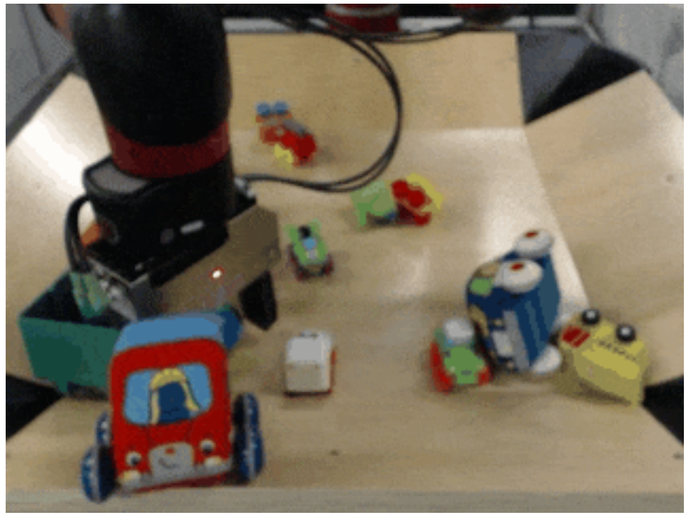
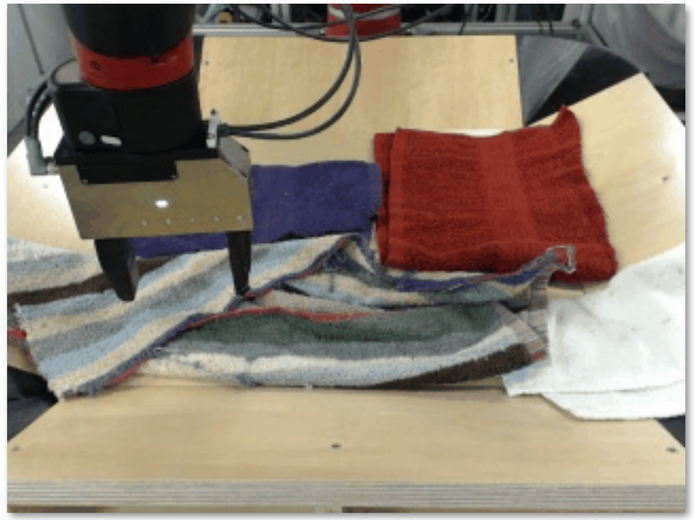




What can we learn from unlabeled interaction?

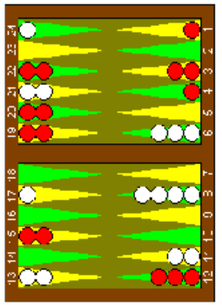
Chelsea Finn

UC Berkeley

Google Brain

Stanford

Impressive Feats in AI



TD Gammon

Watson

helicopter acrobatics

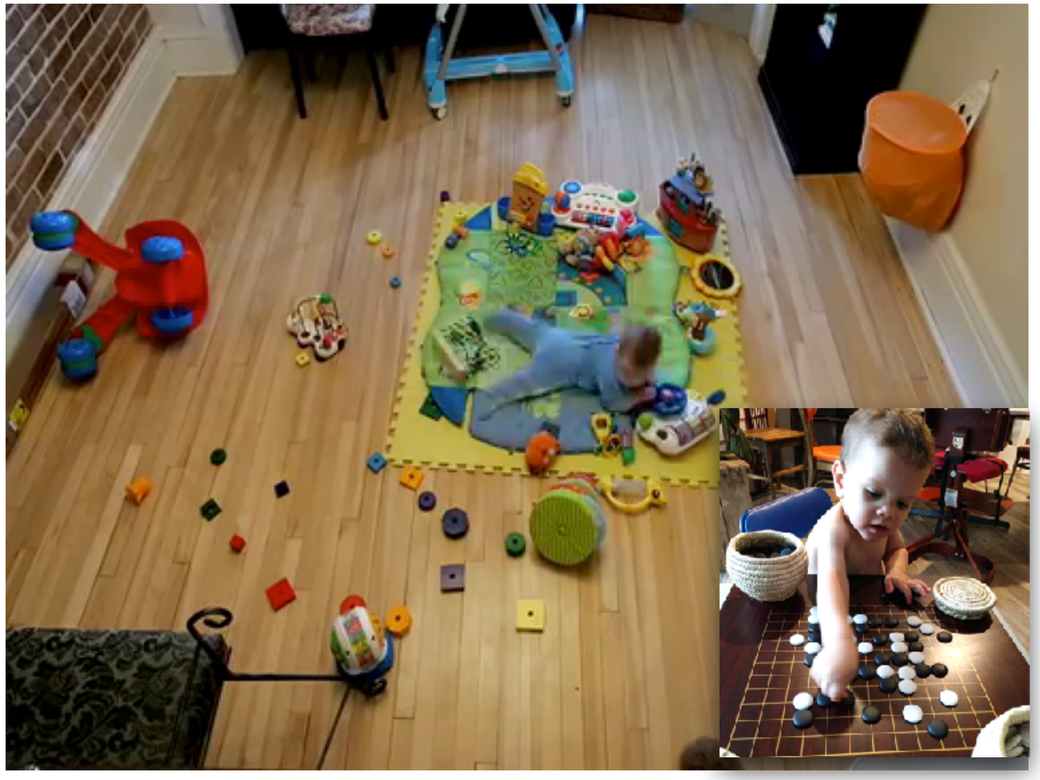
machine translation

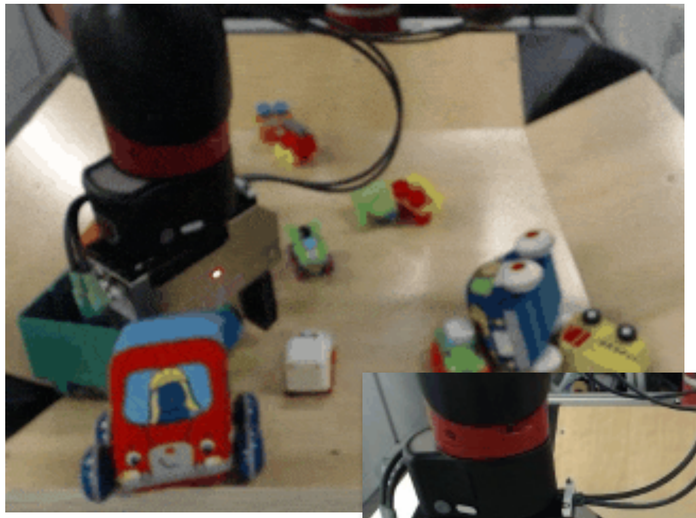
DQN

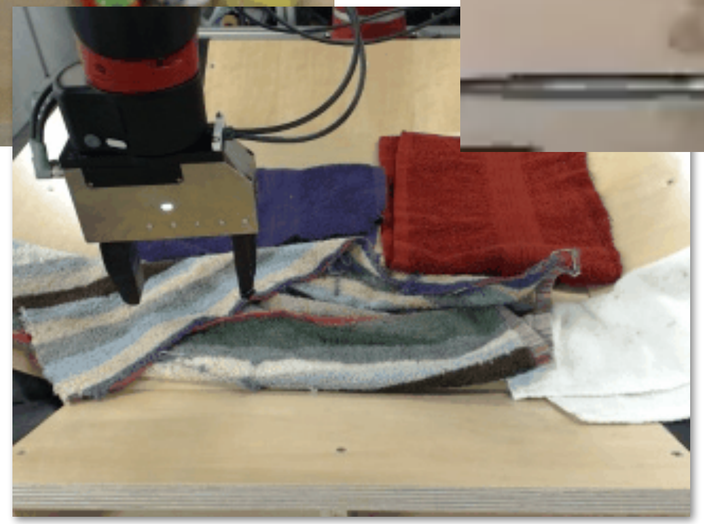
We are pretty good at building optimizers.

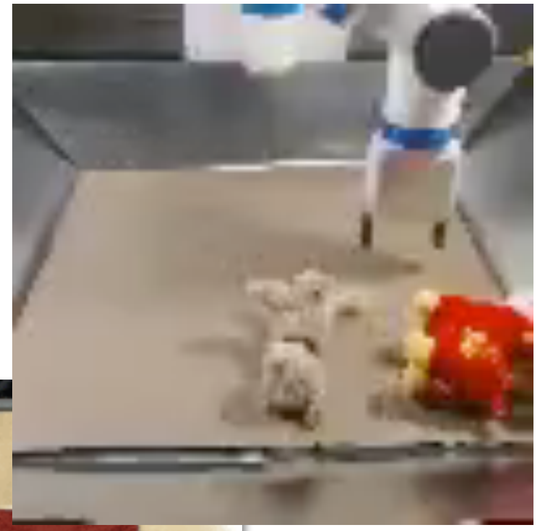
Result: specialized system that can solve **one task** (extremely well)

How can we build ***general-purpose*** & ***flexible*** systems?







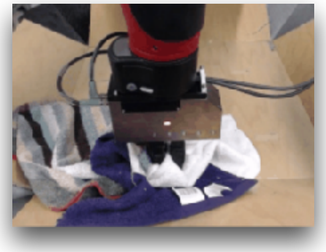


- Not prescribed to any particular task
- Contains rich information about the world
- **No notions** of **progress** or **success**

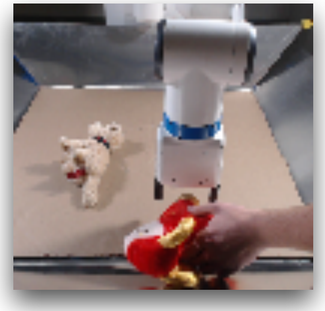
Can we learn from interaction before prescribing to a particular task?

in a way that enables us to later perform a breadth of tasks.

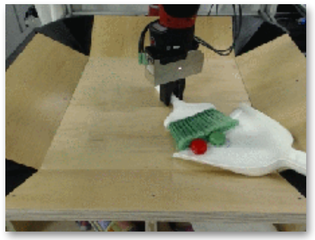
Outline



Learning task-agnostic models.

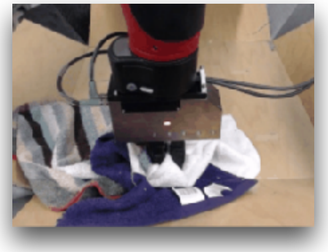


Learning task-agnostic goal representations.



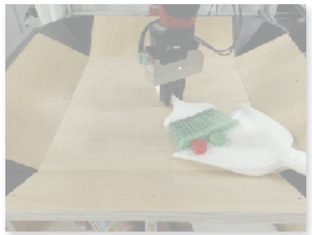
Learning task-agnostic models for complex tasks.

Outline



Learning task-agnostic models.

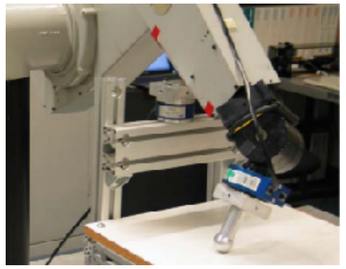
Learning task-agnostic goal representations.



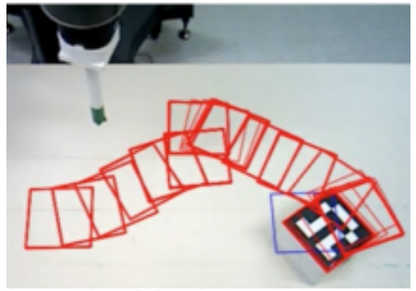
Learning task-agnostic models for complex tasks.

Learning Task-Agnostic Models through Self-Supervision

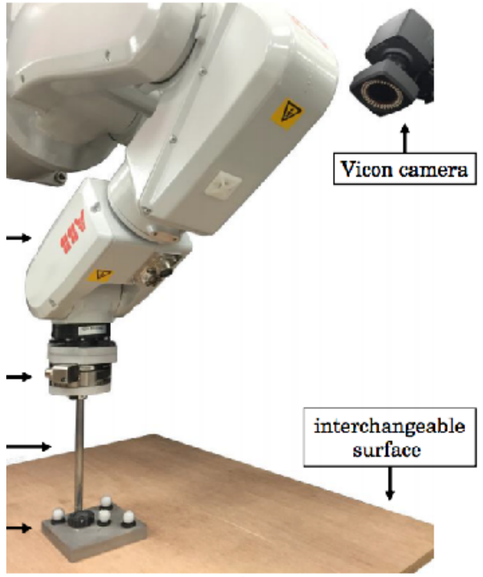
model-based control



Petrovskaya, Park, Khatib '07



Arruda, Mathew, Kopicki, Mistry, Azad, Wyatt '17



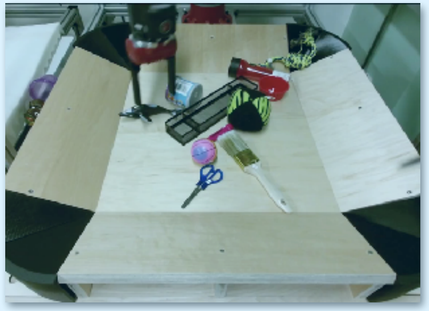
Yu, Bauza, Fazeli, Rodriguez '17

our approach

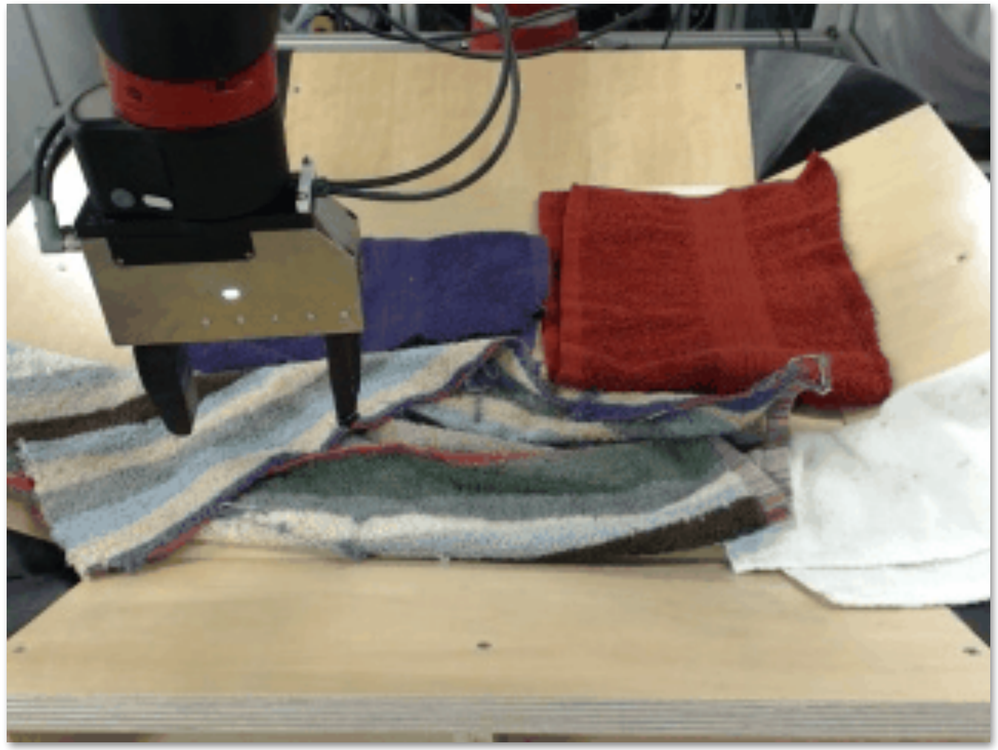
learn model from video data

many objects, raw perceptual inputs

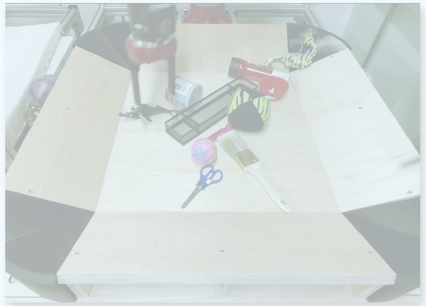
Collect data



Collect **diverse** data in a **scalable** way

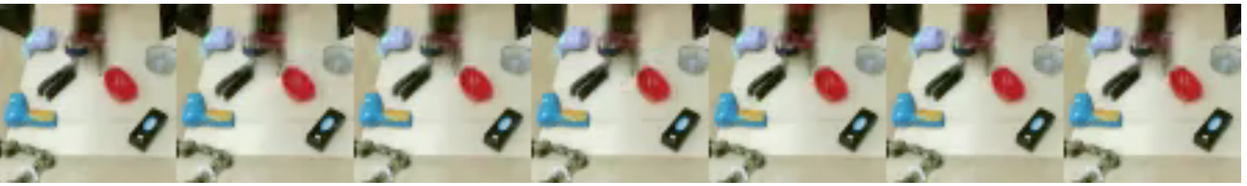


Collect data



Learn to predict

$I_t, a_{t:t+H} \rightarrow I_{t:t+H}$

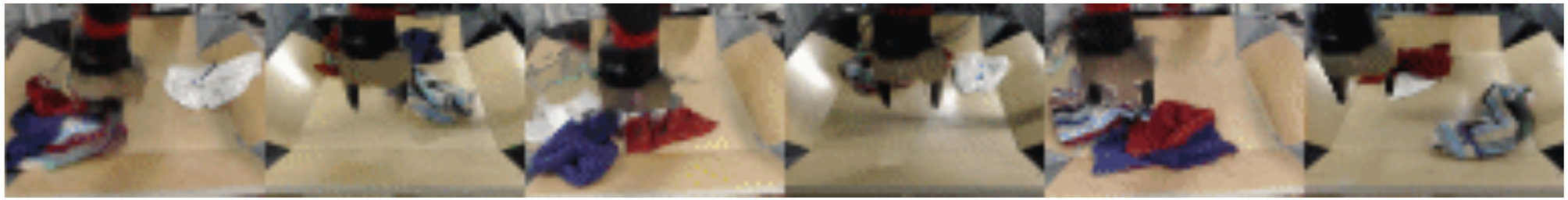


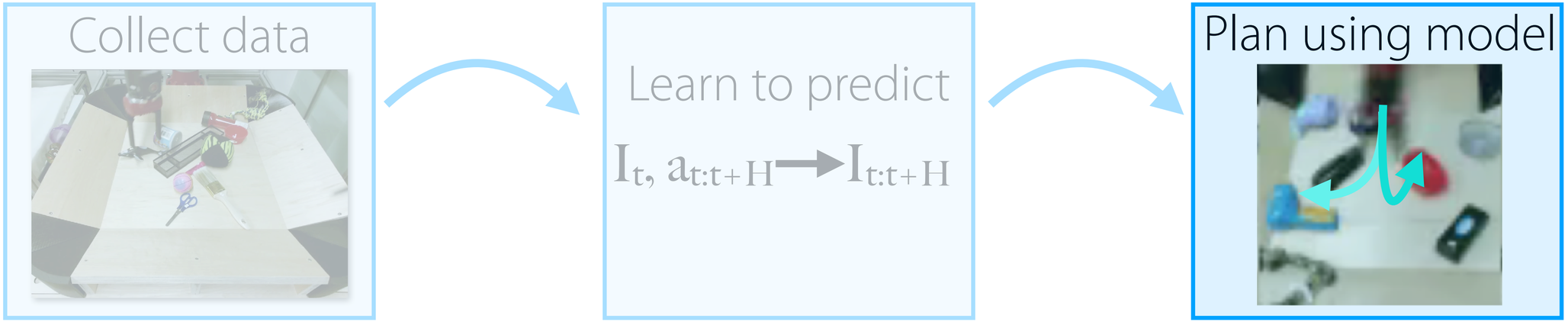
Contrast to:

Models capture **general purpose** knowledge about the world

Use **all** of the available supervision signal.

Also: No assumptions about task **representations**.

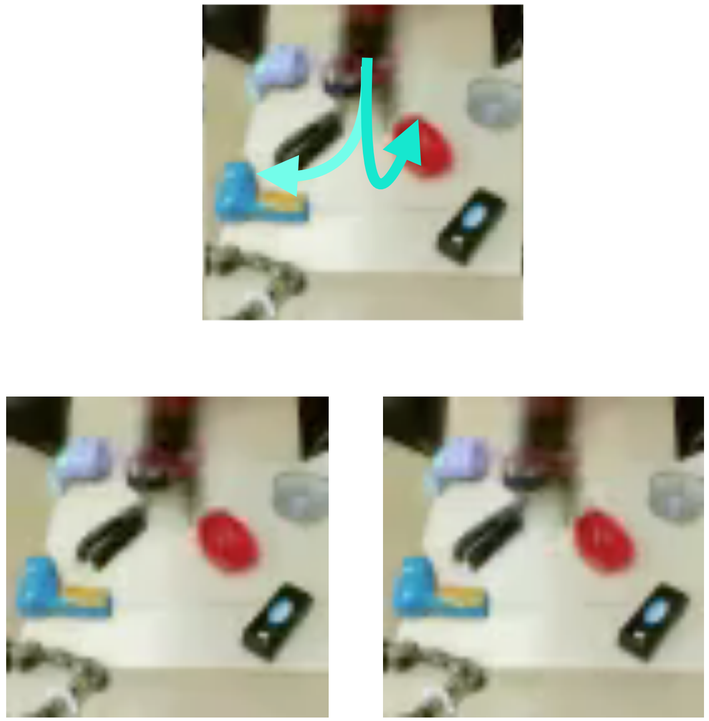




Planning with Visual Foresight

1. Consider potential action sequences
2. Predict the future for each action sequence
3. Pick best future & execute corresponding action
4. Repeat 1-3 to replan in real time

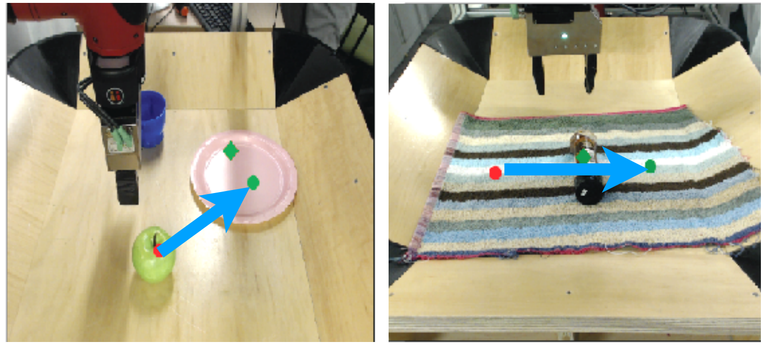
visual "model-predictive control" (MPC)



Finn & Levine. Deep Visual Foresight for Planning Robot Motion. ICRA '17

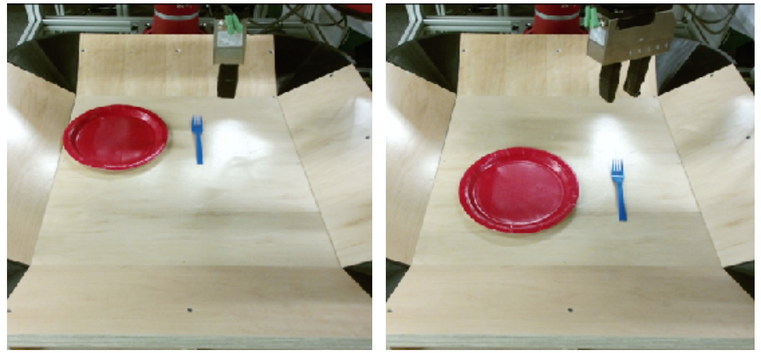
Which future is the best one?

Human specifies a goal by:



Selecting where pixels should move.

Providing an image of the goal.

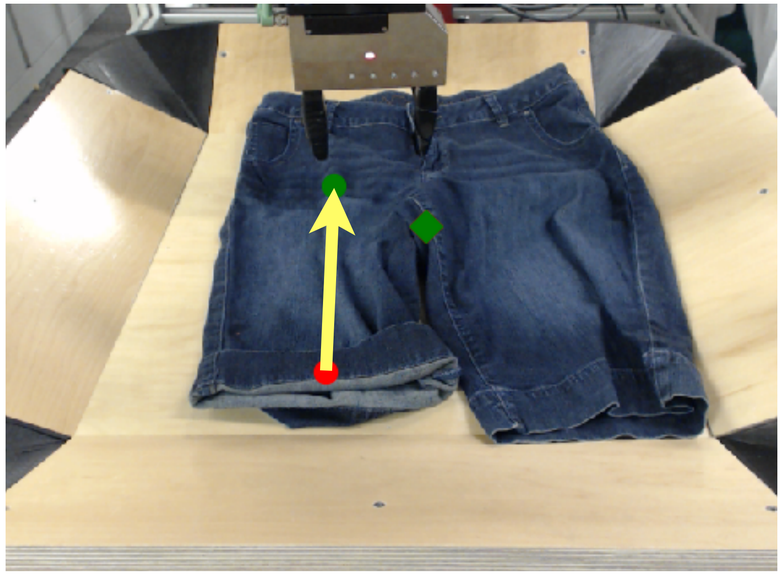


Providing a few examples of success.

Finn & Levine. Deep Visual Foresight for Planning Robot Motion. ICRA '17
Ebert, Lee, Levine, Finn. Robustness via Retrying: Closed-Loop Robotic Manipulation with Self-Supervised Learning. CoRL '18
Xie, Singh, Levine, Finn. Few-Shot Goal Inference for Visuomotor Learning and Planning. CoRL '18

How it works

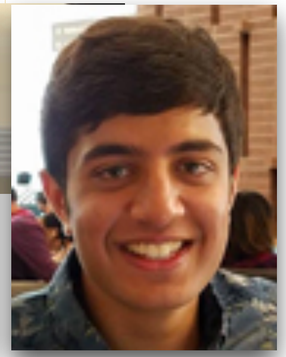
Specify goal



Visual MPC execution

Visual MPC w.r.t. goal

Frederik Ebert Sudeep Dasari



Ebert*, Finn*, Dasari, Xie, Lee, Levine. Visual Foresight, arXiv '18

How it works

Given 5 examples of success

infer goal classifier

visual MPC w.r.t. goal classifier

Visual MPC with learned objective

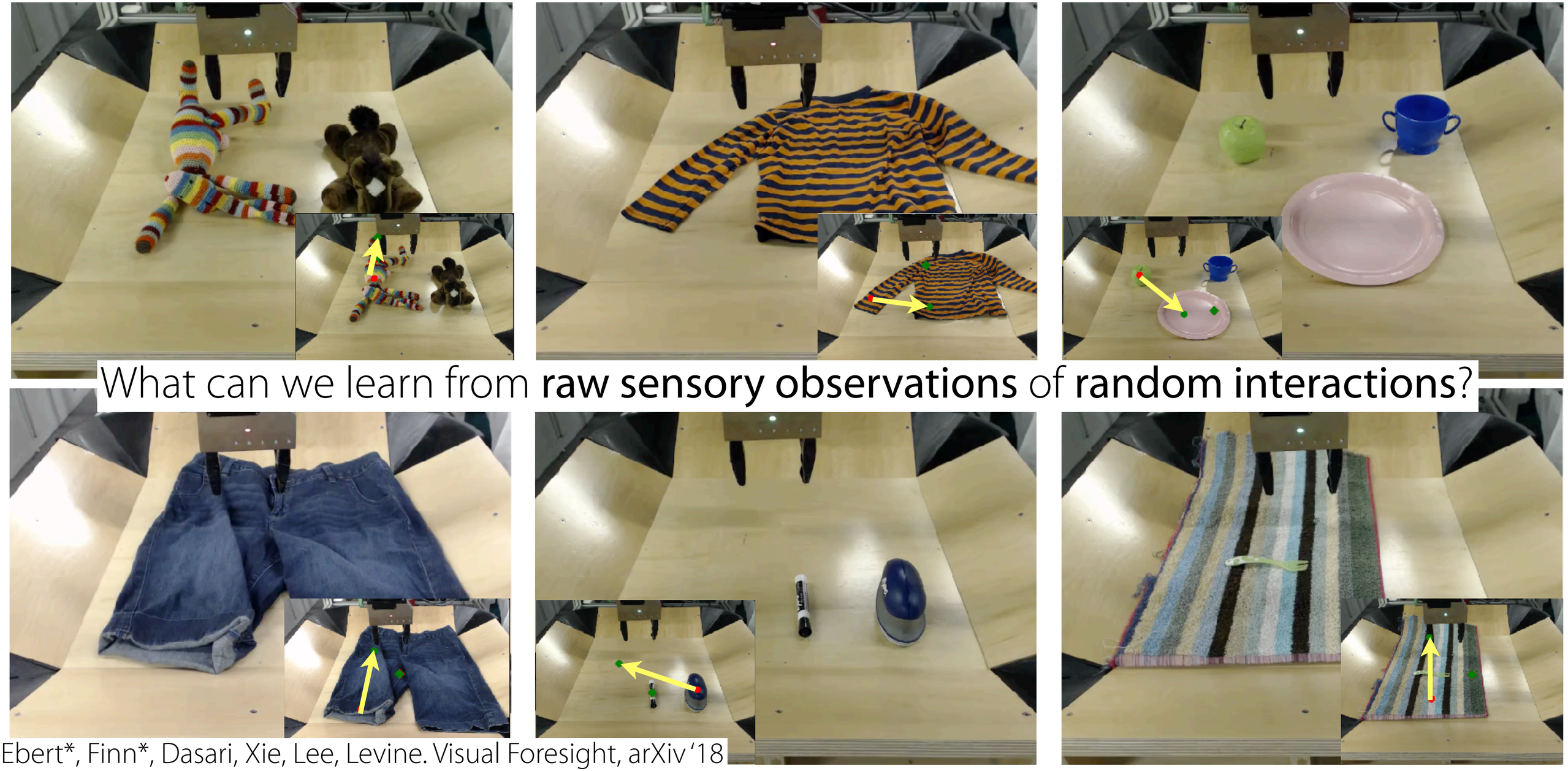
Annie Xie

Avi Singh

Xie, Singh, Levine, Finn. Few-Shot Goal Inference for Visuomotor Learning and Planning. CoRL '18

Planning with a **single model** for many tasks

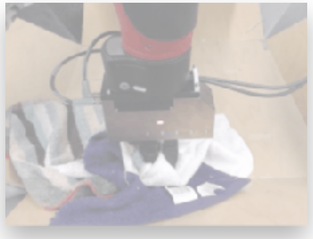
Video speed: 2x



What can we learn from **raw sensory observations** of **random interactions**?

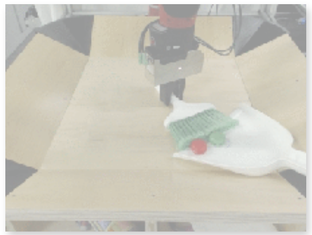
Ebert*, Finn*, Dasari, Xie, Lee, Levine. Visual Foresight, arXiv '18

Outline



Learning task-agnostic models.

Learning task-agnostic goal representations.



Learning task-agnostic models for complex tasks.

Can we learn to reach an image of any particular goal?

Problem: How close is image $\mathbf{I}_1$ to image $\mathbf{I}_2$?

- pixel distance?
- VAE distance?
- inverse model?

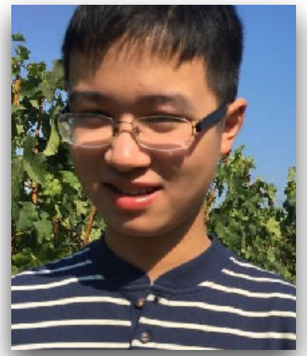
Can we acquire a control-centric goal representation through autonomous, unlabeled interaction?

Key insights:

(a) **Random, unlabeled interaction** is ***optimal*** under the 0/1 reward of reaching the last state. (Kaelbling '93, Andrychowicz '17)

(b) We can use this to optimize for a goal representation.

Tianhe Yu



Gleb Shevchuk

Yu, Shevchuk, Sadigh, Finn. Unsupervised Visuomotor Control through Distributional Planning Networks. RSS '19

Which representation, when used as a reward function, will cause a planner to choose the observed actions?

1. Collect random, unlabeled interaction data: $\{(s_1, a_1, \ldots, a_{t-1}, s_t)\}$
2. Train a latent state representation $s \rightarrow x$ & latent state model $f(x' \mid x, a)$ s.t. if we plan a sequence of actions w.r.t. a goal state $s_t \rightarrow x_t$, we recover the observed action sequence.
3. Throw away latent space model, return goal representation x.

"distributional planning networks"

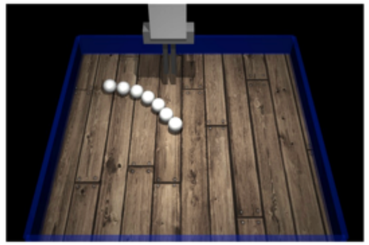
Srinivas, Jabri, Abbeel, Levine Finn. Universal Planning Networks. ICML '18

Yu, Shevchuk, Sadigh, Finn. Unsupervised Visuomotor Control through Distributional Planning Networks. RSS '19

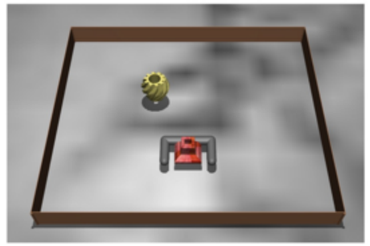
Evaluate metrics on achieving variety of goal images

reaching

rope manipulation

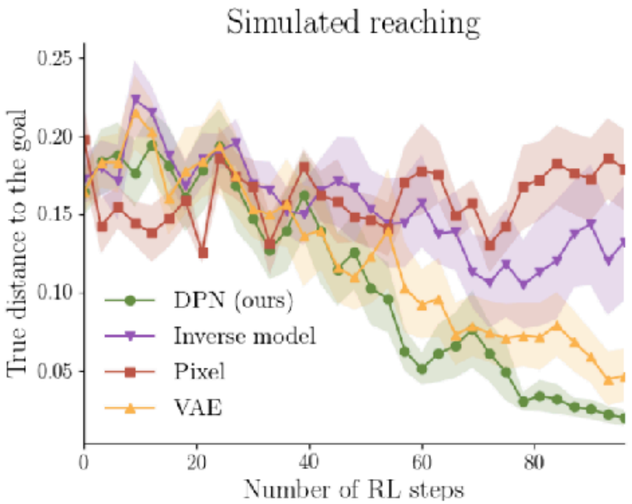


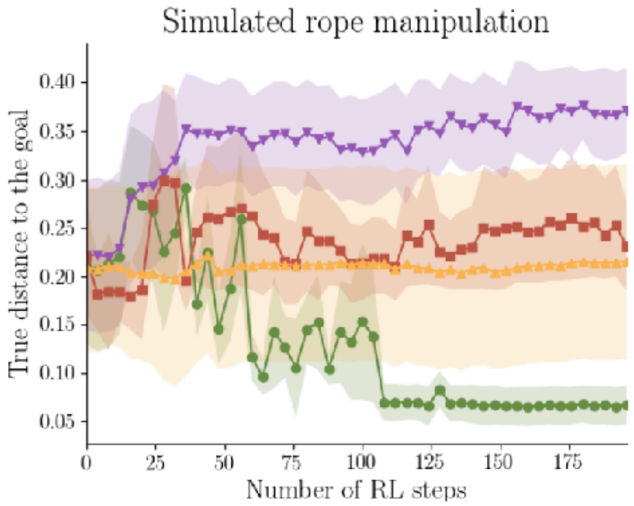
pushing

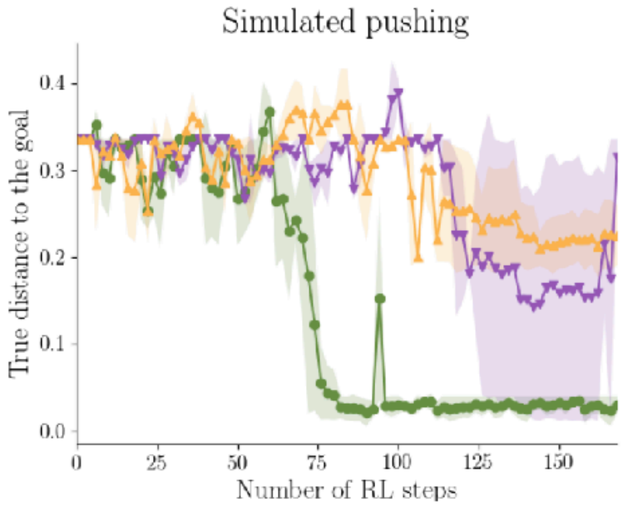


Compare:

- metric from **DPN** (ours)
- **pixel distance**
- distance in **VAE** latent space
- distance in **inverse model** latent space

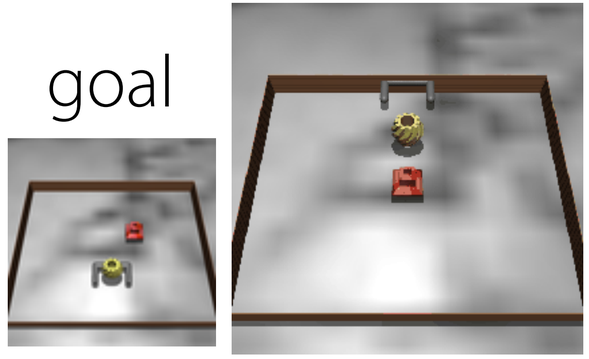






goal

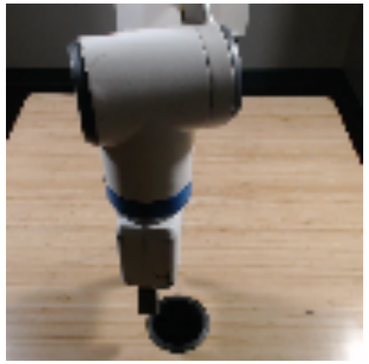
goal



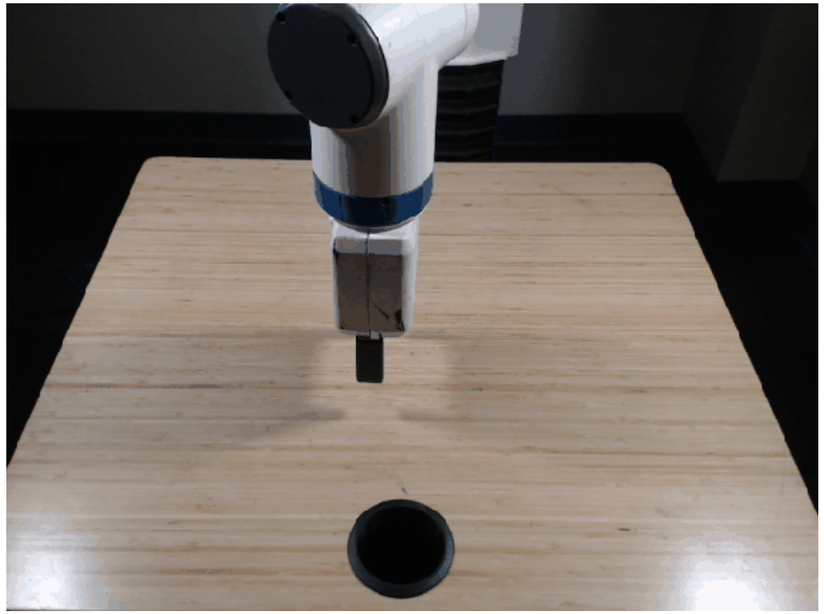
Yu, Shevchuk, Sadigh, Finn. Unsupervised Visuomotor Control through Distributional Planning Networks. RSS '19

Evaluate metrics on achieving variety of goal images

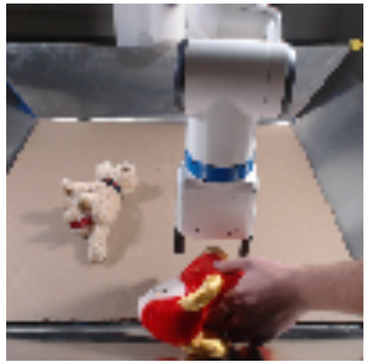
goal



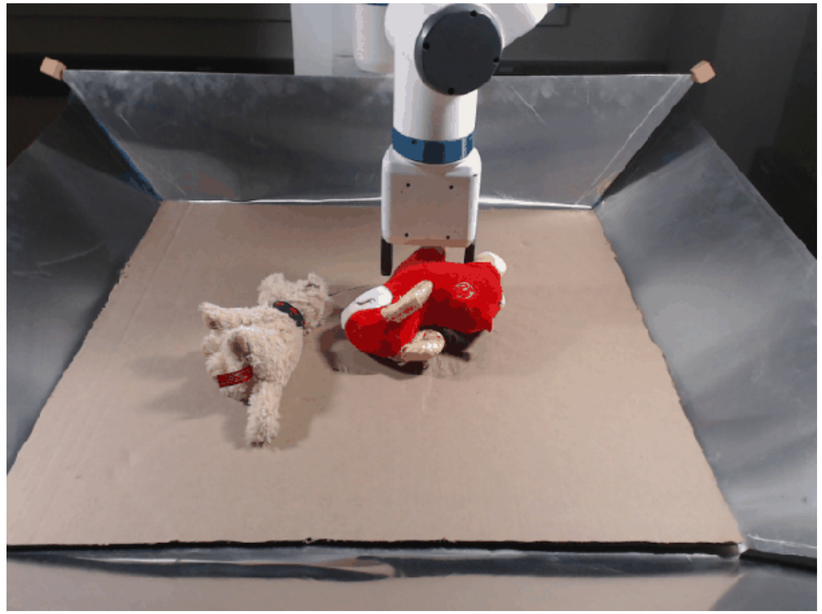
learned policy



goal



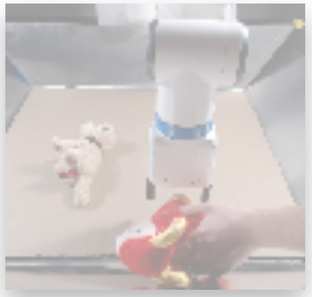
learned policy



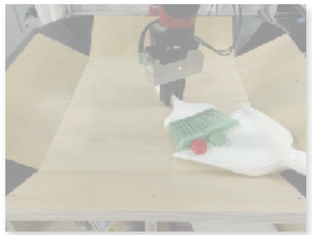
Yu, Shevchuk, Sadigh, Finn. Unsupervised Visuomotor Control through Distributional Planning Networks. RSS '19

Outline

Learning task-agnostic models.



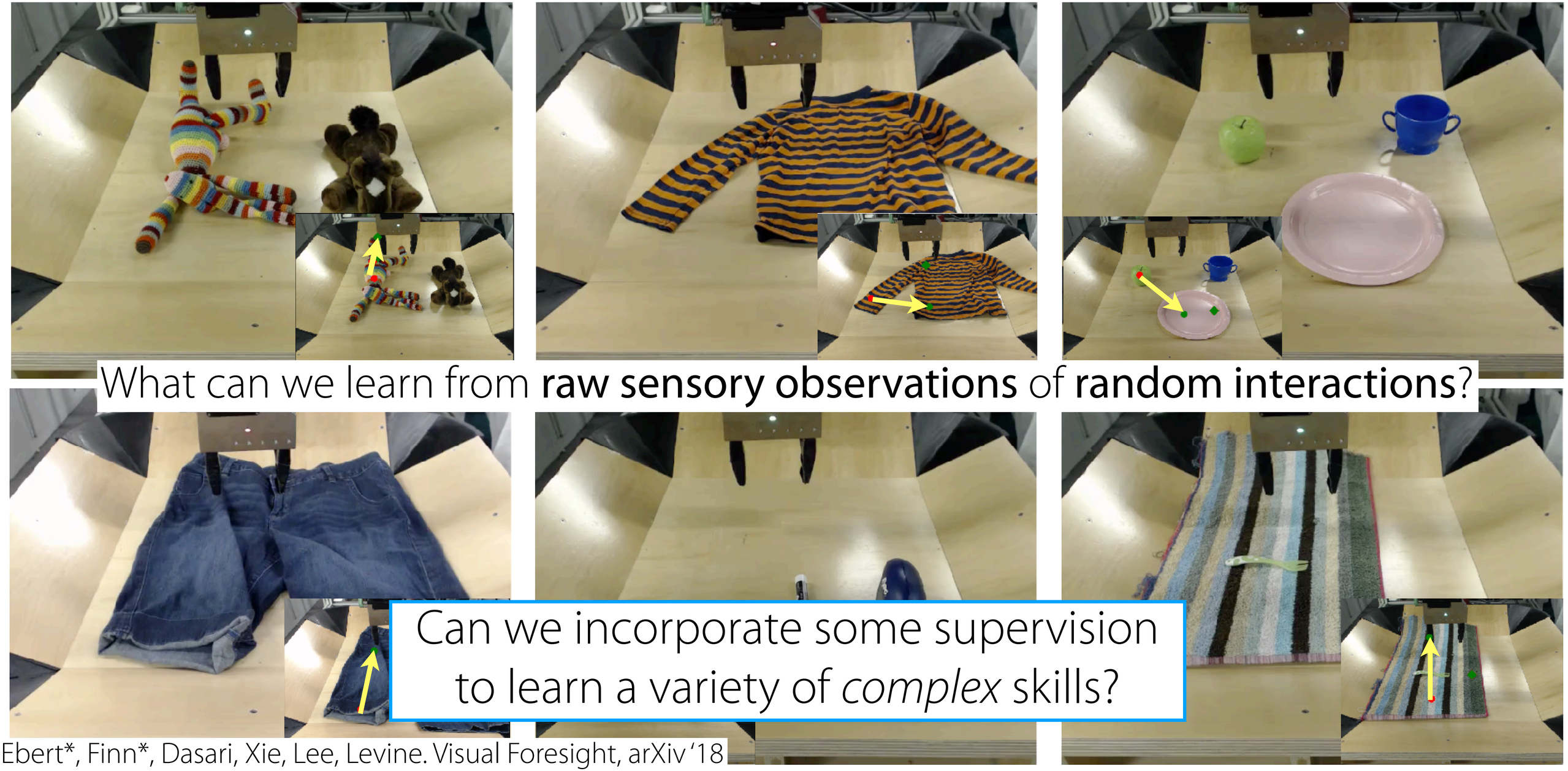
Learning task-agnostic goal representations.



Learning task-agnostic models for complex tasks.

Planning with a **single model** for many tasks

Video speed: 2x



What can we learn from **raw sensory observations** of **random interactions**?

Can we incorporate some supervision to learn a variety of *complex* skills?

Ebert*, Finn*, Dasari, Xie, Lee, Levine. Visual Foresight, arXiv '18

Collect data

Learn to predict

$I_t, a_{t:t+H} \rightarrow I_{t:t+H}$

Plan using model

direct data collection

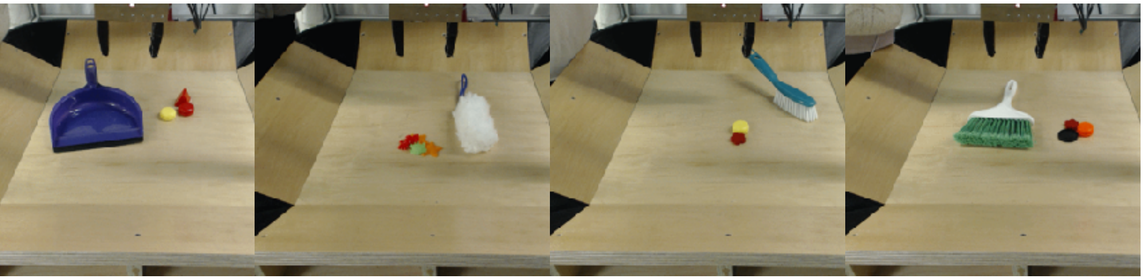
improve model

guide planning

Collect **diverse**, multi-task demonstrations

Fit model of $p(a_{t:t+H} \mid I_t)$ to the demonstration data.

Example multi-task demonstrations:



Samples from action proposal model:

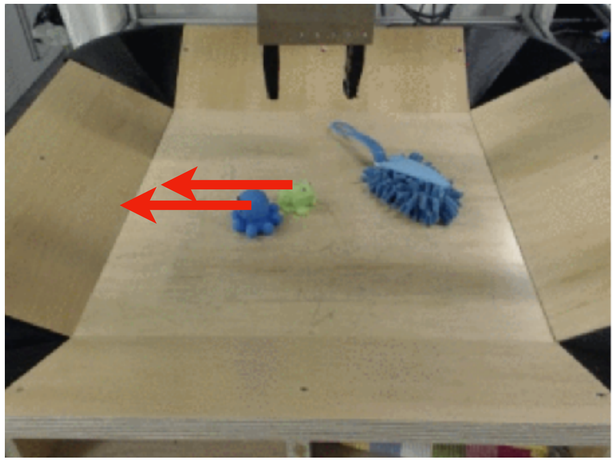
Annie Xie

Frederik Ebert

Xie, Ebert, Levine, Finn. Improvisation through Physical Understanding, RSS '19

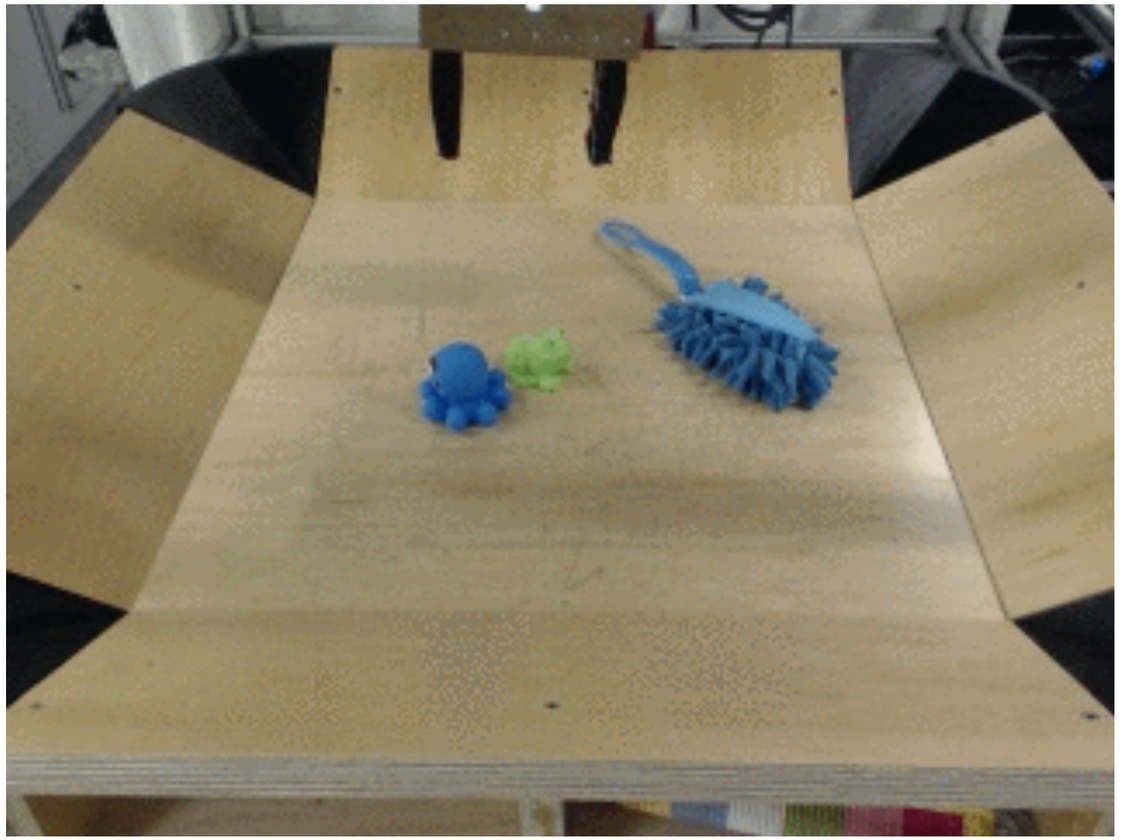
How it works

Specify goal



Guided visual planning w.r.t. goal

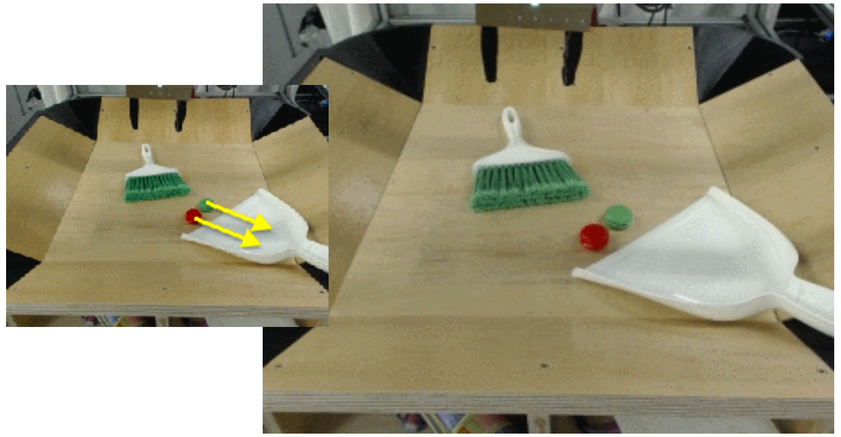
Executing actions



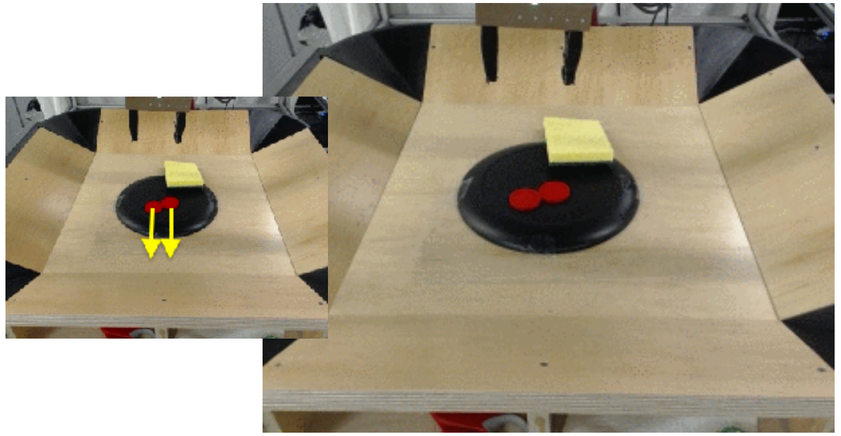
Xie, Ebert, Levine, Finn. Improvisation through Physical Understanding, RSS '19

Qualitative Experiments

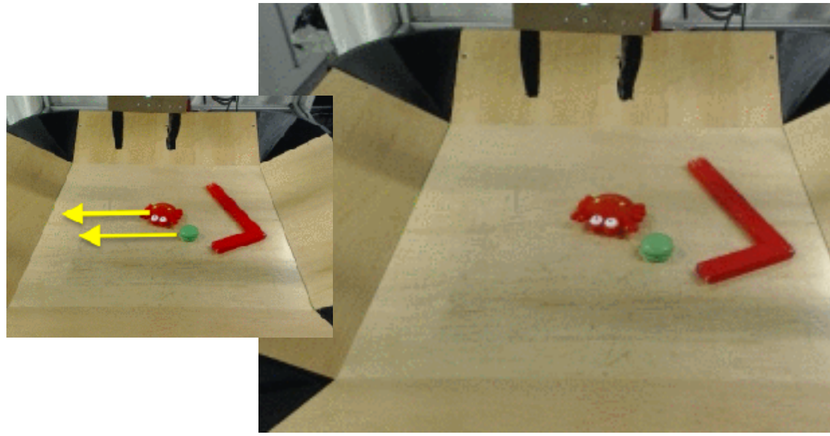
solve new tasks



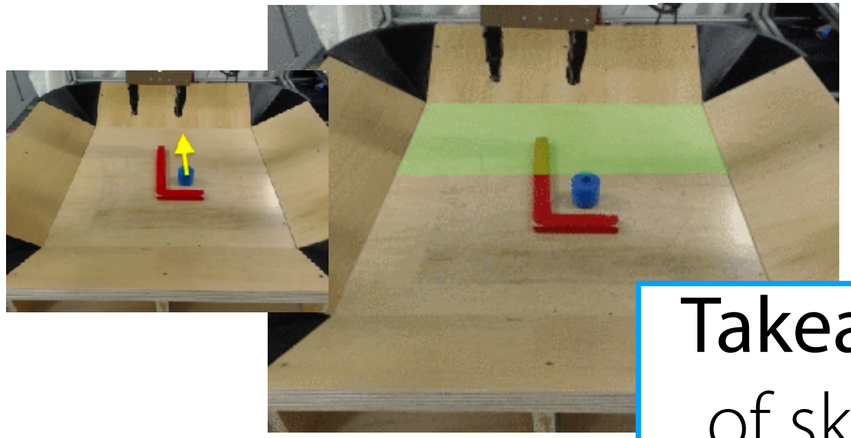
unseen tools



decide when to use a tool…

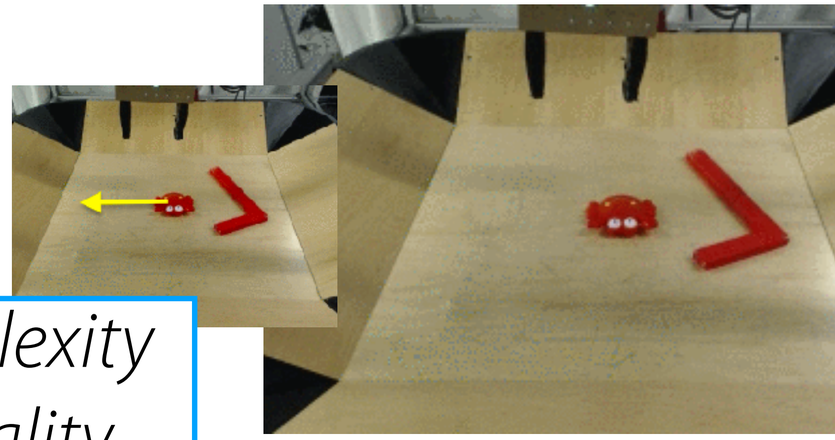


out-of-reach objects



unseen *unconventional* tools

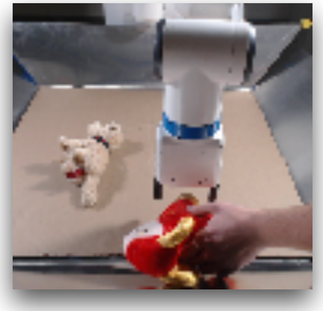
…and when not to



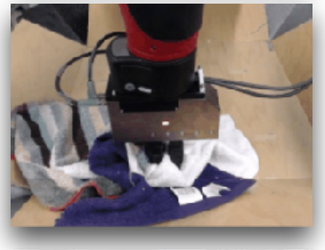
Takeaway: Achieve greater *complexity* of skills while maintaining *generality*.

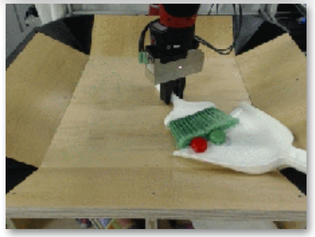
Xie, Ebert, Levine, Finn. Improvisation through Physical Understanding, RSS '19

Takeaways



Task-agnostic goal representations: Learn representation for unsupervised visuomotor control.





Task-agnostic models:
- Learn from diverse unlabeled experiences to generalize to many objects and goals.
- Incorporate diverse, multi-task demonstrations for generality + complexity

Hypothesis: Approaches that learn ***without human supervision*** will enable more ***general*** behavior.

- can collect more data
- forced to not commit to any task

What's next?

Future work: How can we build better task-agnostic models?

Can we model **uncertainty** over future observations?

More and more uncertainty over time.

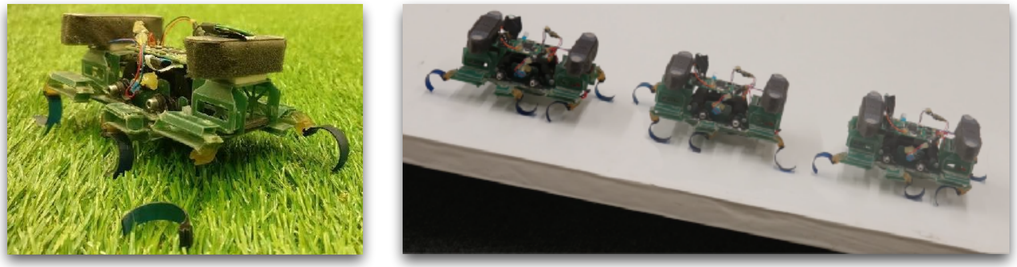
Stochastic adversarial video prediction

Lee, Zhang, Ebert, Abbeel, Finn, Levine. 2018

Can we **adapt the model** with a small amount of experience?

Physical properties unknown until interaction.

Few-shot, online model adaptation



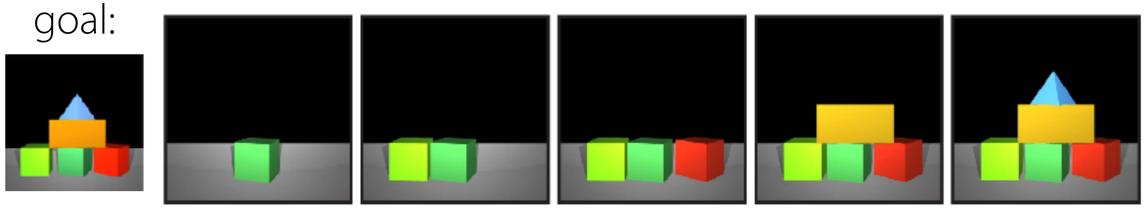
Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. ICLR 2019

Can we learn & plan with **structured representations** of the world?

Structured, lossy representations can enable long-horizon planning and reasoning.

Object-oriented planning & prediction

goal:



Janner, Levine, Freeman, Tenenbaum, Finn, Wu. ICLR 2019

Takeaways

Task-agnostic goal representations for unsupervised visuomotor control.

Task-agnostic models: from diverse unlabeled experiences to generalize to many objects & goals

Hypothesis: Approaches that learn ***without human supervision*** will enable more ***general*** behavior.

- can collect more data
- forced to not commit to any task

Collaborators

Sergey Levine

Pieter Abbeel

Frederik Ebert

Sudeep Dasari

Annie Xie

Alex Lee

Tianhe Yu

Dorsa Sadigh

Gleb Shevchuk

Papers, data, and **code** linked at: people.eecs.berkeley.edu/~cbfinn

Questions?