HEDGE: Hierarchical Event-Driven Generation

Frederik Ebert *, Karl Pertsch *, Oleh Rybkin *,
Chelsea Finn, Dinesh Jayaraman, Sergey Levine

(*equal contribution)

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Penn Engineering | GRASP Laboratory
General Robotics, Automation, Sensing & Perception Lab

USC University of Southern California

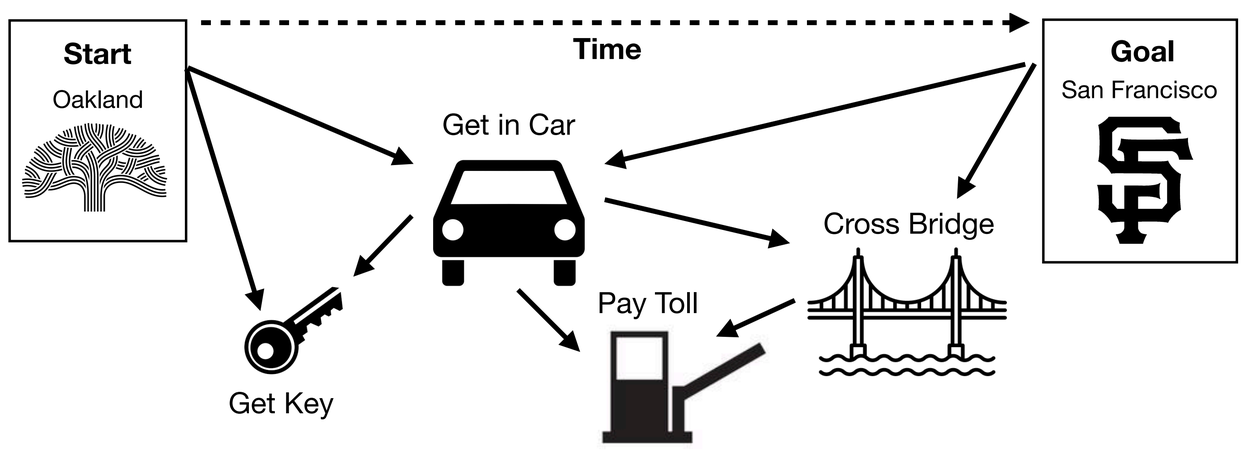
Idea

Task: Modeling the distribution of possible trajectories that lead from the start to the goal.

Motivation: A natural way of predicting for planning tasks is in a time-agnostic manner, starting from the states depicting high-level events, and filling in the finer details later.

Idea: A tree-structured predictive model.

Applications: 80-frame video infilling and planning.

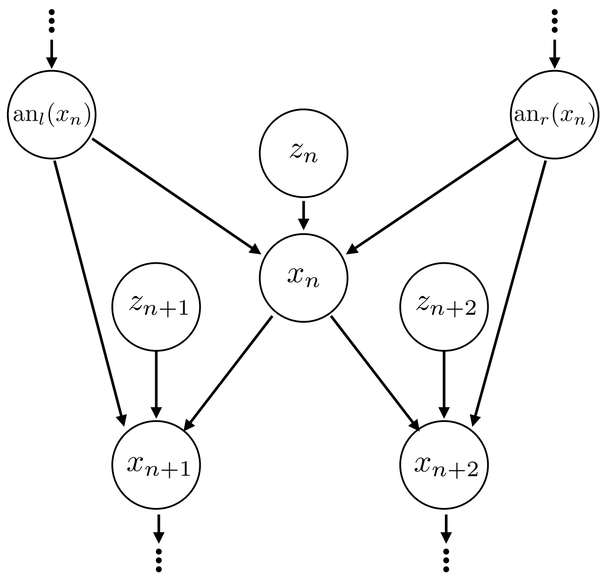


Approach

A Graphical Model of Trees

A **single step** of the model produces one event that lies in between its parent states.

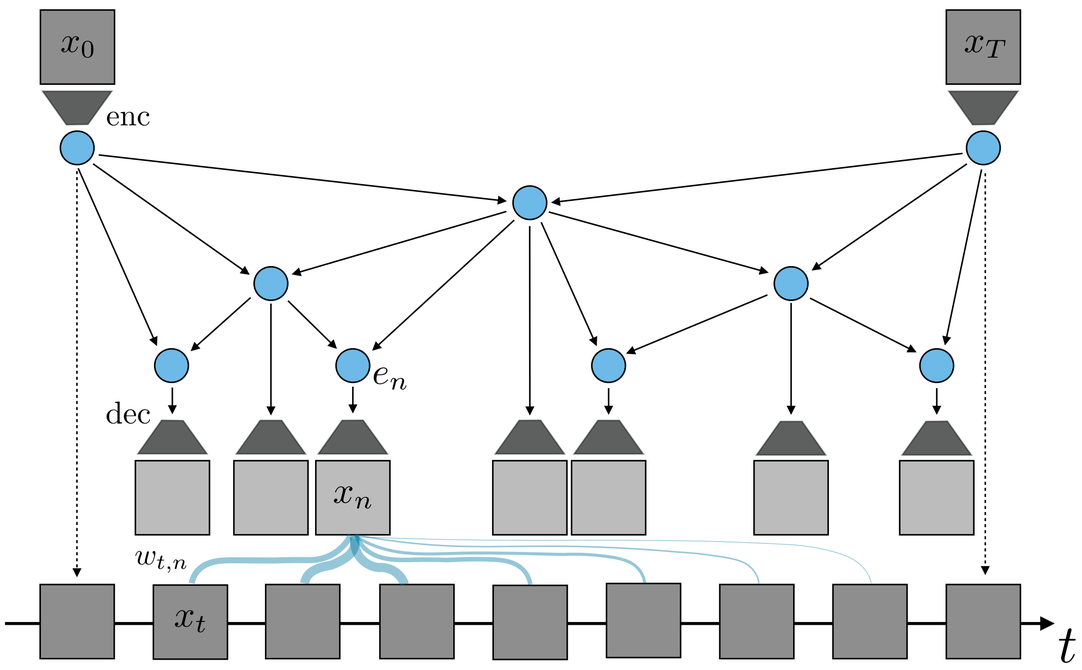
Applying the model **recursively** builds a tree and eventually produces every state between the start and the goal.



$$p(x_{0\sim T}|x_0, x_T) = \prod_{n=0}^{\infty} \mathbb{E}_{p(z_n)} p(x_n|\text{an}(x_n), z_n).$$

Evidence Binding

Problem: How to determine which nodes of the tree match which frames in a time-agnostic model?



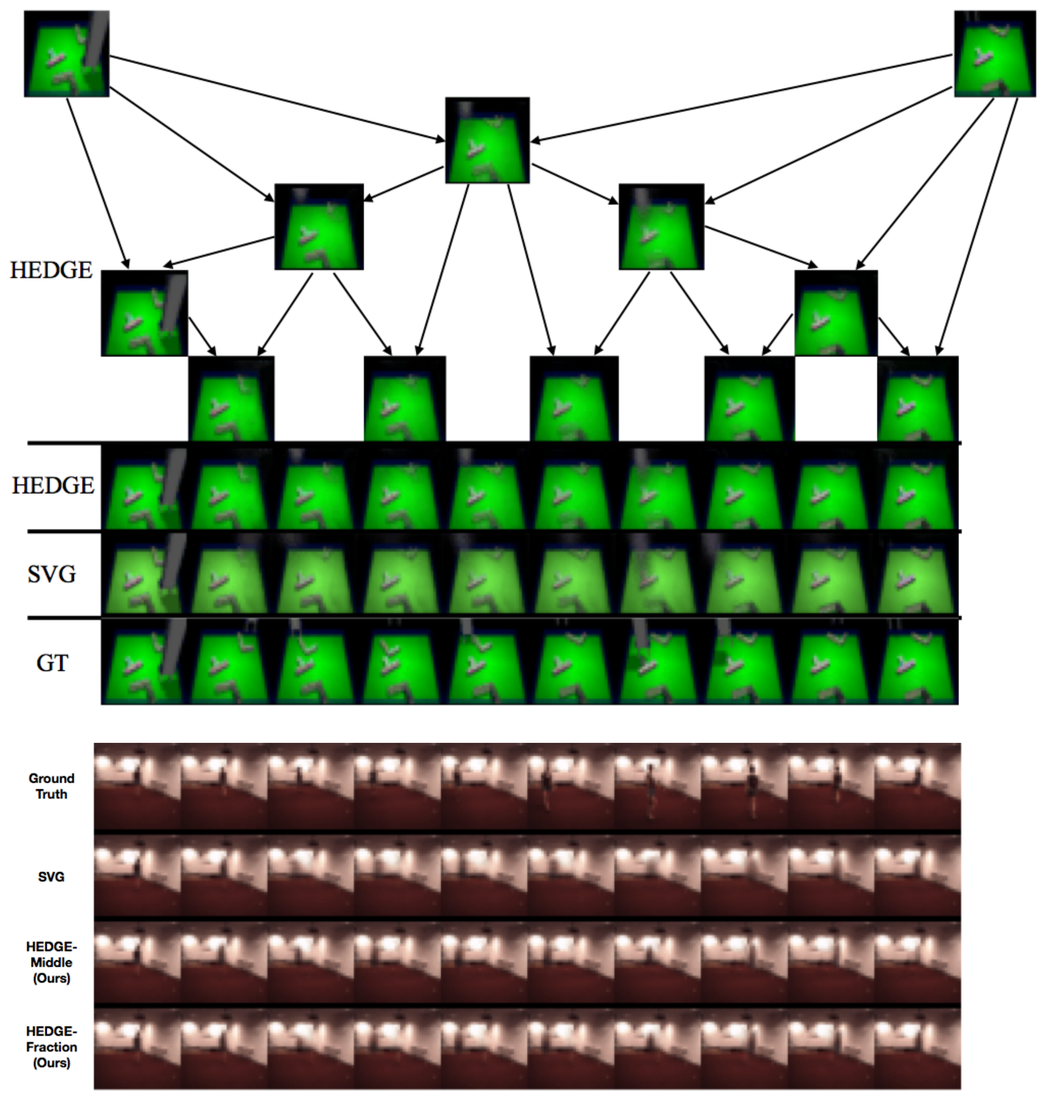
Solution: Introduce a binding distribution p(w | x, z) that selects which frames bind to which nodes.

$$p(x_t|x_0, x_T, z, w) = p(x_t|\bar{x}_{w_t}) p(\bar{x}_{w_t}|x_{0:T}, z)$$

$$\mathcal{L} = \sum_{t=0}^{T} \sum_{n=0}^{N-1} w_{t,n} \, ||x_t - \bar{x}_n||^2 + \beta \sum_{n=0}^{N-1} KL\big[q(z_n|x_{0:T}) \,||\, p(z)\big].$$

Experiments

Long-Term Prediction



How do we evaluate time-agnostic predictions? We use **dynamic time warping** to align to the ground truth, to test if the model predicts the right events regardless of timing.

80-frame video prediction

Dataset	Pushing			Human 3.6M		
Method	PSNR	SSIM	FVD	PSNR	SSIM	FVD
Forward	**22.16** ±1.40	**0.773** ±5e−4	**394**	31.10 ± 3.22	**0.958** ±5e−4	1933
HEDGE-middle (ours)	21.36 ±1.17	0.747 ±2e−4	563	**31.69** ±3.50	**0.958** ±5e−4	**1423**
HEDGE-fraction (ours)	21.26 ± 1.31	0.748 ±2e−4	693	30.11 ± 3.32	0.944 ±2e−4	1497

Control