Learning Plannable Representation with Causal InfoGAN

Thanard Kurutach[1]*, Aviv Tamar[1]*, Ge Yang[2], Stuart Russell[1], Pieter Abbeel[1]

University of California, Berkeley[1]
University of Chicago[2]

Overview

How to extend classical AI planning to image observations?
GAN framework for learning a plannable latent space.
Key idea – structure on dynamics of GAN noise
Visual planning – imagine a goal directed sequence of observations

Motivation

Future robots need to operate inside homes and hospitals alongside humans.
Main challenges:

- long-horizon problems
- unstructured environments
- manipulate objects

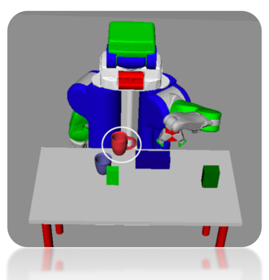
Classical Planning

Pros

- Solve complex planning and **long-horizon** problems efficiently.

Cons

- Need to specify predicates, preconditions, and effects
- Detached from perception



Kim et al, ICRA 2017
Toussaint, IJCAI 2015
Srivastava et al., ICRA 2014

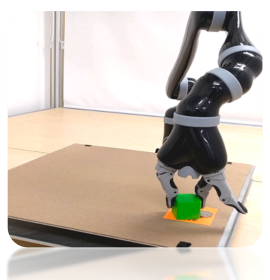
Model-Free Reinforcement Learning

Pros

- Work with **high-dimensional** observations such as images

Cons

- Data demanding
- Difficult to define reward on images
- Task specific



Riedmiller et al., ICML 2018
Gu et al., ICRA 2017
Levine et al., JMLR 2016

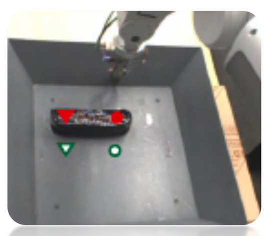
Model-based RL with image inputs

Pros

- Represent system knowledge for solving new tasks in **zero shot**.

Cons

- Currently limited to simple planning, manipulation problems such as object pushing, and shaped reward



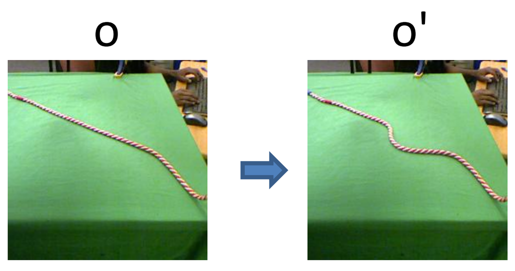
Finn et al, ICRA 2017
Ebert et al., CoRL 2017
Agrawal et al., NIPS 2016
Oh et al., NIPS 2015

Let's come up with a solution that combine the best of three landscapes!

Problem Formulation

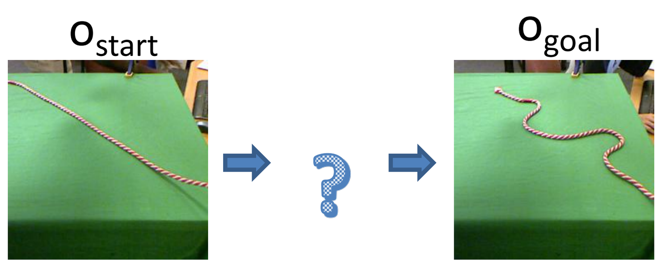
Data: Sequential observations {o, o'} ~ P_{data}

- Captures causality
- Random perturbations



Goal: Given o_{start} and o_{goal} ,generate visual plan

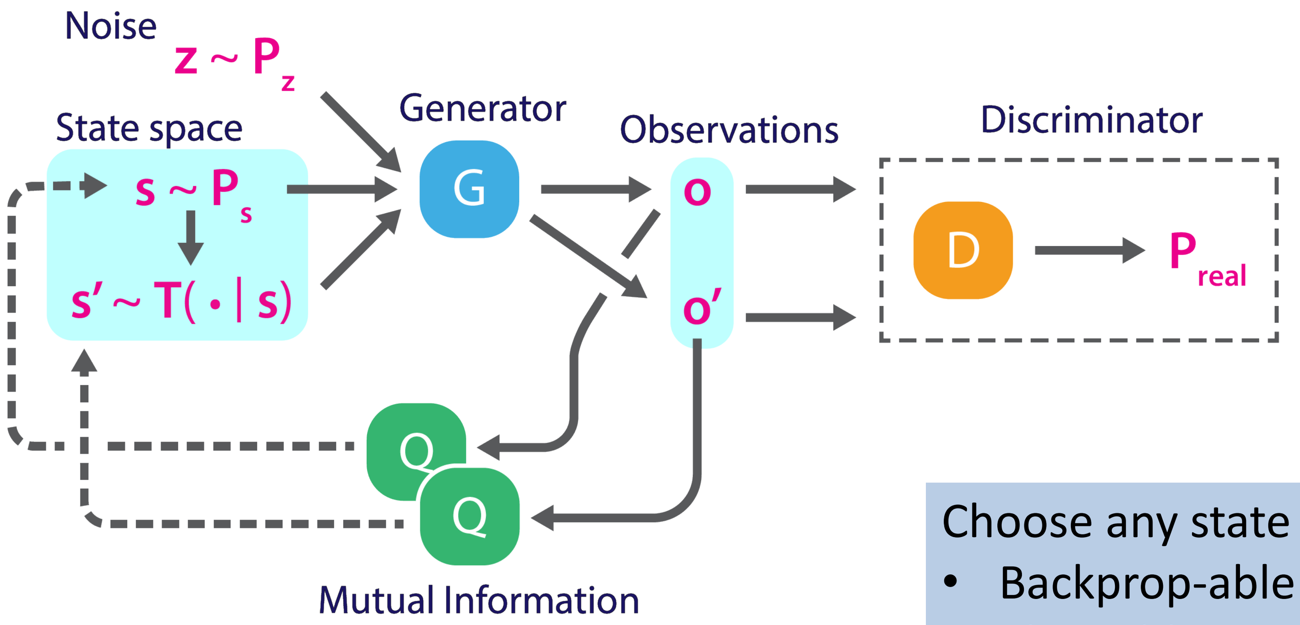
- Break up long horizon task into sub-goals
- Can be later input to inverse model / RL for actions.



Rope data from Nair et al., ICRA 2017

Causal InfoGAN

To solve problem above, we break it down into 3 steps -- map observations to states, plan from start state to goal state, and generate visual plans from state trajectories. Causal InfoGAN allows us to learn the state space such that efficient planning is possible. It is trained to explain causality in the data.



Choose any state space

- Backprop-able
- Plannable

$$\min_G \max_D V(G,D) - \lambda I(s,s';o,o')$$

$$V(G,D) = \mathbb{E}_{o,o' \sim P_{data}}[\log D(o,o')] + \mathbb{E}_{o,o' \sim P_G}[\log(1 - D(o,o'))]$$

$$I(s,s';o,o') \geq \mathbb{E}_{P_{s,s'},P_z}\left[\log Q_{s,s'|o,o'} - \log P_{s,s'}\right]$$

Disentangled Q: classify O → S

$$Q_{s,s'|o,o'} = Q_{s|o} Q_{s'|o'}$$

Rope Manipulation

Goal:

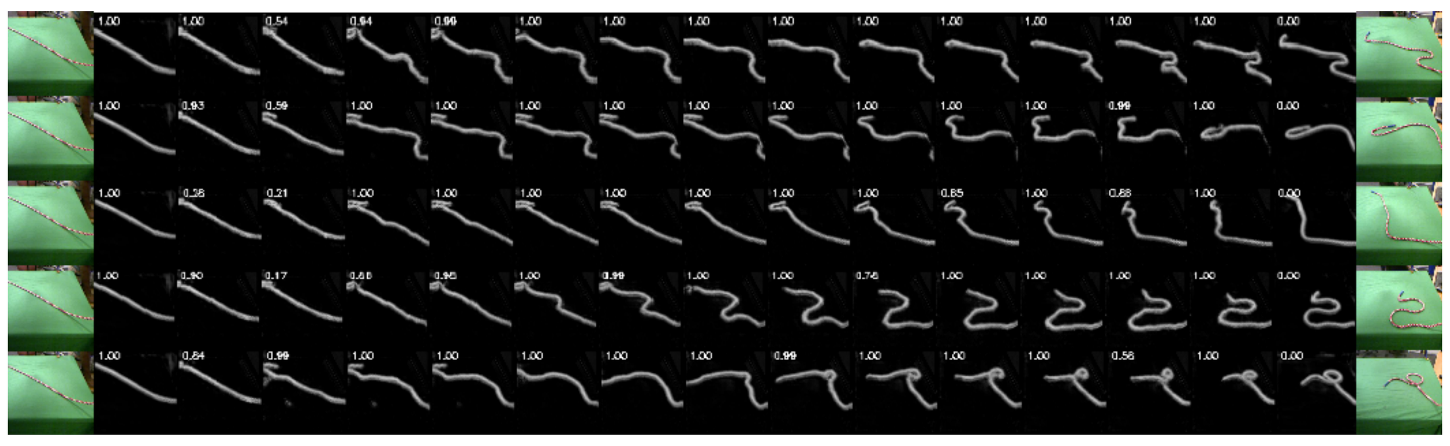
Generate a plan from start observation to goal observation O_{start} to O_{goal}

Key ideas:

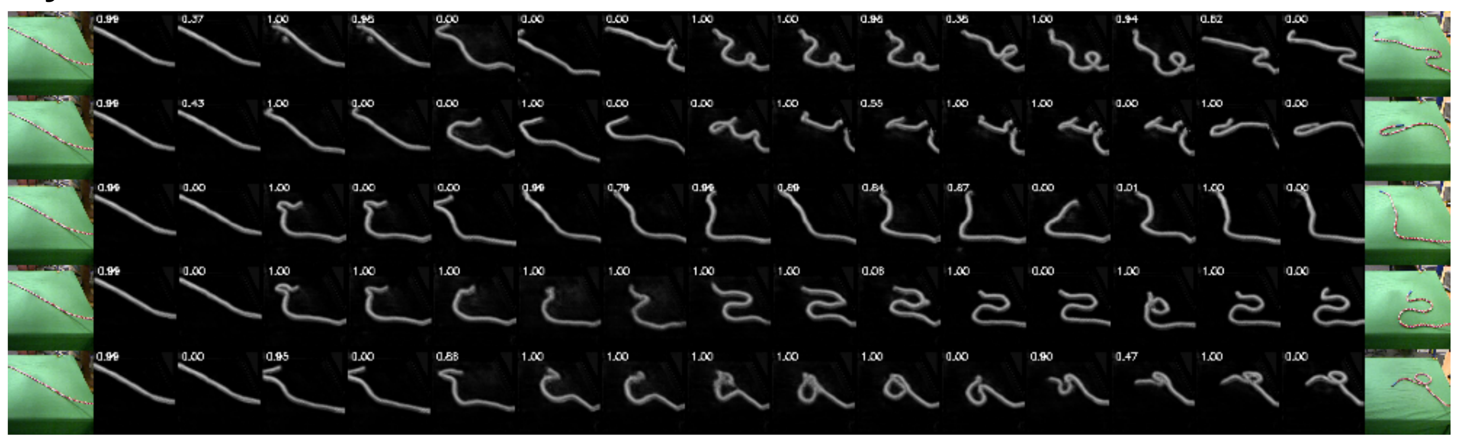
Plan using linear interpolation in continuous latent space
Transition is a local Gaussian model with small variance

Visual Plan Comparison:

Causal InfoGAN

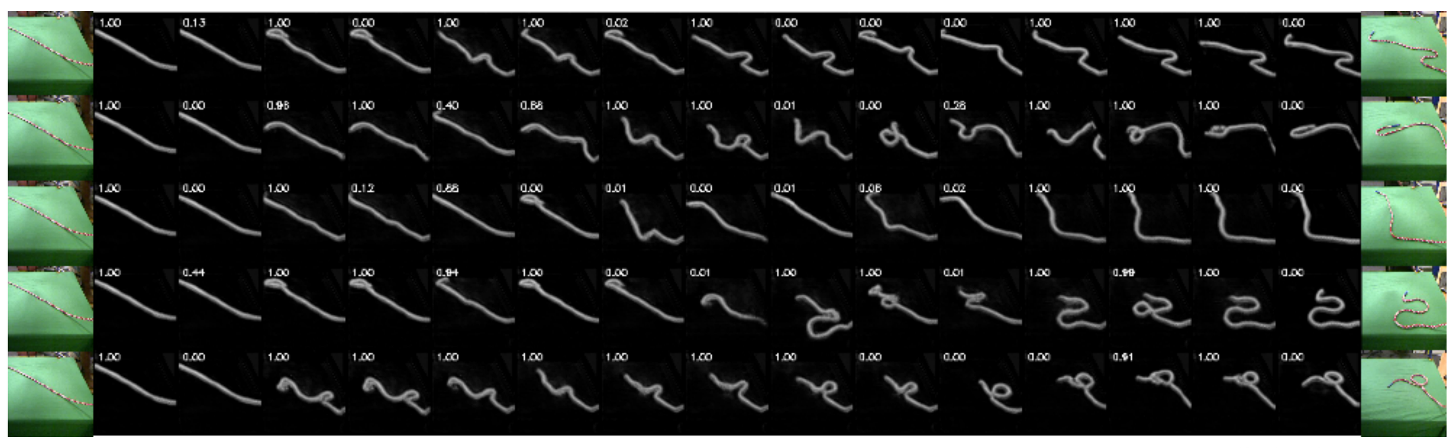


InfoGAN



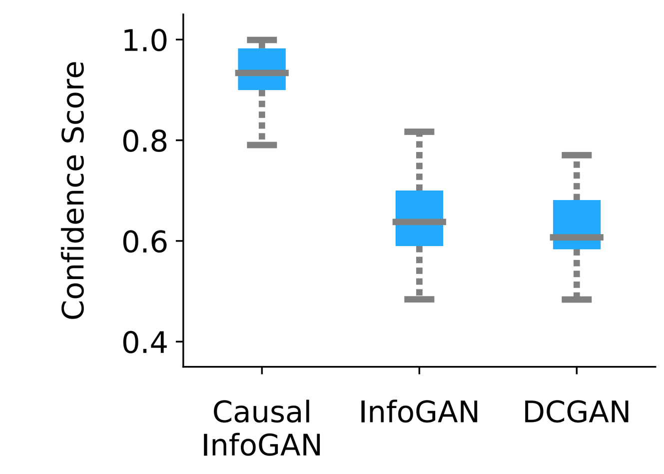
Chen et al., NIPS 2016

DCGAN



Radford et al., ICLR 2016

Numerical Evaluation



We train a classifier to predict the whether an observation pair is sequential, and compute the average of classifier score on visual trajectories.

Classical Planning

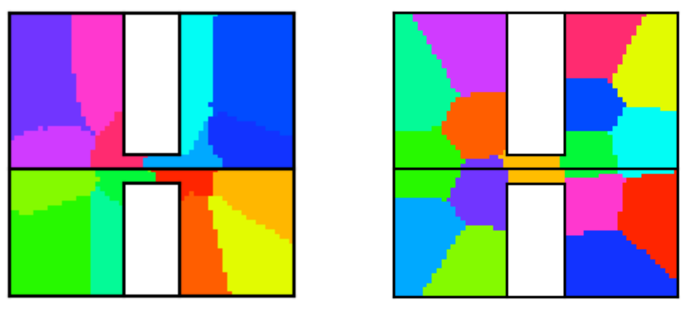
Goal:

Generate a plan from start observation to goal observation O_{start} to O_{goal}

Key ideas:

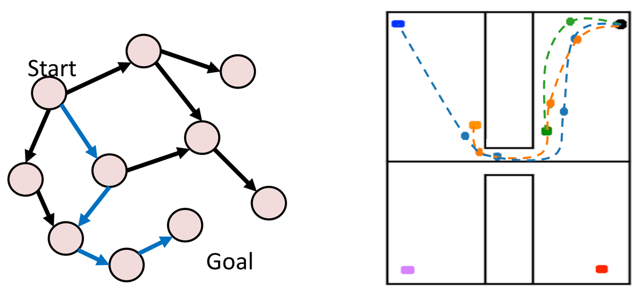
Plan using classical search such as Dijkstra's algorithm
Predicates and transitions are learned.

State Abstractions:



CiGAN K-mean

Planning:



Dijkstra Trajectories

CiGAN **learn abstractions** and **plan according to causality** in the data.

Numerical Evaluation:

	Planning Success Rate	
Causal-InfoGAN	**98%**	
K-means	12.25%	
Temporal K-means	7.0%	[Baram et al, 2016]
Spectral clustering	8.75%	[Srinivas et al., 2016]

Link to paper and code available at: sites.google.com/view/causal-infogan

Conclusion

Causal InfoGAN

- Deep generative models that explain dynamical systems
- General framework for structuring latent space so that is **easy to plan with.**

Visually interpretable plans!
Bridge classical planning + representation learning
Outlook

- Deeper combination of learning and planning.