Mental Adjustments for Changing Autonomous Robots

Jeremy Stober
Department of Computer Sciences, The University of Texas at Austin

Introduction

As the field of robotics advances, the deployment possibilities for autonomous robots increase, along with the expected lifetimes of these robots, requiring robot components that have very low failure rates or degrade in a manner minimizes changes to robot capabilities.

In addition to dealing with degrading physical capabilities, we should expect long lived autonomous robots to be able to augment their physical capabilities by adding or replacing sensors and effectors.

Current approaches to robot control, even those that provide robust control, often depend on the assumption that the robot model is constant over the life of the robot. We identify mental adjustment mechanisms that autonomous robots can use to detect and adapt to physical changes in circumstances where physical constancy does not hold.

Example Physical Changes

- **Sensor Degradation and Augmentation**
  - *Damaged Vision System.* A vision system with camera damage may need to adjust policies to compensate for the damage.
  - *Adding a Sensor.* Visual features extracted from a new sensor need to be understood in terms of other, pre-existing sensors, and existing tasks.
- **Motor Degradation and Augmentation**
  - *Damaged Components.* Damage introduces new physical constraints. Without detection and adaption, plans that do not recognize new constraints will have a higher chance of failure.
  - *Adding Components.* With an additional effectors, the robot needs to adapt to a new action model in order to take advantage of new capabilities.

References

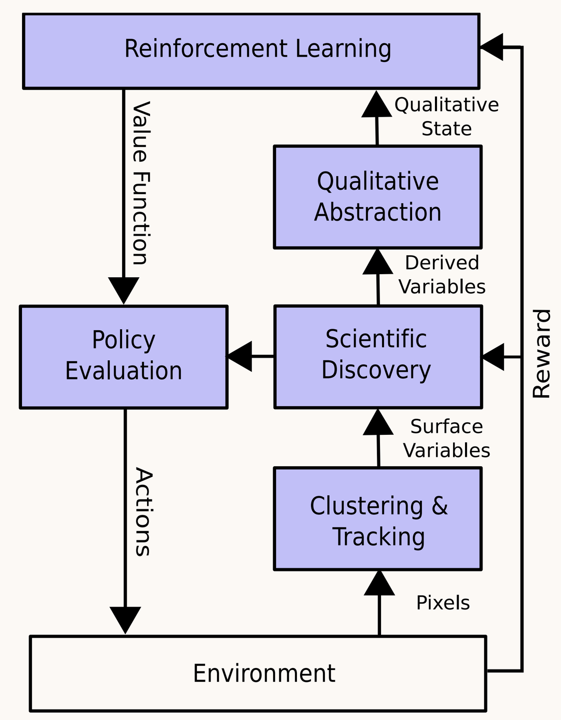
[1] J. Stober, L. Fishgold, and B. Kuipers.
Sensor map discovery for developing robots.
In *AAAI Fall Symposium on Manifold Learning and Its Applications*, 2009.

[2] J. Stober and B. Kuipers.
From pixels to policies: A bootstrapping agent.
In *7th IEEE International Conference on Development and Learning*, pages 103–108, 2008.

Autonomous Feature Discovery [2]

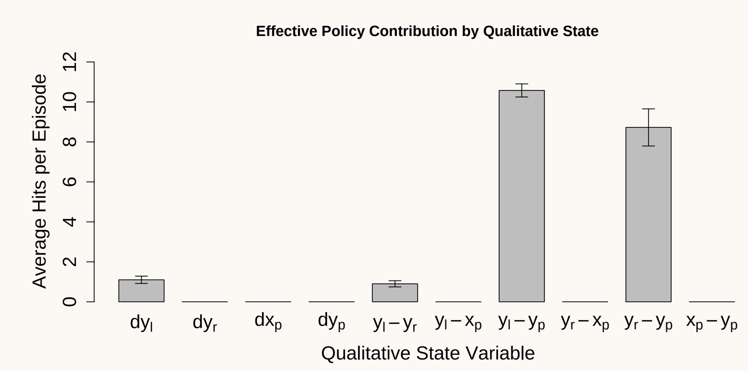
- A major issue in sensor augmentation is discovering what features, if any, a new sensor provides that would be relevant for completing a robot's assigned tasks.
- Often, in reinforcement learning or control studies, the features are carefully designed in advance, and the focus is on learning optimal control policies in terms of these pre-determined features.
- An autonomous robot, augmenting itself, may not have the benefit of an outside opinion regarding the importance of new sensor features. To address this, we present an approach to autonomous feature discovery.

From Pixels to Policies



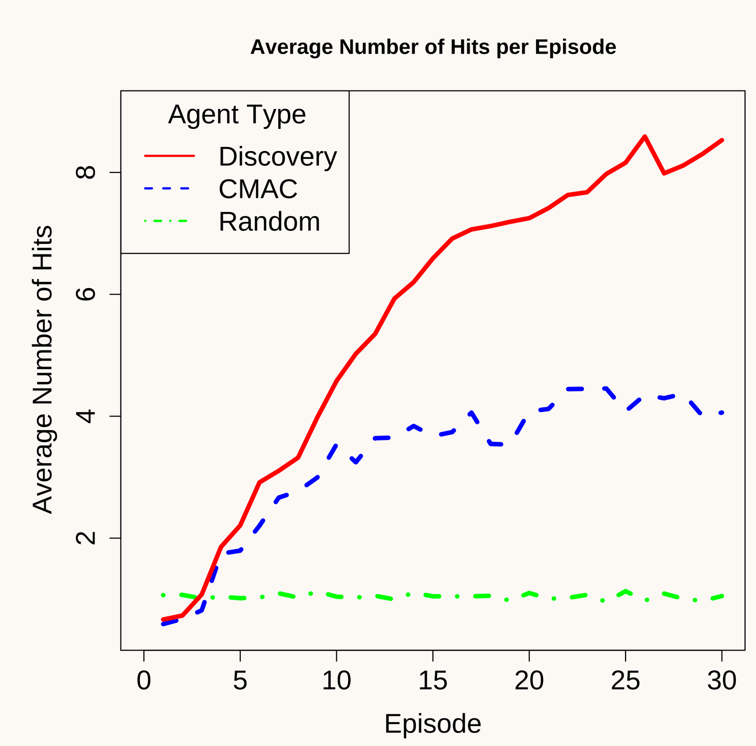
- We demonstrated a method for discovering task-relevant features by searching over expressions composed of commonly available visual properties of objects.

Generating Features



- Starting with trackers into the sensor stream, the agent uses heuristic search and qualitative abstraction to generate features for reinforcement learning.
- The agent measures the *effective policy contribution* of generated features (above plot) to determine which features should remain in the representation.

Results

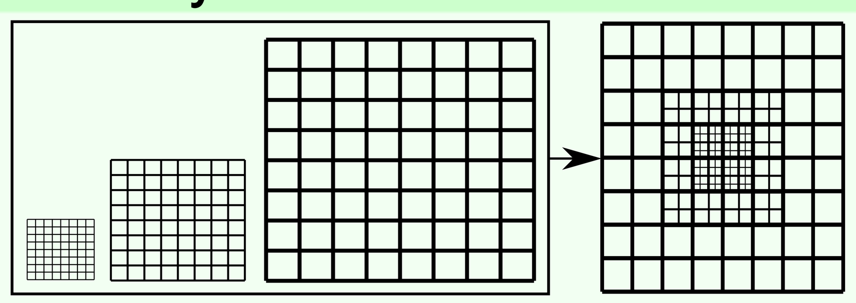


- Agents attempt to learn a simple volley task in a Pong environment.
- With autonomous feature discovery, a simple agent learns faster than when using a naive feature set.

Sensorimotor Embedding [1]

- A useful skill for an active vision system is the ability to center on a salient object in a scene.
- We show how to robustly learn this centering skill using reinforcement learning and a foveated retina.

Learning the Policy



- Each receptive field measures the saliency of a portion of a scene, $I_k \in \mathcal{I}$, and depends on the global state of the entire retina, $(\theta, \phi) \in \mathcal{S}$.
- Each receptive field implements an activation function, $\delta : \mathcal{I} \times \mathcal{S} \to [0, 1]$. Retina activation is the sum of receptive field activations

$$R_{\mathcal{I}}(s) = \sum_{I_k \in \mathcal{I}} \delta(I_k, s).$$

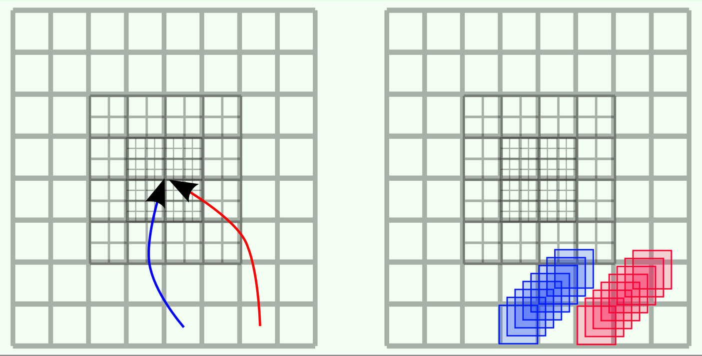
- We factor the policy decision into individual votes by receptive fields

$$\pi^*(s) = \frac{1}{\mathcal{R}_{\mathcal{I}}(s)} \sum_{I_k \in \mathcal{I}} \delta(I_k, s) \cdot \pi_k$$

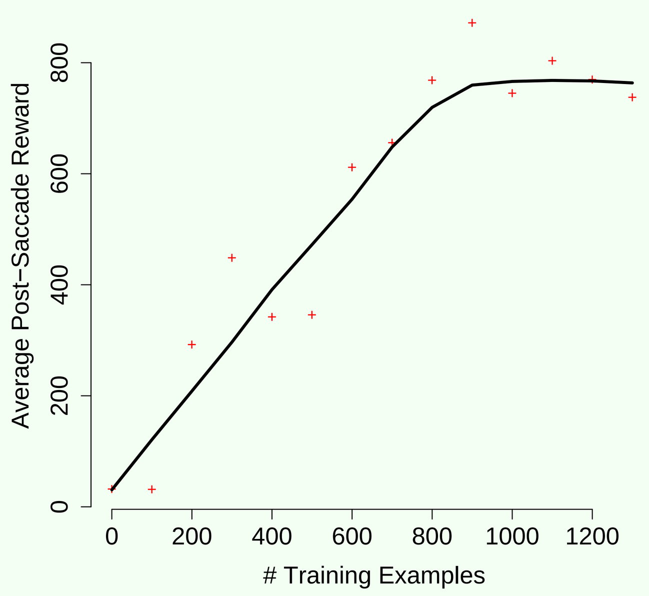
where each π_k is a constant policy learned through stochastic estimation methods.

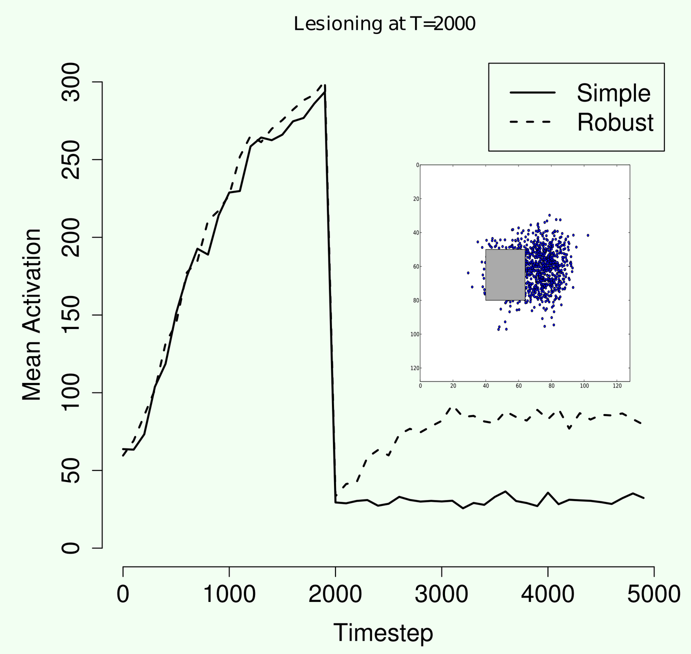
Learning the Structure

- Policy estimates encode the positions of the receptive fields (left).
- Provides an alternative to manifold learning (right).



Results





- The left and right graphs show how reward increases with training time and how this approach adapts to lesioning.