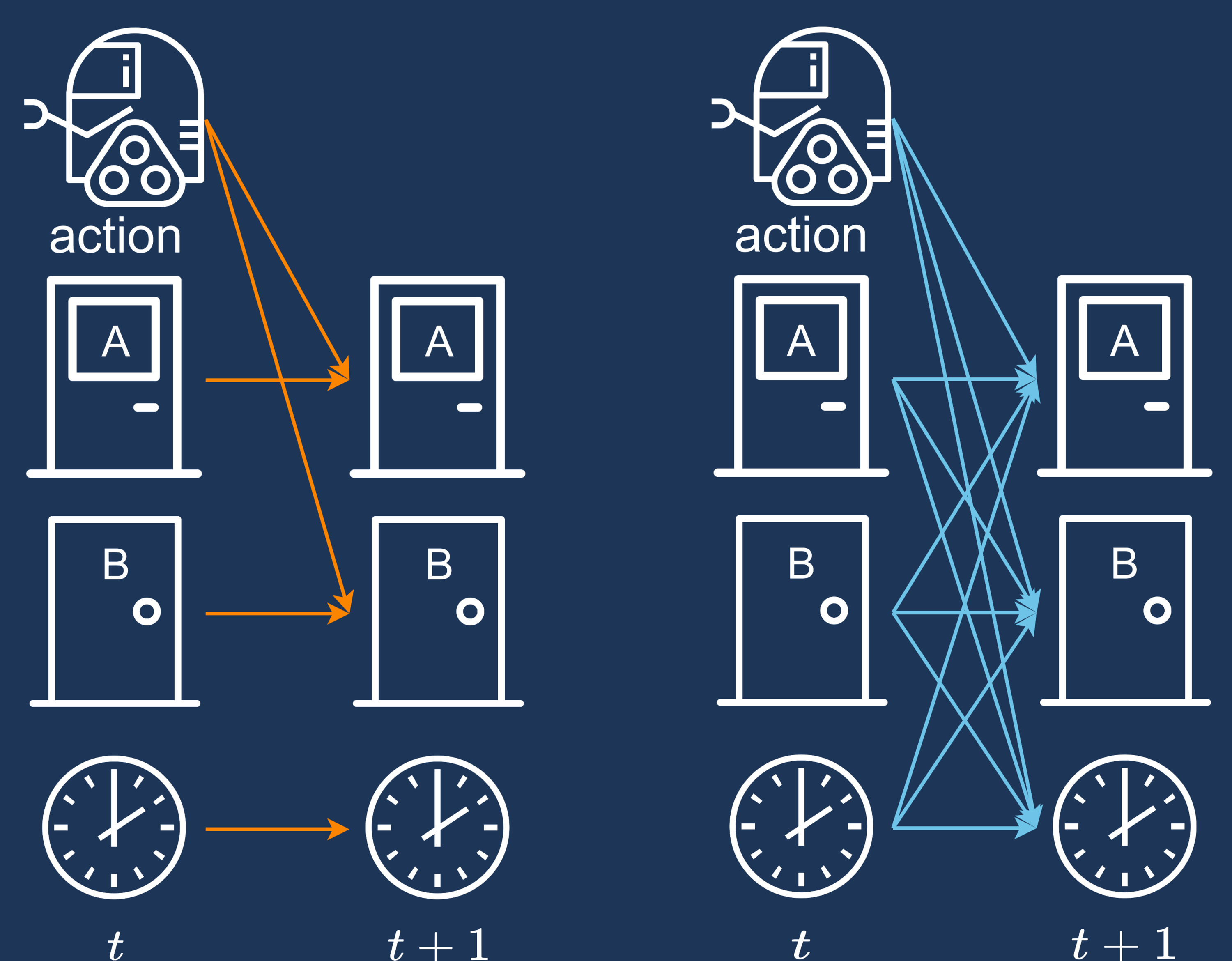
Although real-world dynamics are often sparse, most model-based RL still learns dense dynamics.

In the real world, the next step value of each state variable often only depends on a few current state variables.

But dense models depend the dynamics of each state variable on all variables and the action, making them vulnerable to spurious correlations.



Sparse dynamics models not only generalize better than dense ones, but also enable a state abstraction that is task (reward)-independent.

Causal Dynamics Learning for Task-Independent State Abstraction

Zizhao Wang, Xuesu Xiao, Zifan Xu, Yuke Zhu, and Peter Stone

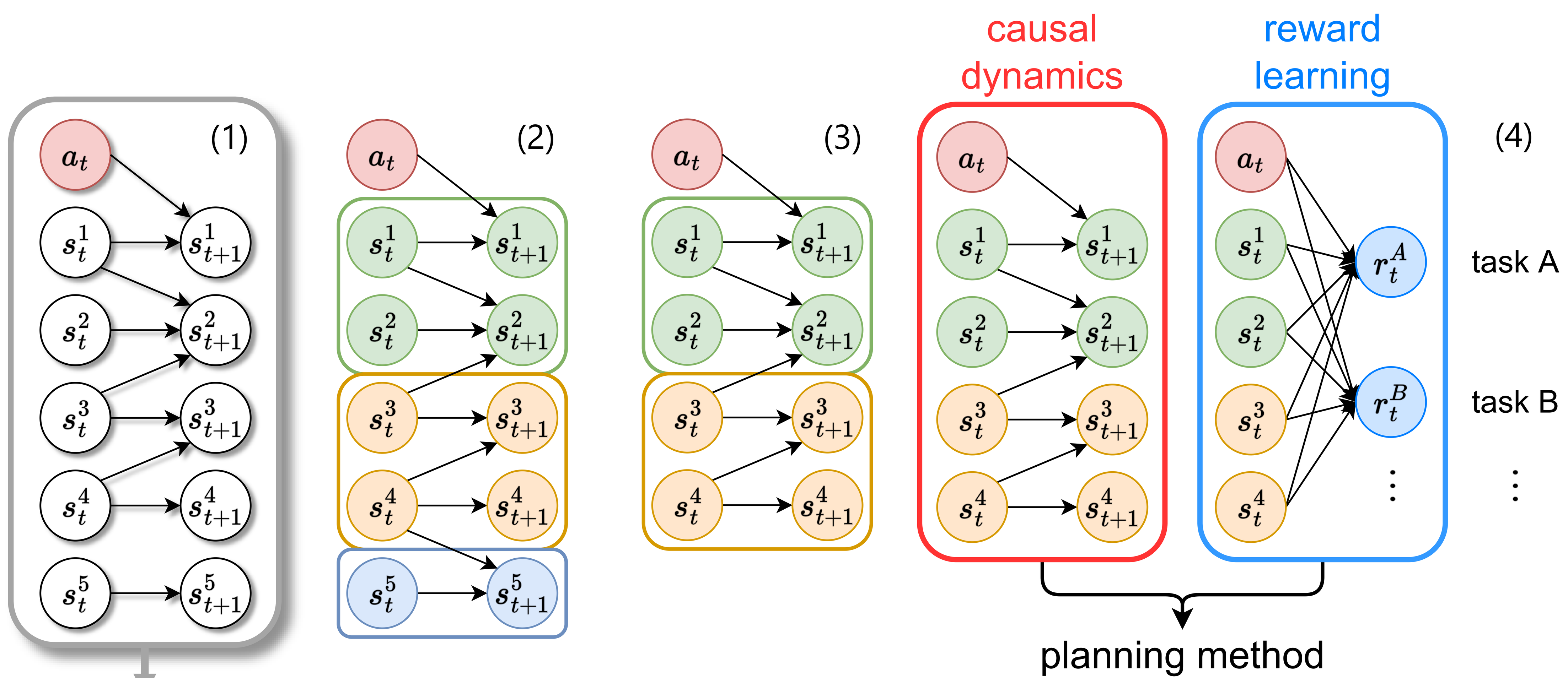
LARG Learning Agents Research Group, The University of Texas at Austin

RPL UT Robot Perception & Learning Lab

Method

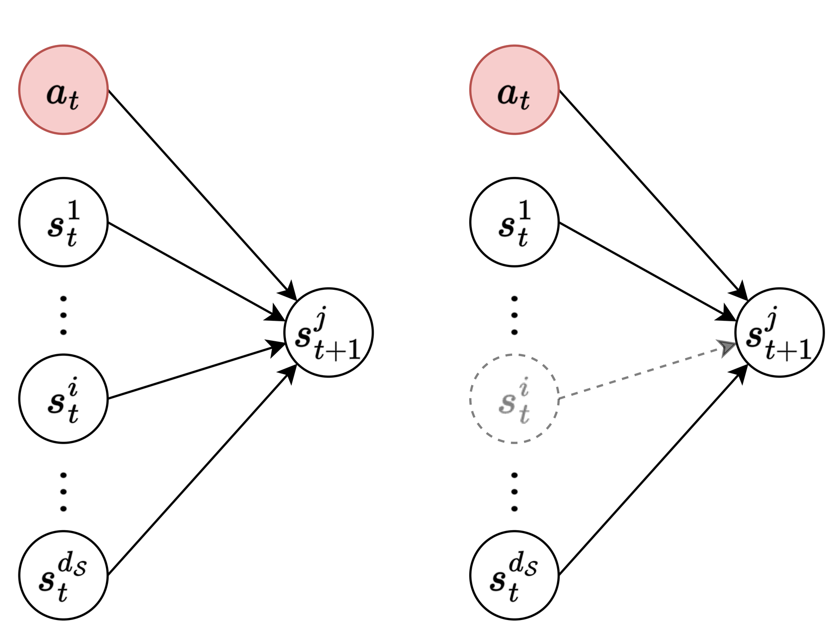
Given a high-level state space and action space, our method:

1. learns a causal dynamics model from transition data
2. splits state variables into three types:
   - controllable: can be affected by the action
   - action-relevant: can't be affected by the action, but affect action's results
   - action-irrelevant: all other variables
3. derives a state abstraction by ignoring action-irrelevant variables
4. uses the abstracted causal dynamics to learn (many) downstream tasks

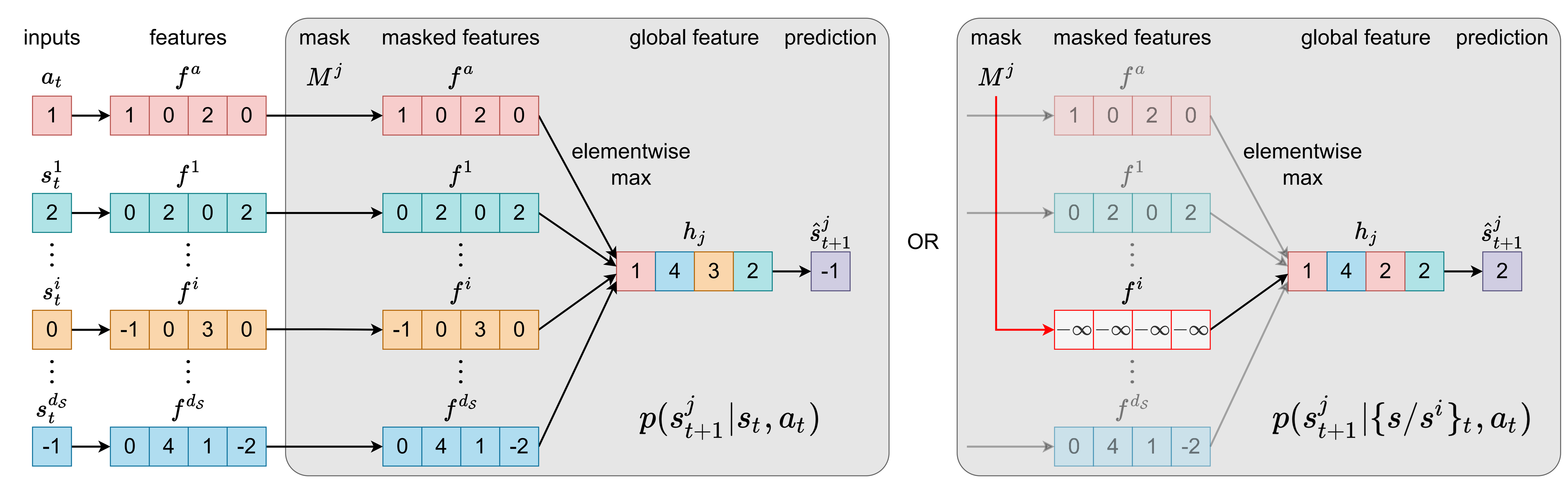


In step 1, to determine if $s_t^i \to s_{t+1}^j$ between each state variable pair, we check if s_t^i is needed to predict s_{t+1}^j.

$$p(s_{t+1}^j | s_t, a_t) \overset{?}{=} p(s_{t+1}^j | \{s_t / s_t^i, a_t\})$$

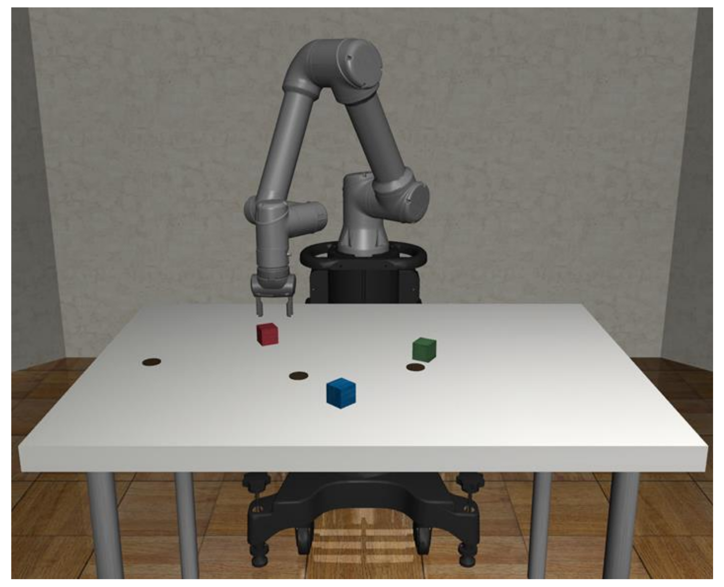


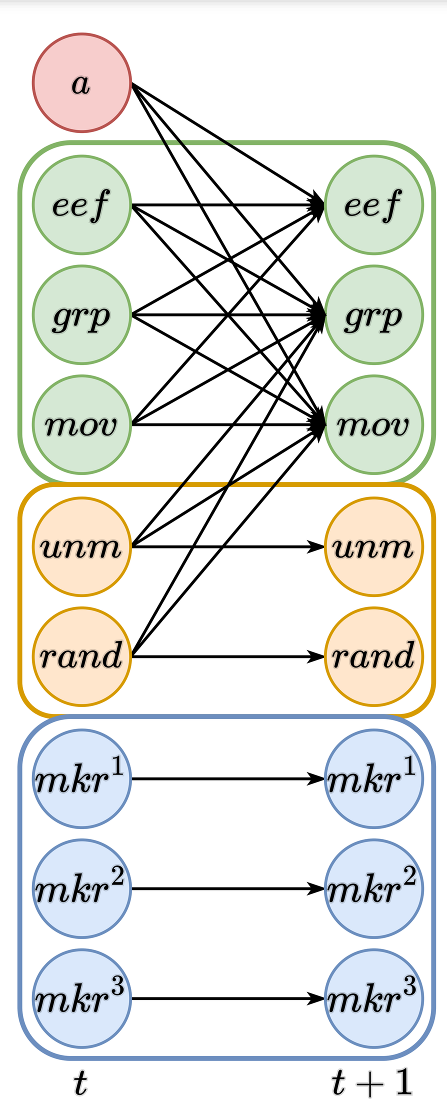
Learning $p(s_{t+1}^j | s_t, a_t)$ & $p(s_{t+1}^j | \{s/s^i\}_t, a_t)$ for every i, j pair needs to train many models...
Instead, our novel architecture can represent all models of $p(s_{t+1}^j | \cdot)$ in one network.



Results

In a synthesized and a table-top manipulation environment, our method learns the correct causal graphs.





Our method generalizes better than dense dynamics for dynamics and policy learning.