Planning with Diffusion for Flexible Behavior Synthesis

Michael Janner* Yilun Du* Joshua B. Tenenbaum Sergey Levine

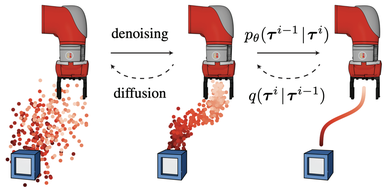
UC Berkeley MIT *equal contribution

diffusion-planning.github.io

Summary

Diffuser is a denoising diffusion model for reinforcement learning tasks that plans by iterative refinement

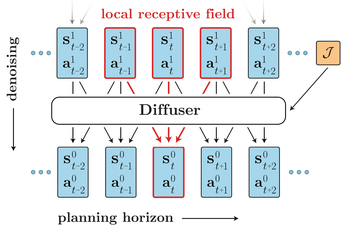
- Trajectories are generated non-autoregressively; all timesteps are sampled simultaneously
- Instead of using a better generative model in an existing planning algorithm, diffusion-based sampling is itself the planning algorithm



Denoising lends itself to **flexible conditioning**:

- maximize reward with classifier-guided sampling
- plan with constraints via image inpainting
- a single diffusion model can be reused for multiple tasks

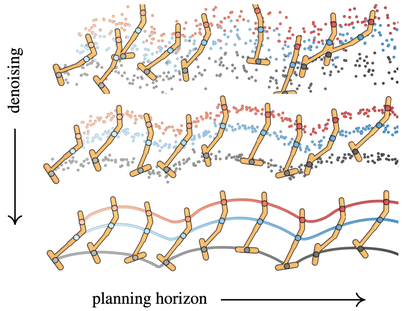
A Generative Model for Trajectory Planning



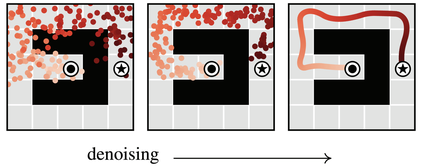
Diffuser is an image diffusion model with temporal convolutions

- a small receptive field constrains the model to only enforce local consistency during a single denoising steps
- by composing many denoising steps together, local consistency drives global coherence of a sampled plan

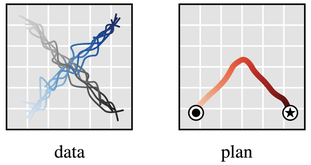
A guide function $\mathcal{J}$ can bias plans toward those optimizing a test-time objective or satisfying a set of constraints



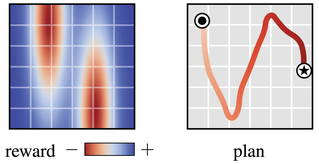
Properties of Diffusion Planners



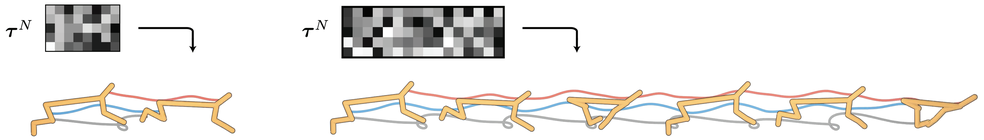
① **Learned long-horizon planning** that does not suffer from the myopic failure modes common to shooting-based planning algorithms



② **Temporal compositionality** for stitching together trajectory snippets to generate novel plans



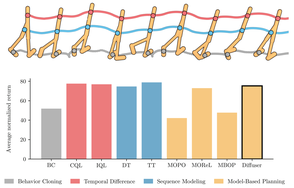
③ **Task compositionality** for conditioning samples on new reward functions unseen during training



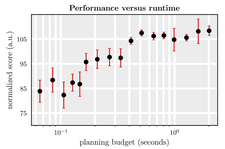
④ **Variable-length planning** by changing the size of the noise initialization in the reverse diffusion process

Experimental Evaluation

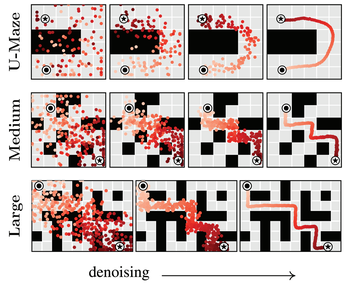
D4RL offline locomotion

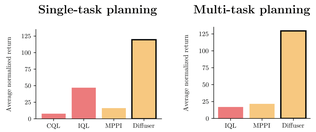
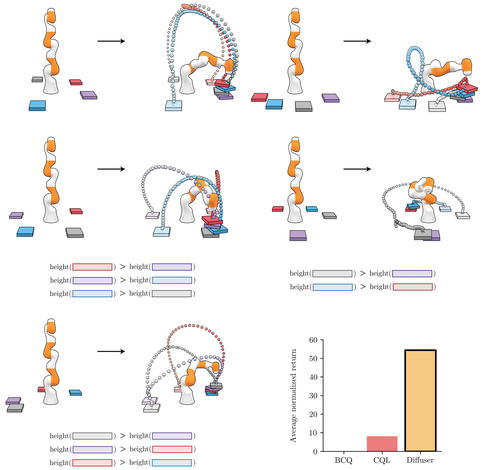


Warm-starting: vary the number of reverse diffusion steps to speed up receding-horizon planning



Maze2D: sparse-rewards & long-horizons





Kuka block stacking: test-time generalization