Learning Generalizable Models through Unsupervised Interaction

Chelsea Finn
Impressive Feats in AI

TD Gammon
Watson
machine translation
DQN
helicopter acrobatics
AlphaGo

Why are these impressive?
They perform a complex task very well, sometimes even better than a human.

“specialists”

What is equally important: but not impressive (on the surface)
Generality: ability to perform many tasks

How can we build generalists?
It turns out — the simpler, but broader capabilities are really hard. (Moravec’s Paradox)

This talk: can we do the unimpressive things?
Can we build an agent that can do many tasks?

Learning a policy in a closed universe

Learn general-purpose model + plan with model for many tasks

Model-based control

From raw observations, with limited supervision, in the physical world
Can we build an agent that can do many tasks?

- **learning a policy** in a closed universe
- **learn general-purpose model** + **plan with model** for many tasks
- **adaptable** model for dynamic environments
- **visual model of diverse, open-world environments** from **raw observations**, with **limited supervision**, in the **physical world**
Model-based Reinforcement Learning

Collect data

\{s_t, a_t\}

Fit model

\[
\begin{align*}
\hat{p}_\theta \\
\end{align*}
\]

Plan using the model

Can we learn a model that can be adapted online in dynamic environments? (online system identification with neural networks)

Few-Shot Adaptation via Model-Agnostic Meta-Learning

Key idea: Learn how to adapt to new scenarios from small amounts of data

\[
\text{MAML: } \min_\theta \sum_{\text{task } i} \mathcal{L}_{\text{test}}^i(\theta - \alpha \nabla_\theta \mathcal{L}_{\text{train}}^i(\theta))
\]

Can we learn a model that can be adapted online in dynamic environments?
Can we learn a model that can be adapted online in dynamic environments?

online adaptation = few-shot learning

“tasks” are temporal slices of experience

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real-World Environments through Meta-RL
online adaptation = few-shot learning  "tasks" are temporal slices of experience

Meta-training:
$$\min_\theta \sum_{\text{task } t} \mathcal{L}^t_{\text{test}}(\theta - \alpha \nabla_\theta \mathcal{L}^t_{\text{train}}(\theta))$$

- model error on $s_{t:t+H}, a_{t:t+H-1}$
- model error on $s_{t-M:t}, a_{t-M:t-1}$

Result: meta-learned initialization $\theta$

Test time:
At every time step $t$:
1. Reset to learned initialization $\theta$
2. Adapt model parameters $\theta' = \theta - \alpha \nabla_\theta \mathcal{L}^t_{\text{train}}(\theta)$
3. Plan using adapted model
Dynamic Environments with Online Adaptation with MAML

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real-World Environments through Meta-RL
VelociRoACH Robot

Meta-train on variable terrains

Meta-test with slope, missing leg, payload, calibration errors

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real-World Environments through Meta-RL
VelociRoACH Robot

Meta-train on variable terrains  Meta-test with slope, missing leg, payload, calibration errors

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real-World Environments through Meta-RL
VelociRoACH Robot

Meta-train on variable terrains  Meta-test with slope, **missing leg**, payload, calibration errors

Our method  Model-Based RL (no adaptation)

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real-World Environments through Meta-RL
Can we build an agent that can do **many tasks**?

- **learn general-purpose model**
- **plan with model for many tasks**

**learning a policy in a closed universe**

**adaptable model for dynamic environments**

**modeling diverse, open-world environments**

from **raw observations**, with **limited supervision**, in the **physical world**
Our goal: generalize to **novel objects** and, also to **many tasks** (by learning a *general-purpose* model)

**Overall approach:** Collect data, learn model, plan to achieve many tasks
Collect diverse data in a scalable way

No supervision —> Scalable data collection —> Diverse data
(data available online!)
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**.
Are these models useful?
How can we use these models to plan?
(to achieve many human-specified goals)
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)
How to predict video?

- deep recurrent network
- multi-frame prediction
- action-conditioned
- explicitly model motion

Key idea: output how pixels will move, rather than pixels themselves
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.
How it works

Specify goal

Visual MPC w.r.t. goal

Visual MPC execution

Novel Object Positioning via Visual Foresight

Given 5 examples of success

inferred goal classifier

visual MPC w.r.t. goal classifier

Visual MPC with learned objective

How it works

Xie, Singh, Levine, Finn. Few-Shot Goal Inference for Visuomotor Learning and Planning
Planning with a **single model** for many tasks

Completely self-supervised

Only human involvement during training is initial programming and providing objects
Demo at NIPS 2017: Long Beach, CA

planning with visual models

Demo at AI4ALL Outreach Camp

one-shot imitation

The students were unimpressed.
(but still had fun)
Can we build an agent that can do many tasks? from raw observations, with limited supervision, in the physical world.

**Takeaways**

- adaptable model for dynamic environments
- modeling diverse, open-world environments
- dynamic environments
- data-efficient
- significant object diversity
- minimal supervision
What’s next? How can we build better, more useful models of the world?

Can we model **uncertainty** over future observations?
More and more uncertainty over time.

Can we learn **structured** representations & models from video?
Structured representations can lead to better data efficiency and generalization.

How should we **model the objective**?
Agents need internal representation of the goal in the real world.

Stochastic adversarial video prediction
Lee, Zhang, Ebert, Abbeel, Finn, Levine. 2018

Learning object-centric models
Janner, Levine, Freeman, Tenenbaum, Finn, Wu. 2018

Goal inference from images
Xie, Singh, Levine, Finn. CoRL 2018
Collaborators

Frederik Ebert  Sudeep Dasari  Annie Xie  Avi Singh  Anusha Nagabandi  Ignasi Clavera  Simin Liu  Pieter Abbeel  Sergey Levine

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

Questions?