Predict the Contact That Changes the Decision

Matched action branches for ranking robot recovery moves

Abstract

Robot world models may predict plausible futures while ranking near-identical recovery actions incorrectly. We propose to train a small contact-event critic on action branches replayed from the same simulator state. The critic will rank recovery actions by the likelihood of task-relevant events such as clearing a jam or losing a grasp. The study will measure whether paired supervision improves downstream decisions at fixed model capacity and rollout cost, including under novel object geometry.

Keywords / world models, robot manipulation, counterfactual actions, contact prediction

Problem and prior work

In tabletop manipulation, two nearby actions can place the end effector on opposite sides of a contact transition. Mean video fidelity can hide this distinction. DreamTrue [1] improves action-faithful prediction using counterfactual post-training and human-annotated defects. What-If World [2] evaluates causal sensitivity using paired interventions. Here we use observed branch outcomes at contact boundaries as supervision, then ask whether gains in branch ranking translate to closed-loop recovery.

Hypothesis and method

When a nominal skill stalls, a fixed proposer emits short candidate action chunks, such as a lateral retreat or a small reorientation. An RGB-D and proprioception encoder, shared across all learned critics, scores predicted contact milestones for each action. Fork the simulator before contact and run each candidate with shared initial state and nuisance seed. Train with pointwise event prediction and a pairwise ranking loss on candidates with different measured outcomes. Ignore tied pairs in the ranking term and retain them for event calibration. A receding-horizon selector executes the first portion of the chosen action, then observes and replans.

What the pairing isolates

Train a pointwise control on exactly the same branch trajectories. For a second control, spend the same number of simulator steps collecting independent states and actions. These comparisons separate the effect of the ranking objective from the effect of targeted data collection. Privileged contact labels are available during training and evaluation, while runtime uses observation and candidate actions.

Validation plan

Experimental design

In ManiSkill3 [3], define three recoveries for jammed insertion, partially missed handle grasp, and a blocked tabletop push. A starting data budget is 2,000 fork states with four candidates per state. Split by object geometry and contact configuration before collecting branches. Test on 300 held-out starts, varying friction and controller noise. Compare behavior of the nominal skill, a pointwise critic trained on independent states, a pointwise critic trained on paired data, a video-prediction scorer trained from the same observations if feasible, and a simulator oracle. Match candidate sets, decision frequency, and rollout horizon. Use a fixed wall-clock ceiling for runtime comparisons.

Outcomes and rejection criteria

Primary endpoints are closed-loop recovery success and same-state action regret, measured against the best branch observed under repeated, seeded rollouts. Report contact-event calibration, collisions and latency. Use paired intervals clustered by start scene and report geometry and friction shifts separately. If ranking improves but closed-loop success does not, inspect whether the chosen actions exhaust the proposal library or whether successive interventions invalidate the short-horizon critic. Treat a zero recovery gain as a negative control result for the central hypothesis.

Feasibility and deliverables

Weeks 1-2 implement state replay and event labels. Weeks 3-4 train compact critics. Weeks 5-6 run held-out closed-loop evaluation. If identical state replays show contact noise, estimate success with repeated branch executions and fit to win probabilities. Release fork-state manifests, contact labels, and decision evaluation code. Scope the initial study to simulated rigid objects.

References

[1] DreamTrue: Action-Faithful Robot World Model with Counterfactual Post-Training. arXiv:2610.12468.

[2] What-If World: A Causal Benchmark for General World Models in Embodied Scenarios. arXiv:2605.27589.

[3] ManiSkill3: GPU Parallelized Robotics Simulation and Rendering for Generalizable Embodied AI. arXiv:2410.00425.