Temporal Logic Guidance for Action-Only Diffusion Policies with World Models

Moritz Zoellner, Anastasios Manganaris, and Rohan Paleja
Department of Computer Science, Purdue University
{zoellner, amangana, rpaleja}@purdue.edu

Motivation

Diffusion policies [1] enable rich, multimodal robot behavior, but offer limited control at deployment time. In human-centric environments, it is a necessity to **adapt behavior according to user preferences**. While the desired behaviors may already exist within the learned distribution, policies are not able to reliably select these modes. We aim to enable inference-time steering of robot behavior based on **Signal Temporal Logic (STL) [2]** specifications.

Signal Temporal Logic

STL is a formal language to specify **constraints over trajectories**. It extends propositional logic with temporal operators such as *Always* (**G**) and *Eventually* (**F**), and evaluates specifications using a **continuous robustness score**. Example:

`F(transport_can) & G(upright_can)`,
where `ρ(upright_can) = 5° - θ`$_{can}$

Related Work

Universal Guidance (Bansal et al., 2023) [3]
→ Gradient-based diffusion guidance based on external signals

LTL-DOG (Feng et al., 2024) [4]
→ Temporal logic guidance within diffusion denoising using a policy that outputs both actions and states

STLCG++ (Kapoor et al., 2025) [5]
→ Efficient differentiable STL robustness for gradient optimization

Method

Human Intent **"Don't spill the drink"** $\xrightarrow[\text{as STL}]{\text{Expressed}}$ $\phi = \mathbf{G}(\theta_{tilt} < 5°)$

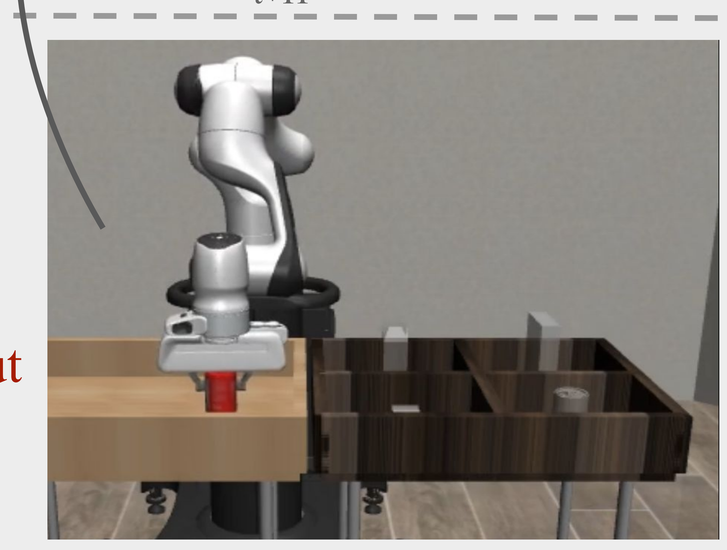
(Our Contribution)
Inference-Time Guidance

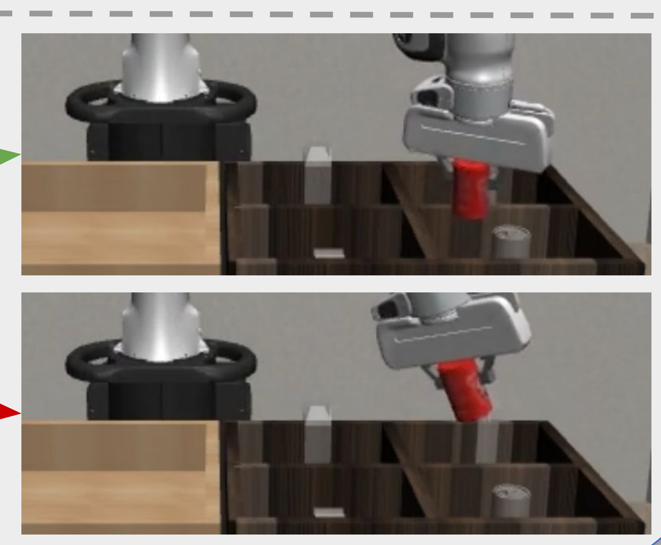
1. World Model Rollout
2. STL-based Robustness
3. Diffusion Steering Update

a_t, a_{t+1}, … s_t, s_{t+1}, s_{t+2}, …, s_{t+H}

$$J = \rho_\phi(\hat{s}_{t+1:t+H}) \quad a^{k-1} \leftarrow a^k + \lambda \nabla_a J$$

Diffusion Policy Rollout



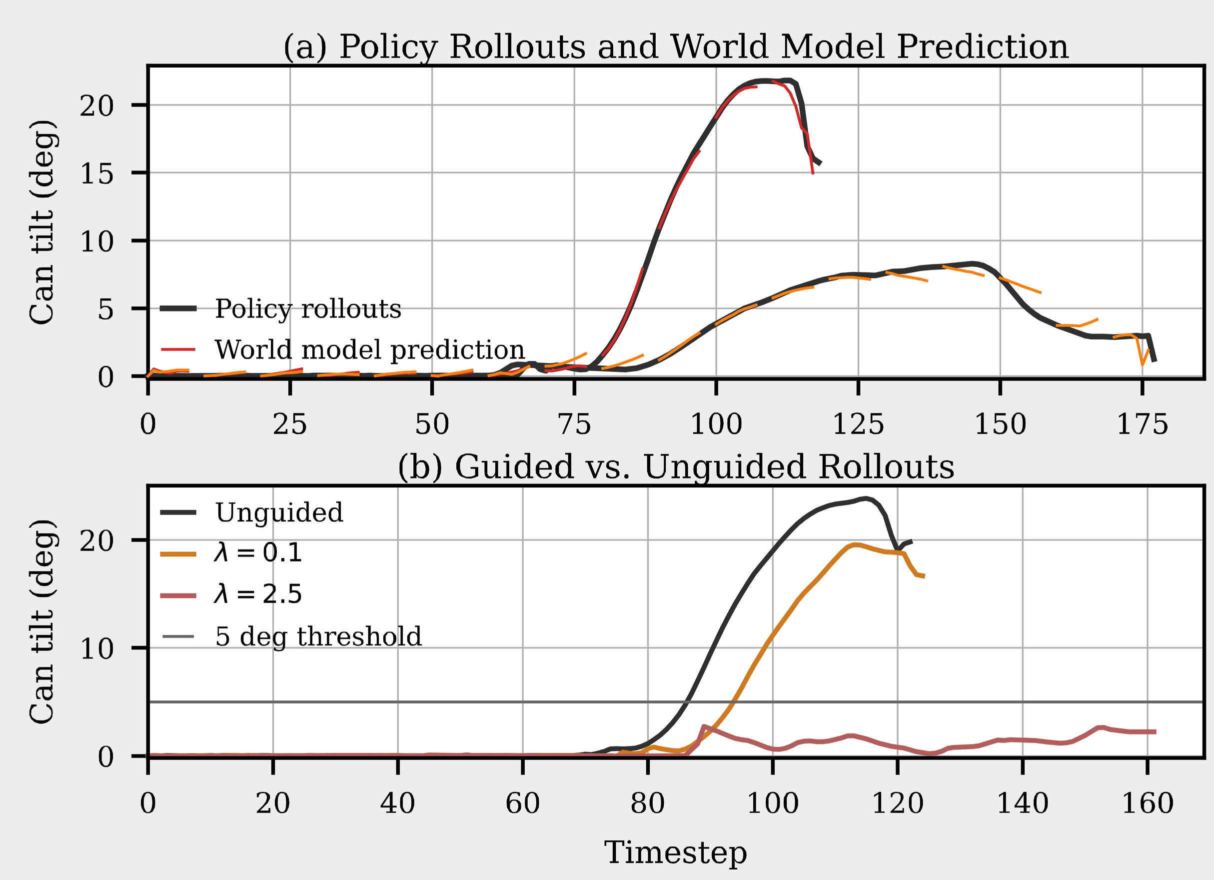


Conclusion

We enable *STL-based inference-time steering of diffusion policies*. However, current STL guidance methods face a **horizon mismatch**: policies predict **short-horizon action chunks**, while specifications often require **long-horizon evaluation**. To enable arbitrarily complex temporal tasks, we aim to use a **high-level world model** that predicts how actions change the atomic propositions, rather than future states.

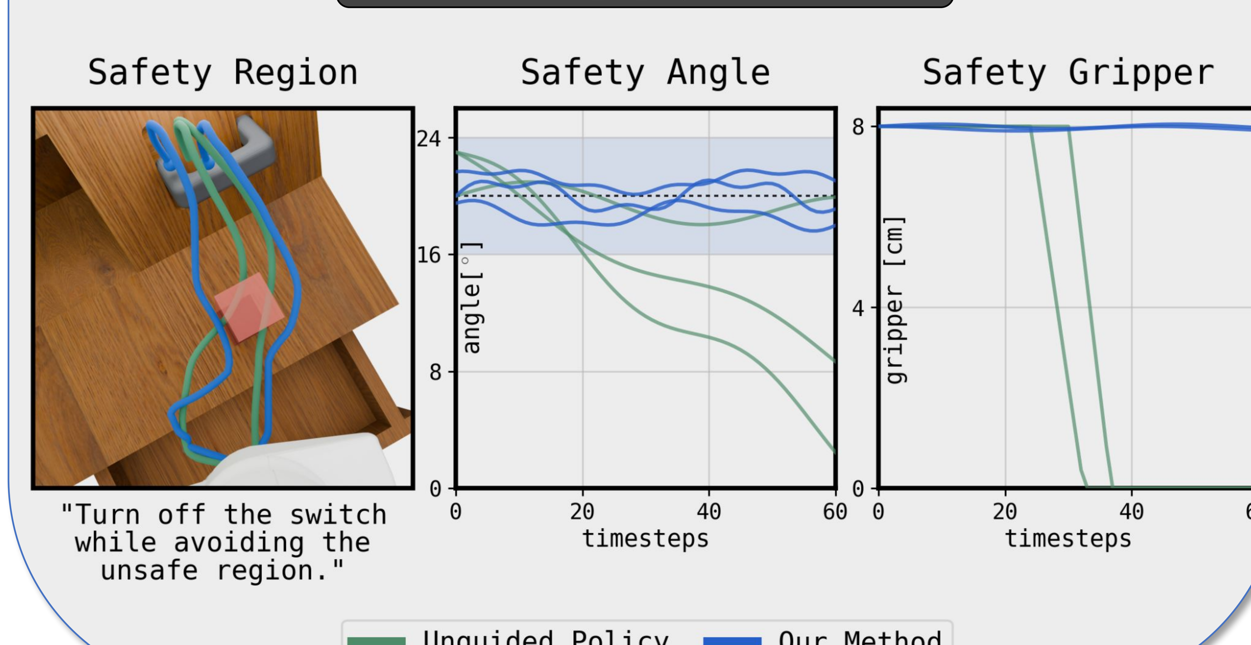
Results

Robomimic [6] (*Can Transport*)



Method	Avg. Tilt (°) ↓	Succ. (%) ↑	Viol. (%) ↓
Base Policy	8.51	100.0	84.0
Sample & Rank	8.42	98.0	82.0
Guidance (Ours)	**1.93**	**100.0**	**4.0**

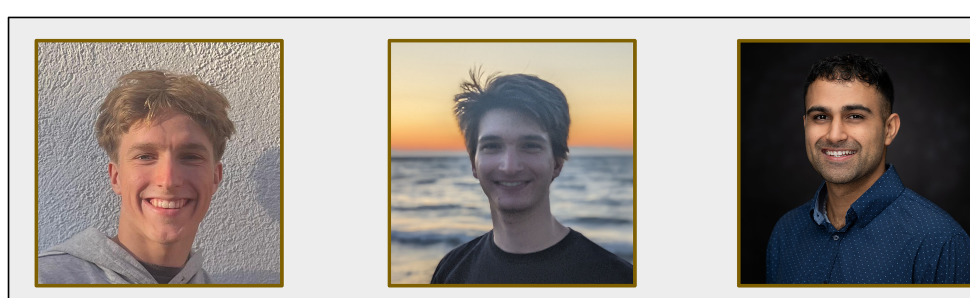
CALVIN [7] (*Switch on*)



References

[1] C. Chi et al., "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion," The International Journal of Robotics Research, 2025.
[2] O. Maler and D. Nickovic, "Monitoring Temporal Properties of Continuous Signals," in Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems, 2004.
[3] A. Bansal et al., "Universal Guidance for Diffusion Models," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[4] Z. Feng, H. Luan, P. Goyal, and H. Soh, "LTLDoG: Satisfying Temporally-Extended Symbolic Constraints for Safe Diffusion-Based Planning," IEEE Robotics and Automation Letters, 2024.
[5] P. Kapoor, K. Mizuta, E. Kang, and K. Leung, "STLCG++: A Masking Approach for Differentiable Signal Temporal Logic Specification," IEEE Robotics and Automation Letters, 2025.
[6] A. Mandlekar et al., "What Matters in Learning from Offline Human Demonstrations for Robot Manipulation," in Proc. Conference on Robot Learning, 2022.
[7] O. Mees, L. Hermann, E. Rosete-Beas, and W. Burgard, "CALVIN: A Benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks," IEEE Robotics and Automation Letters, 2022.

Authors



Moritz Zoellner   Anastasios Manganaris   Rohan Paleja