

18. Visuomotor Policies (via behavior cloning)

Last Lecture outline

- Manipulator Control

- multibody dynamic equation

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} = \tau_g(q) + u + \sum J_i(q)f_i$$

- trajectory tracking
- Force Control (direct / indirect)
- all of above are controlling the DOF of robot, but the big control problem:
- $q = [q_{robot}, q_{environment}]'$

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} = \tau_g(q) + Bu + \sum J_i(q)f_i$$

which becomes an optimal control problem

- Perception with deep learning

- perception also gets harder, estimate q
- Sometimes we have to use probabilities


- We are going to control the robot & environment using

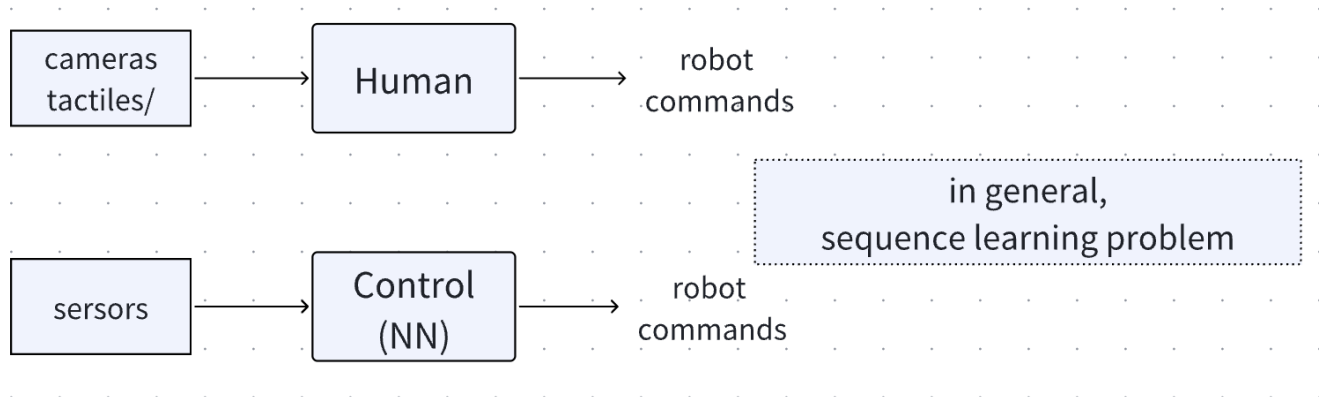
- imitation learning
- Reinforcement learning
- Model-based Reinforcement Learning

Today's Topic outline

- policy extension

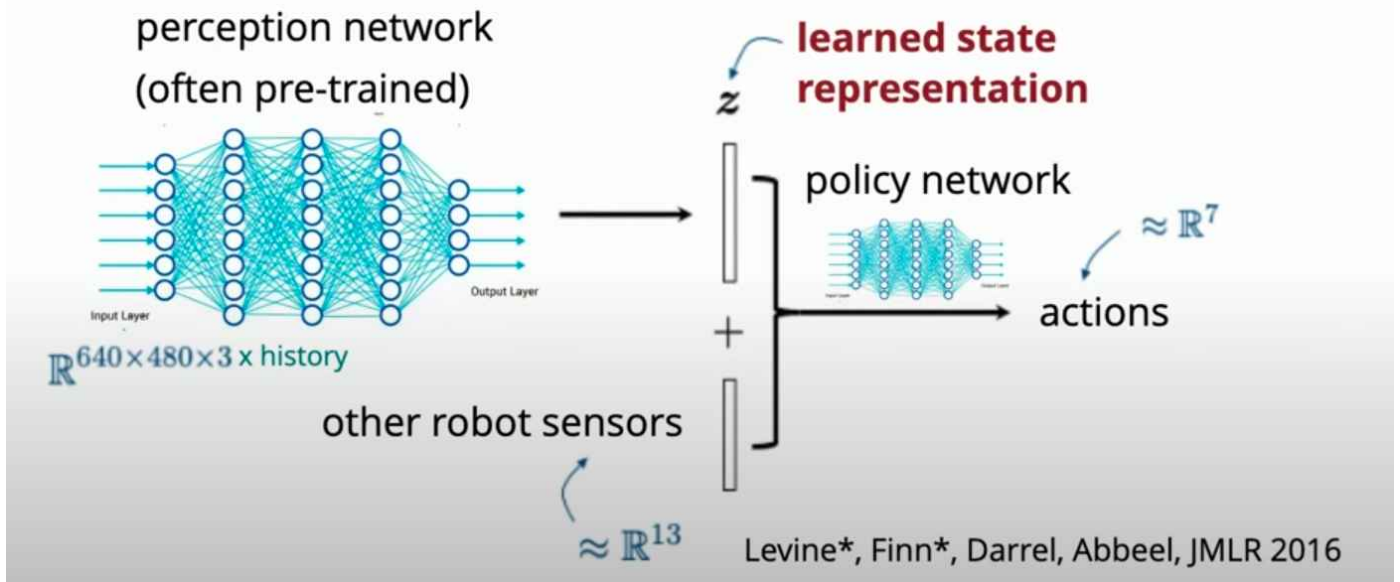
- control "policy": $u = \pi(q, \dot{q})$, $q = q(\text{robot})$
- try to extend to: $u = \pi(q, \dot{q})$, $q = q(\text{robot} + \text{environment})$
- why not $u = \pi(\text{camera}, \text{tactile}, \text{etc})$
- write a single controller that works for an entire class of objects

- Try to do the controller more locally, in an easier way
 -  [Deep Imitation Learning for Complex Manipulation Tasks from Virtual Reality Teleoperation.pdf](#)
 - Control becomes a supervised learning problem
 -



- Imitation Learning
 - Behavior cloning (mimic input/output from human)
 - Inverse Optimal Control / Inverse Reinforcement Learning (just learn the reward, then do RL)
- Visuomotor policies
 - Pre-train on Image-net, we don't care about elephants/dogs etc, drop off the last layer, and add a layer to emphasize dishes and mug, fine tuning.
 - **Close a loop on the RGB sensors directly, Closing the loop on cameras!! So valuable**

Visuomotor policies



[End-to-End Training of Deep Visuomotor Policies.pdf](#)

- a lot of people use Transformers, google RT1, RT2 , decision transformers
- [Dense Descriptors](#)

1. Start on tasks that can not be simulated, food, deformable object

- Diffusion Policy, Visuomotor Policy Learning via Action Diffusion
- Does the task, humans need the force feedback to get it done?

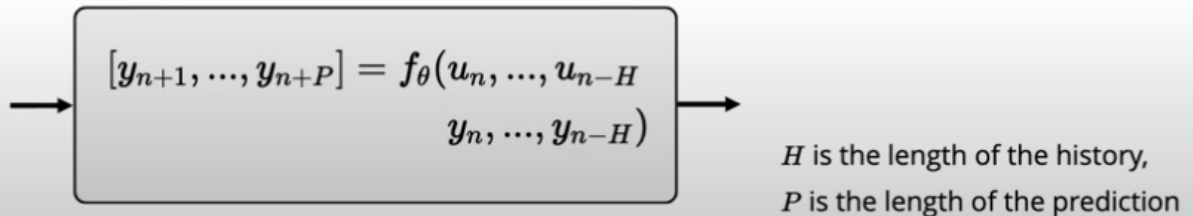
2. Diffusion, another self supervised learning

- take any image in your data set and add noise, random gaussian
- train the backwards process

Representing dynamic output feedback



"Diffusion Policy" is an auto-regressive (ARX) model with *forecasting*



Conditional denoiser produces the forecast, conditional on the history

- [auto-regressive model](#) ; predicte a sequence of actions

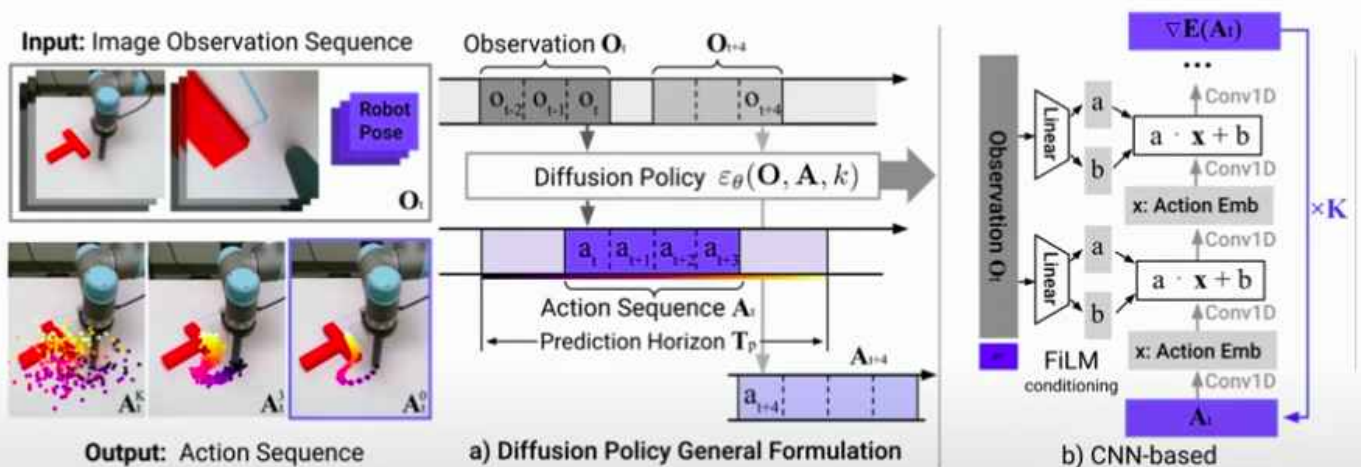


Image backbone: ResNet-18 (pretrained on ImageNet)

Total: 110M-150M Parameters

Training Time: 3-6 GPU Days (\$150-\$300)

Why (Denoising) Diffusion Models?

- High capacity + great performance
 - Small number of demonstrations (typically ~50)
 - Multi-modal (non-expert) demonstrations
 - Training stability and consistency
 - no hyper-parameter tuning
 - Generates high-dimension *continuous* outputs
 - vs categorical distributions (e.g. RT-1, RT-2)
 - Action-chunking transformers (ACT)
 - Solid mathematical foundations (score functions)
-
- I've discussed training *one* skill
 - Wanted: few shot generalization to new skills
 - multitask, language-conditioned policies
 - connects beautifully to internet-scale data
 - Big Questions:
 - How do we feed the data flywheel?
 - What are the scaling laws?
 - I don't see any immediate ceiling

Maybe control becomes like Vision or NLP. I won't stop trying to understand things

I think there is a handful of fruits of controls theory, but sometimes you can see better with models GCS. try to understand both and push, to get a much deeper way looking through