BECAUSE: Bilinear Causal Representation for Generalizable Offline MBRL

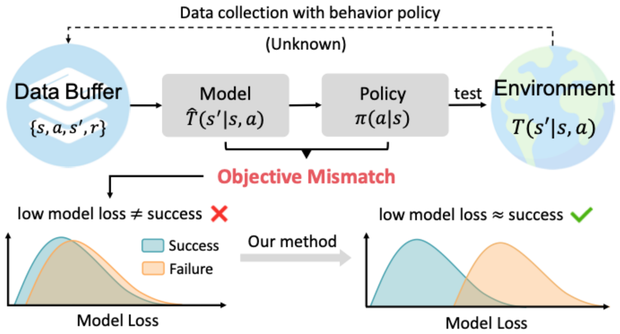
Haohong Lin, Wenhao Ding, Jian Chen, Laixi Shi, Jiacheng Zhu, Bo Li, Ding Zhao

NEURAL INFORMATION PROCESSING SYSTEMS · Caltech · Carnegie Mellon University · THE UNIVERSITY OF CHICAGO · Massachusetts Institute of Technology

Motivation

Objective mismatch between **dynamics** and **policy**

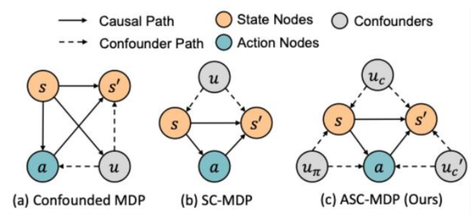
- Imperfect behavior policies π_β
- Intrinsic Uncertainties in the Environment Transitions T
- Both can be modeled as **Spurious Correlations**!



Problem Formulation

Confounded MDP: $[\boldsymbol{u_\pi}, \boldsymbol{u_c}]$. **Properties:** $s' \perp u_\pi \mid (s, a, u_c)$.

- Policy confounder $\pi_\beta(a|s; \boldsymbol{u_\pi})$.
- Dynamics confounder $T(s'|s, a; \boldsymbol{u_c})$.



Bilinear MDP

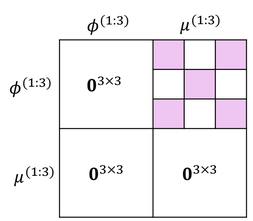
- Feature ϕ, μ
- Core Matrix M

$$T(s'|s,a) = \phi(s,a)^T M(u_c)\, \mu(s')$$

State-action Representation · Core Matrix · Next State Representation

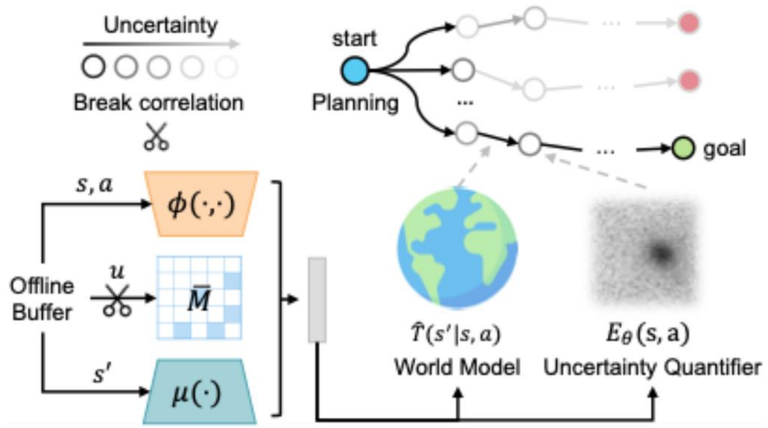
Bilinear MDP is a Bipartite Structured Causal Model

- Causal nodes with d concepts:
  - $\{\phi(s,a)^{(i)}\}_{i=1}^{d}, \{\mu(s')^{(j)}\}_{j=1}^{d}$
- Causal edges between
  - M_{ij}: $\phi(s,a)^{(i)} \to \mu(s')^{(j)}$



Methodology

BECAUSE: Bilinear causal structure helps mitigate objective mismatch in MBRL!



What is BECAUSE?

- Concept ϕ, μ
- De-confounded M
- Enforce Sparsity
- Learn a world model that alleviate mismatch

How to learn BECAUSE?

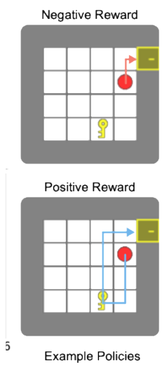
- Transition confounder u_c:
  - Estimate sparse $\widehat{M}(u_c)$
- Policy confounder u_π
  - Reweighting on M : $\widehat{M}(u_c) \to \bar{M}$

How to use BECAUSE?

- **Planning (w/ MPC):**
  - Learned model $\hat{T}(s'|s,a) = \hat{\phi}(s,a)\bar{M}\,\hat{\mu}(s')$
- **Planning under uncertainty:**
  - Uncertainty Quantifier $\Gamma_\theta(s,a)$

  EBM: $\Gamma_\theta(s,a) = E_\theta(\phi(s,a), M)$

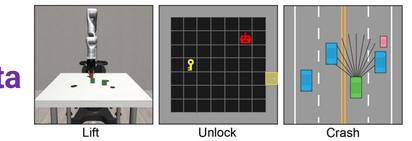
  Adjust Return $r(s,a) - \Gamma_\theta(s,a)$

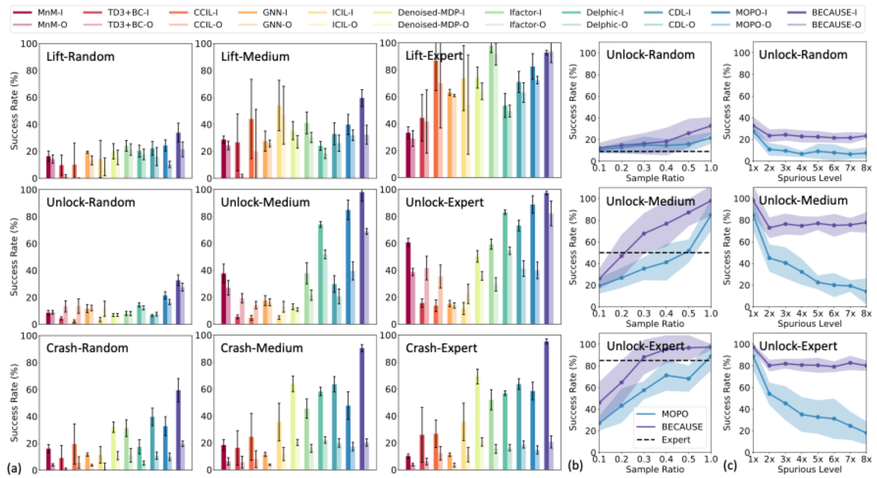


Experiments and Analysis

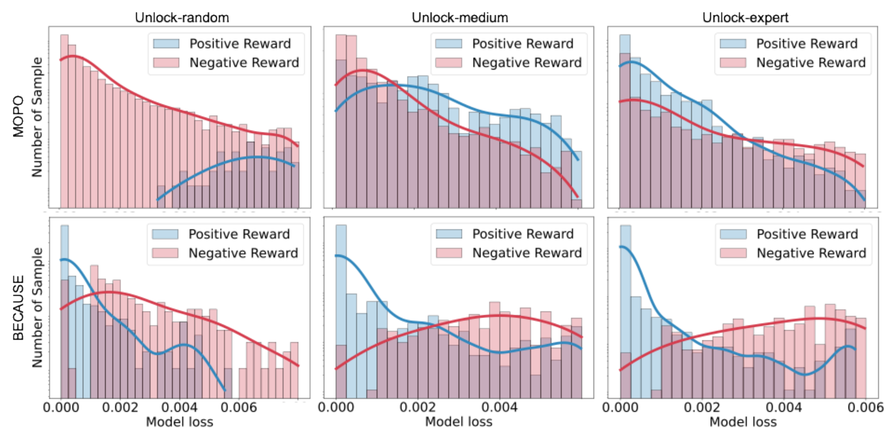
Q1. Online generalization

Q2. Robustness to batch data scales and spurious levels





Q3. Mitigation of Objective Mismatch



External Links

Paper Preprint · Project Page · Personal Homepage