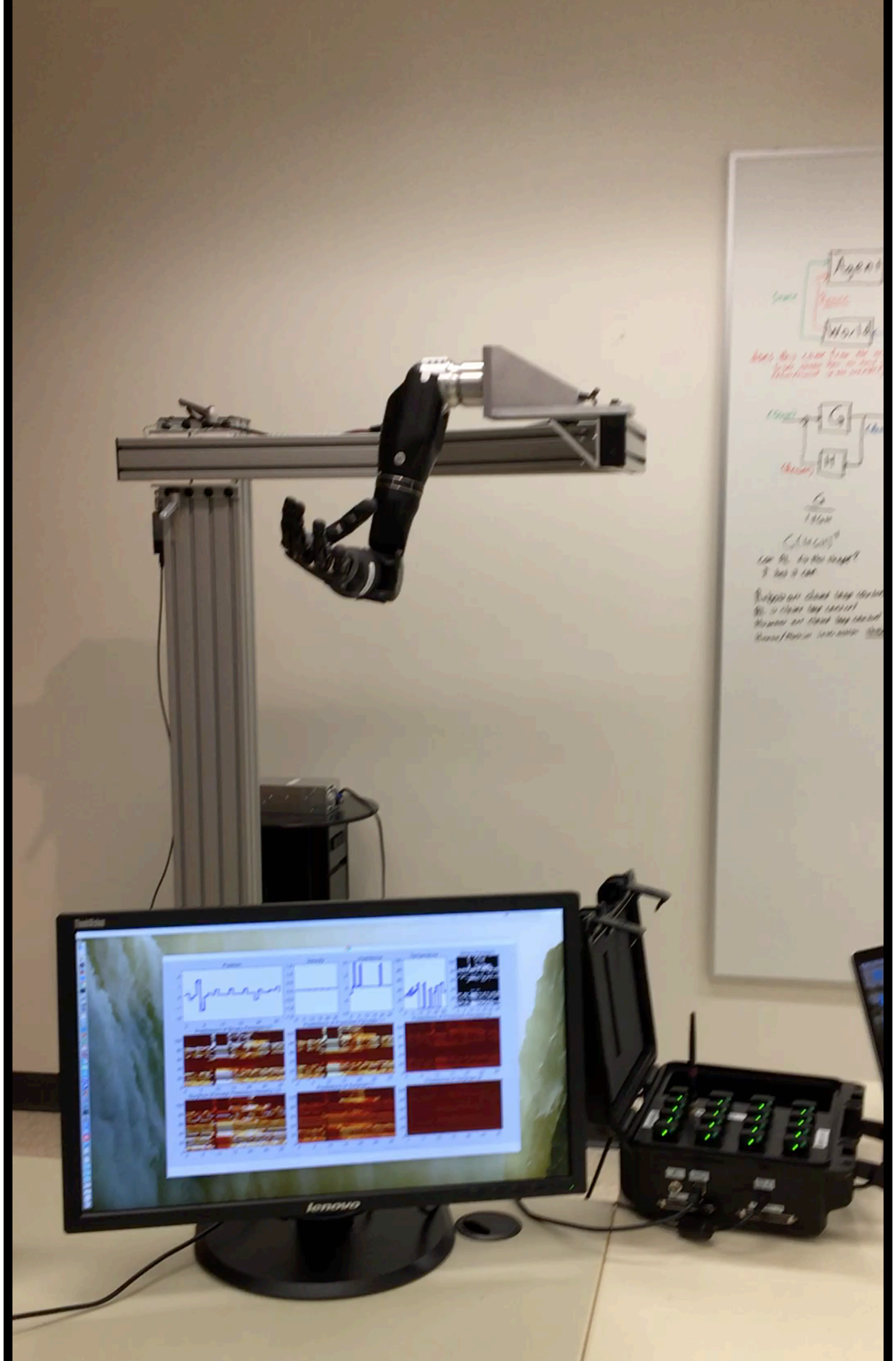
Reinforcement Learning (Without the Math)

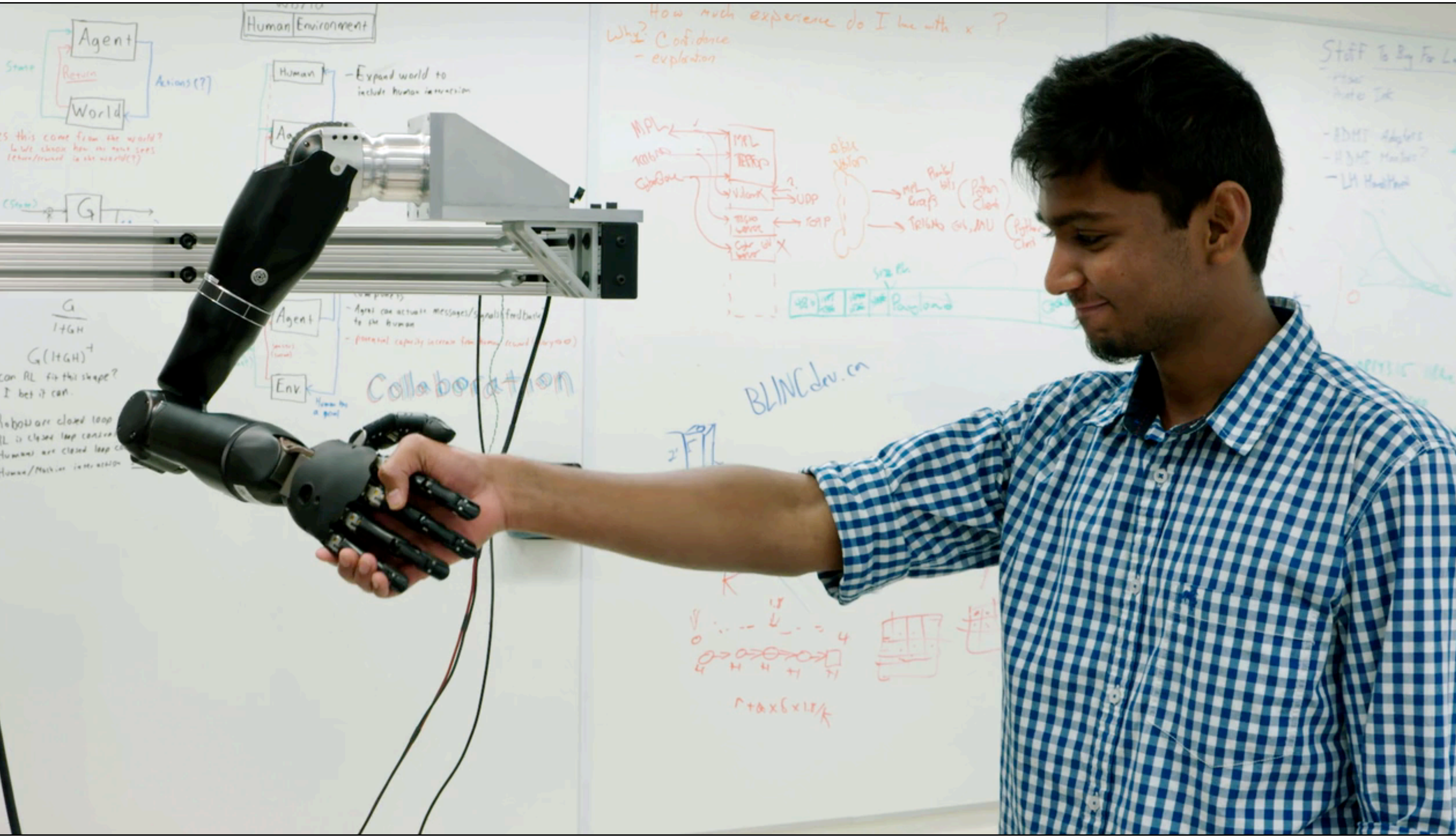
Patrick M. Pilarski, Ph.D.
Canada Research Chair & Assistant Professor

Division of Physical Medicine and Rehabilitation, Dept. Medicine
PI, Alberta Machine Intelligence Institute (Amii)
PI, Reinforcement Learning and Artificial Intelligence Laboratory

UNIVERSITY OF ALBERTA
EDMONTON · ALBERTA · CANADA



Pilarski Lab
Jan. 2017



Pilarski Lab
August 2016

Why Intelligence?

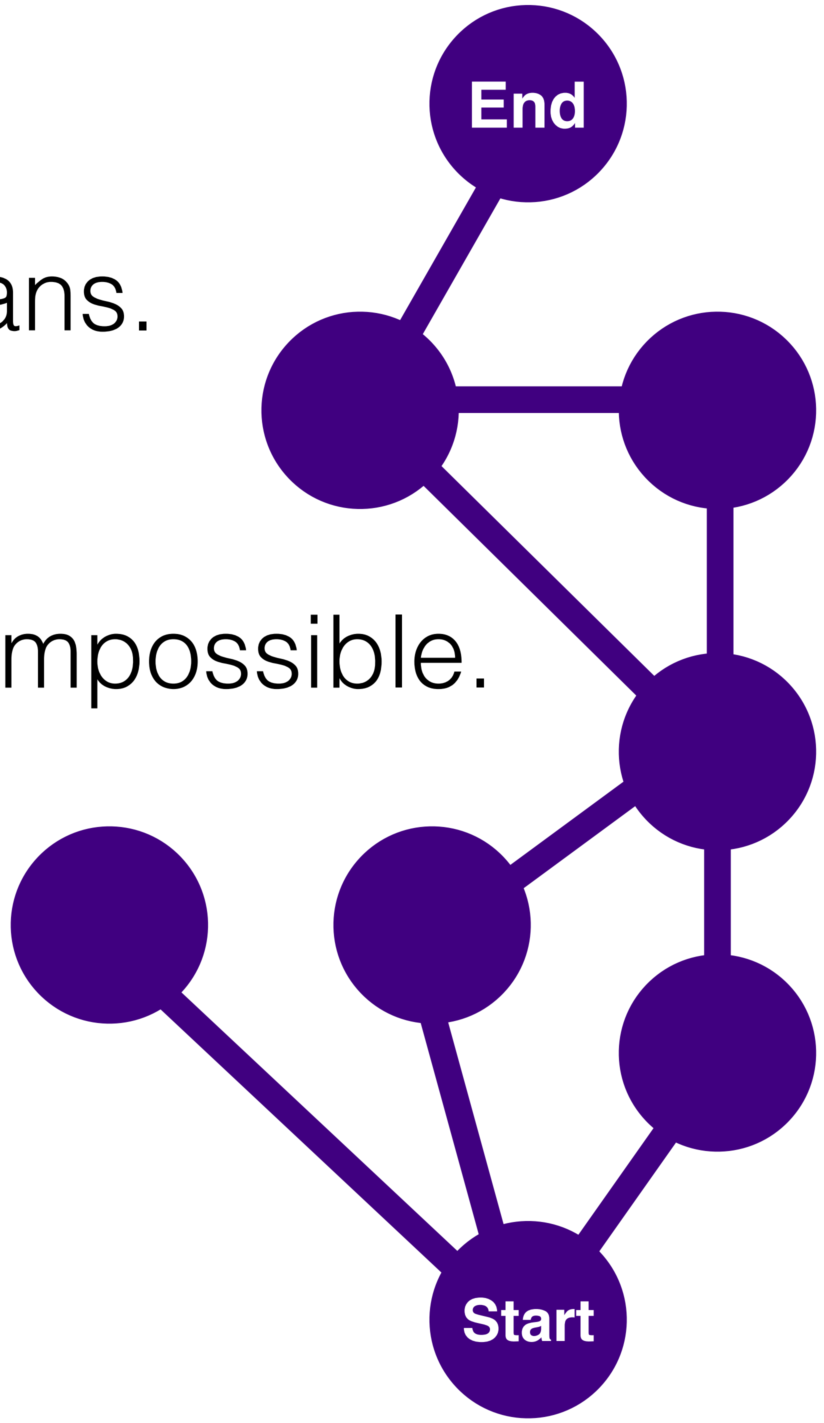
- **Enhanced control** over a changing and increasingly complex world.
- **Anticipation** of future events and outcomes.
- **General tools** for solving hard problems.
- "Optimizing the control of complex systems and extracting knowledge from massive amounts of data."
- Examples: finance, healthcare, energy, resources, transport, information processing.

Why Learning?

- **Things are Unknown:** known ends but unclear means.
- **Things are Complex:** scaling up is demanding or impossible.
- **Things Change:** systems need to adapt!

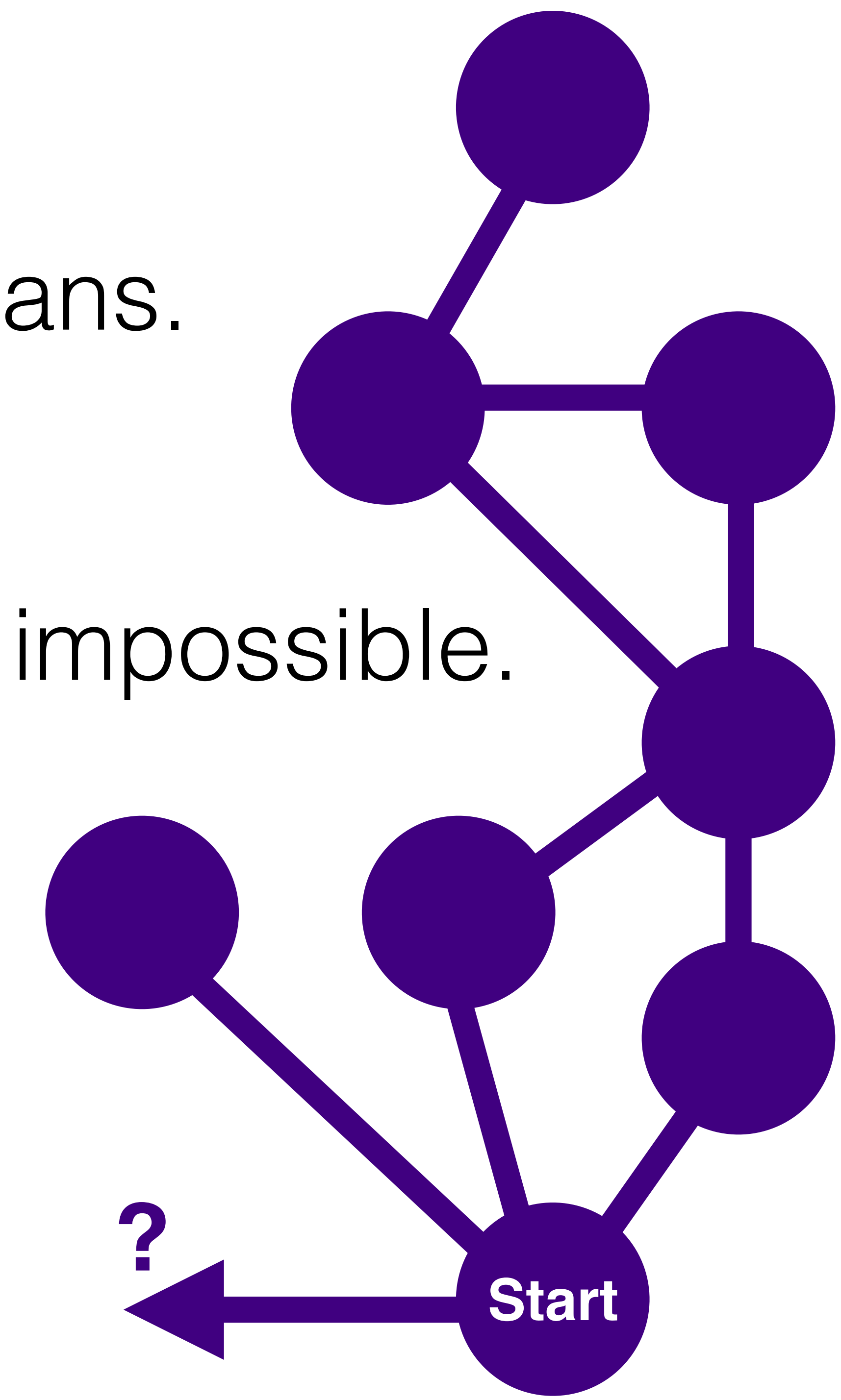
Why Learning?

- **Things are Unknown:**
  known ends but unclear means.
- **Things are Complex:**
  scaling up is demanding or impossible.
- **Things Change:**
  systems need to adapt!



Why Learning?

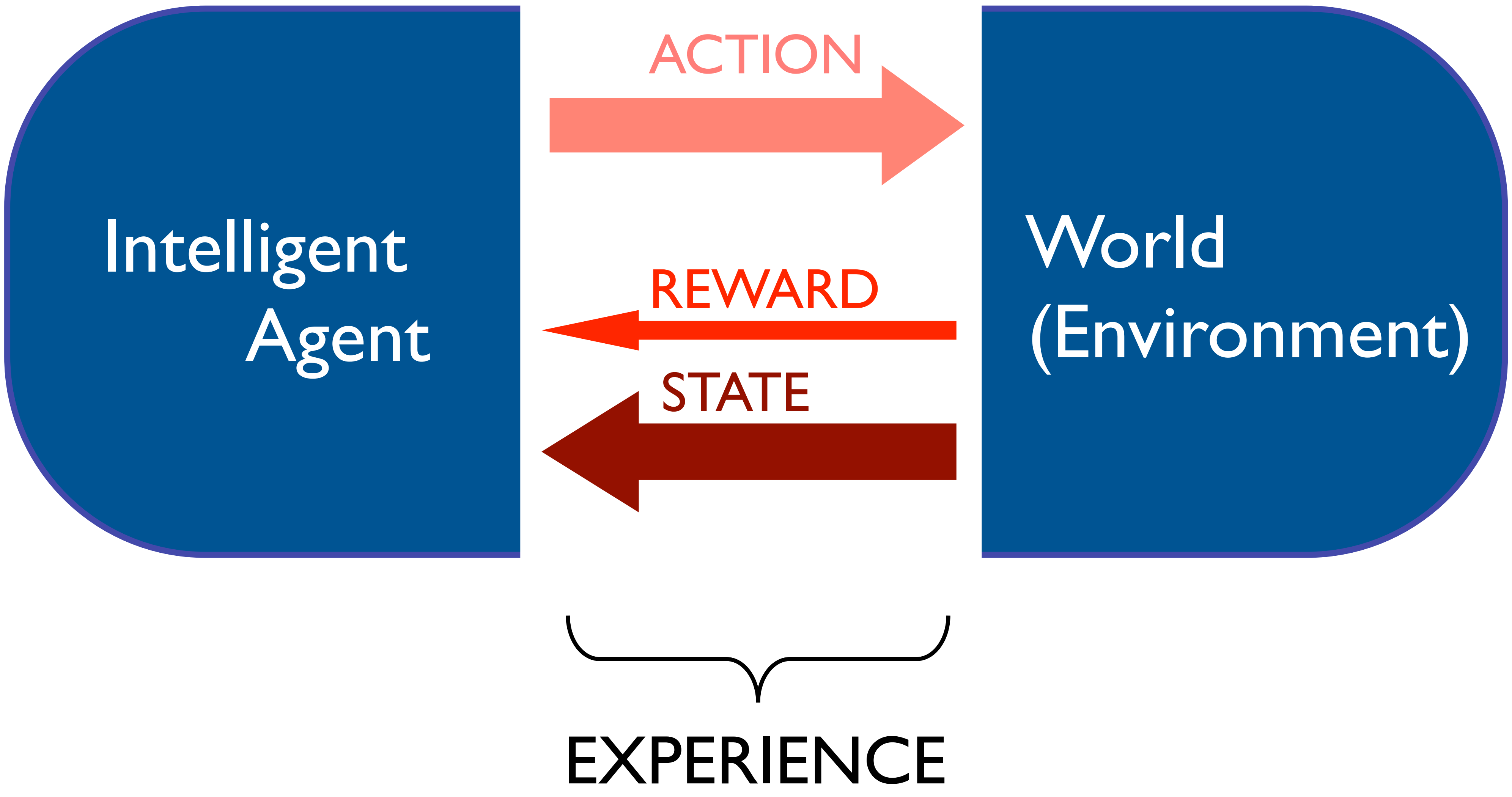
- **Things are Unknown:** known ends but unclear means.
- **Things are Complex:** scaling up is demanding or impossible.
- **Things Change:** systems need to adapt!



KEY IDEA

Our ability to directly engineer an intelligent system no longer scales up to our goals or to the complexity of our problems of interest.

Reinforcement Learning



Reinforcement Learning is an approach to:

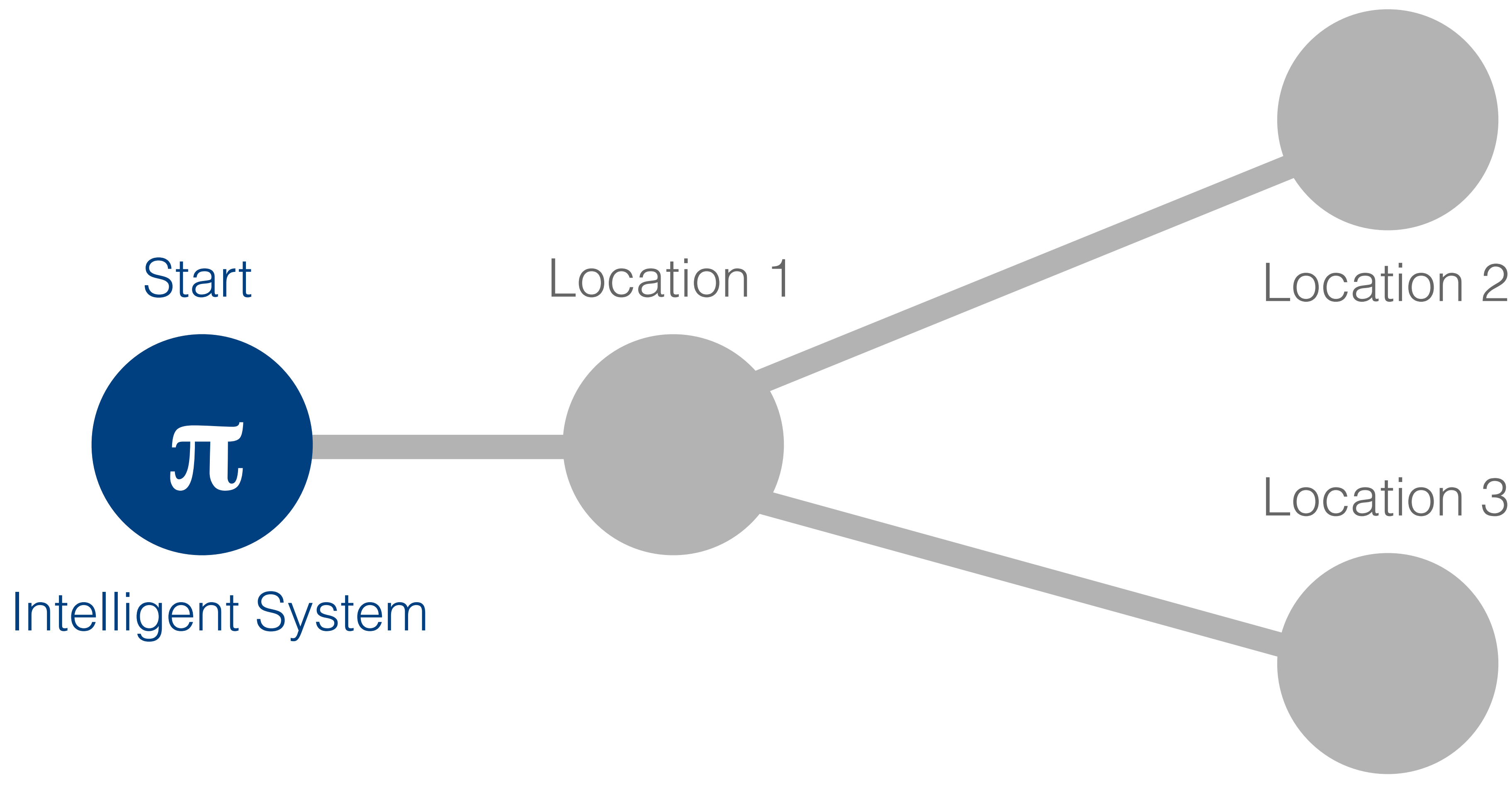
- Natural intelligence
- Artificial intelligence
- Optimal control
- Operations research
- Solving partially observable Markov decision processes

(and the perspective that all of these are the same)

RL Headlines

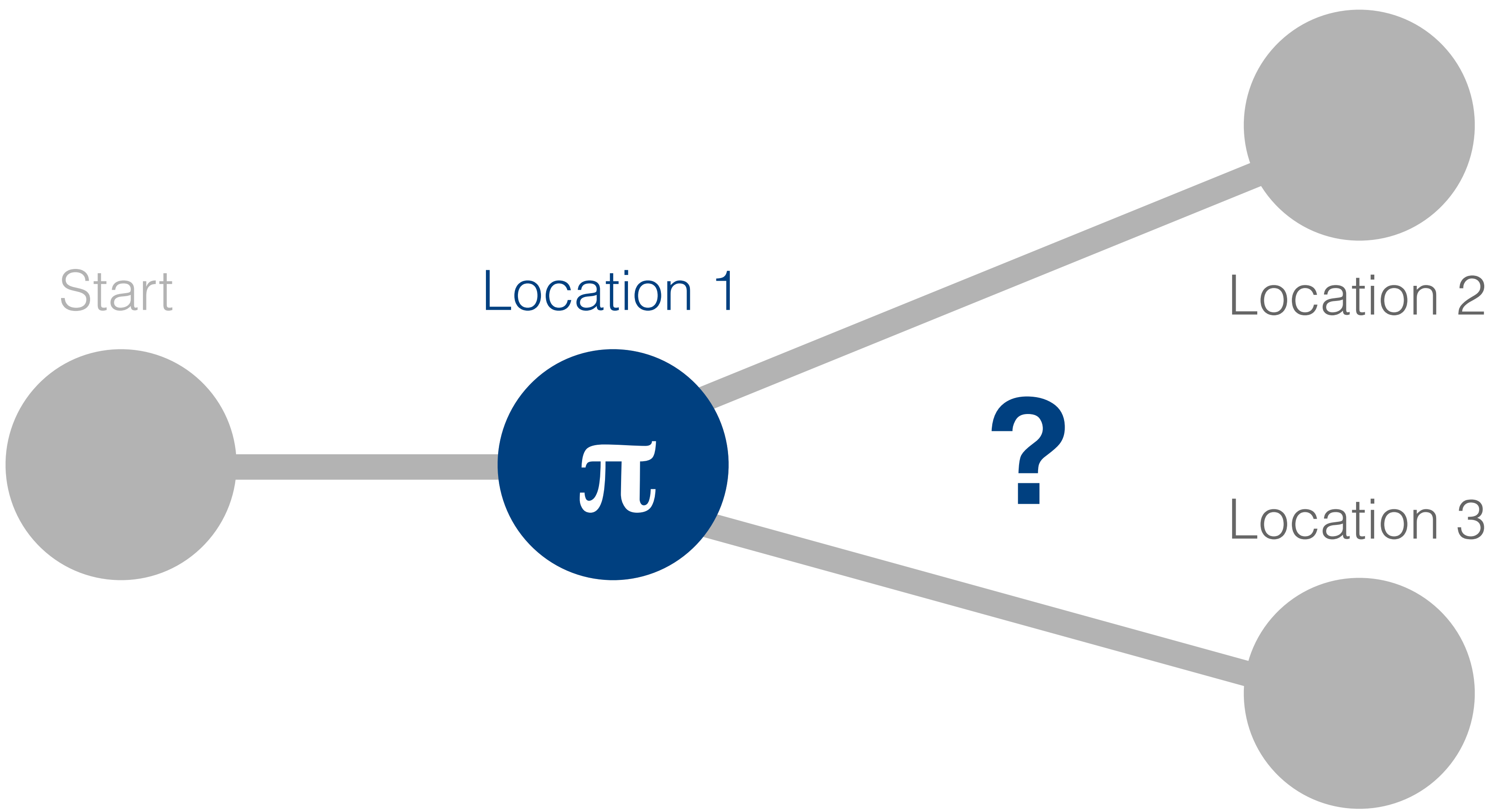
- RL is widely used in robotics
- RL algorithms have found the best known approximate solutions to many games
  (RL is part of the revolution in solving Go)
- RL algorithms are now the standard model of reward processing in the brain
- High-impact combinations of RL algorithms with deep neural networks.

Decision Making



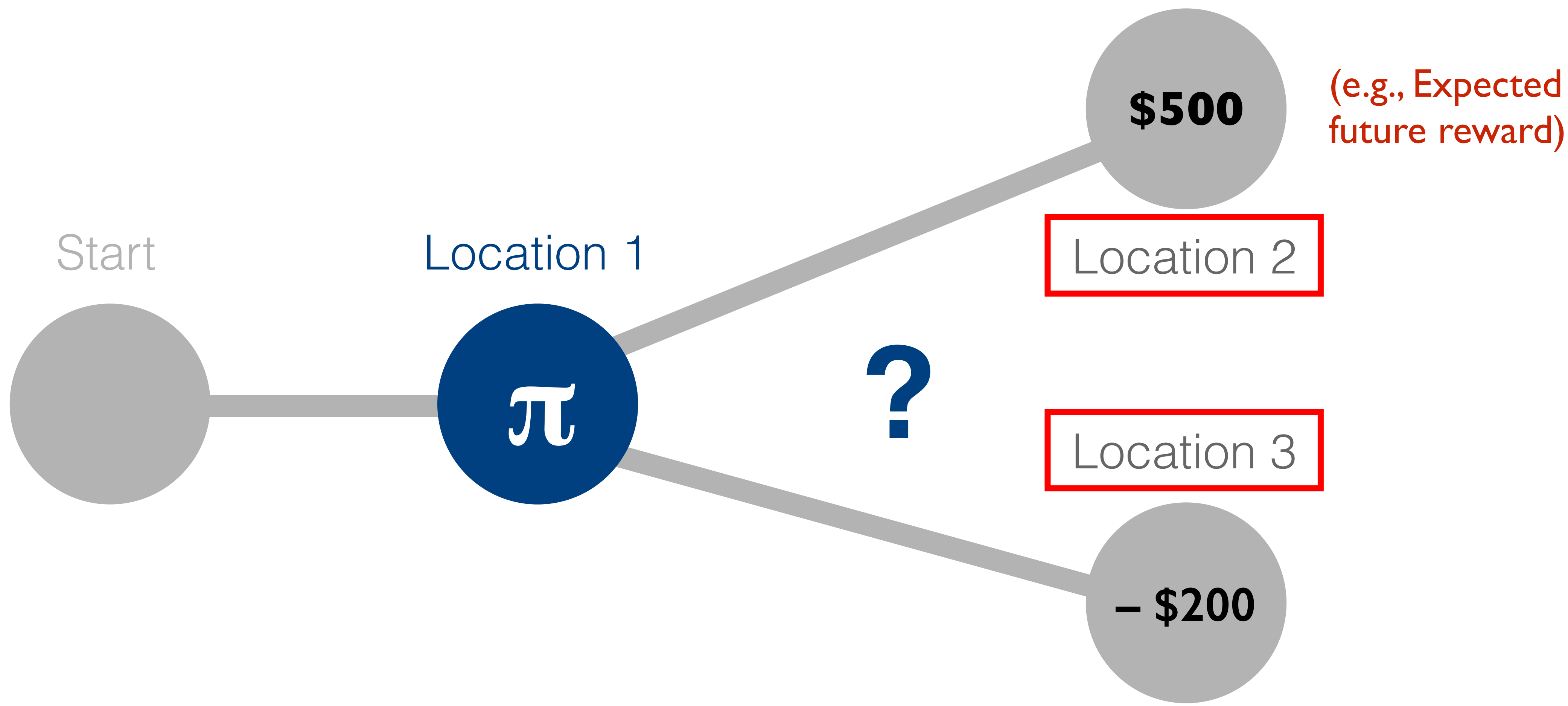
How to act so as to maximize reward?

Decision Making



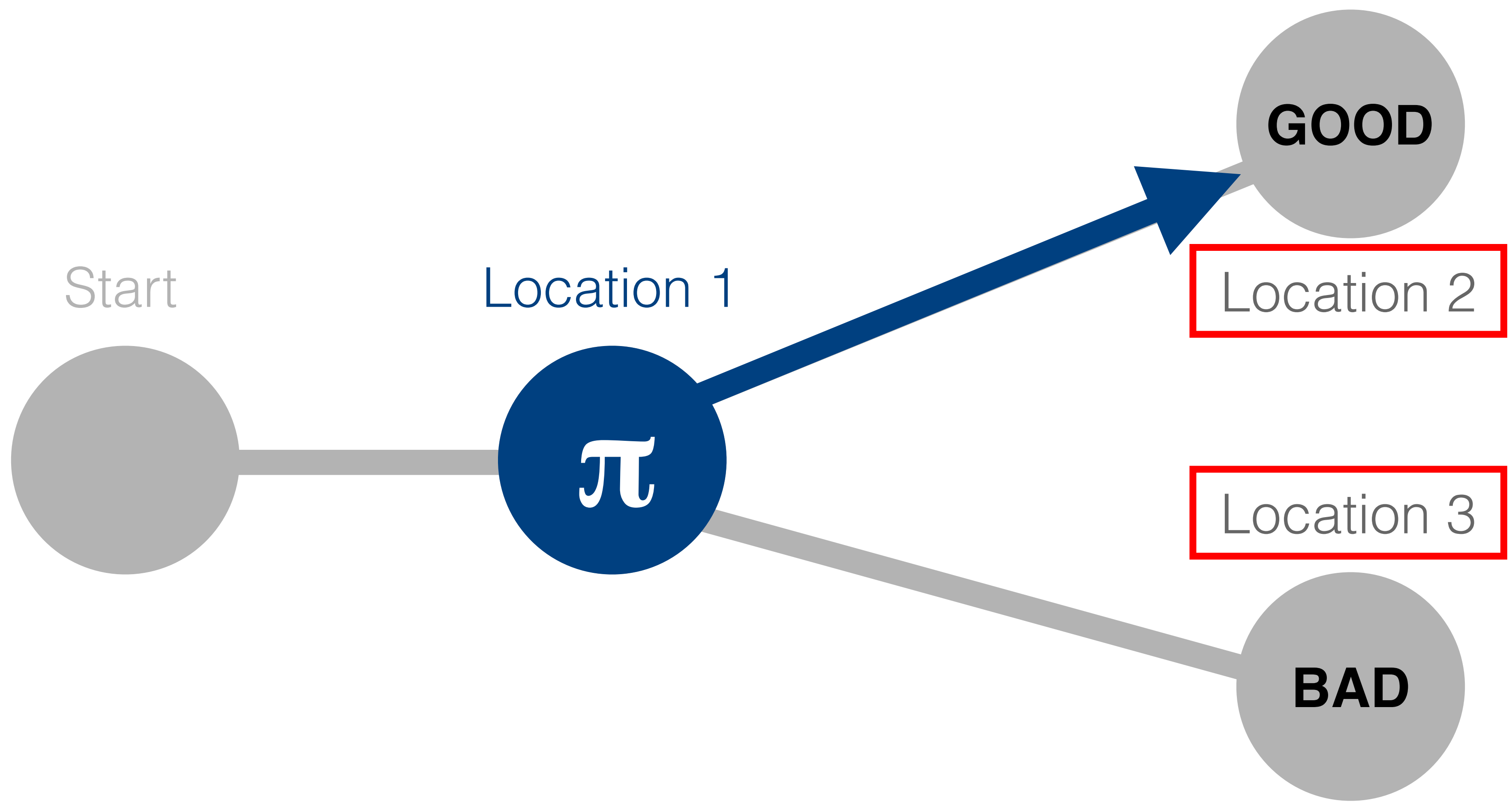
Using a learned model or values . . .
(that can be learned directly from experience)

Decision Making



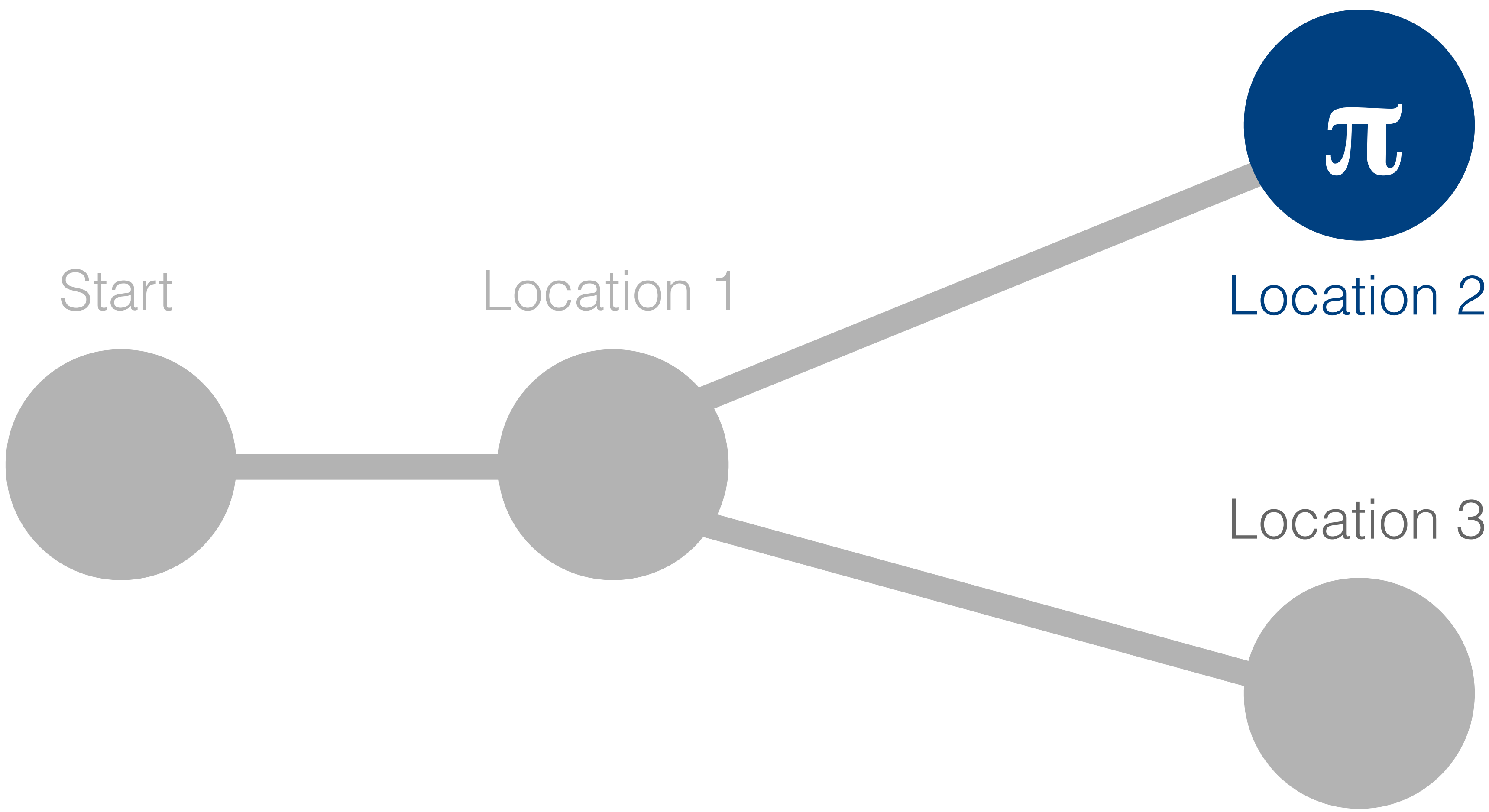
Using a learned model or values . . .
(that can be learned directly from experience)

Decision Making



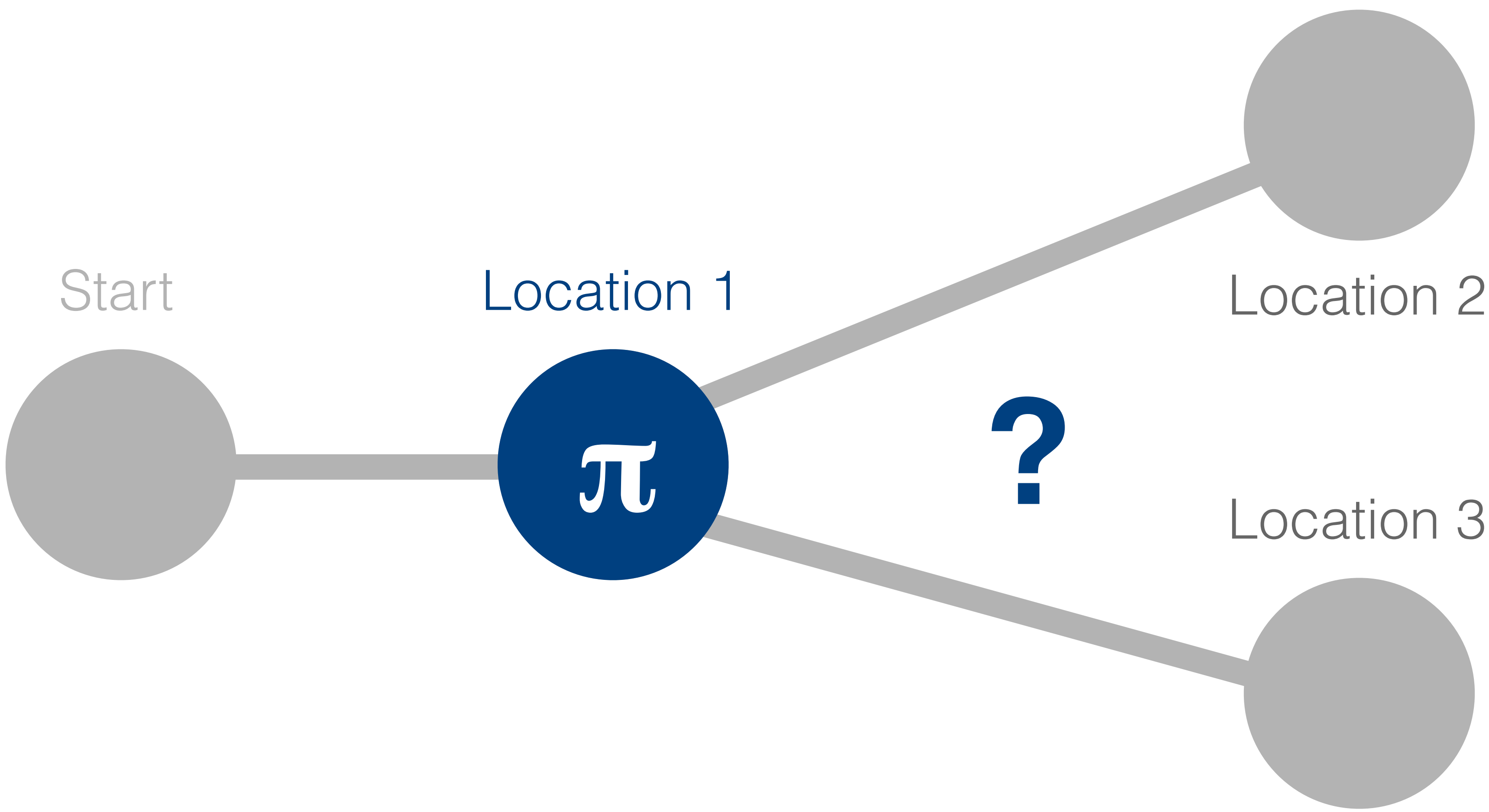
Using a learned model or values . . .
(that can be learned directly from experience)

Decision Making



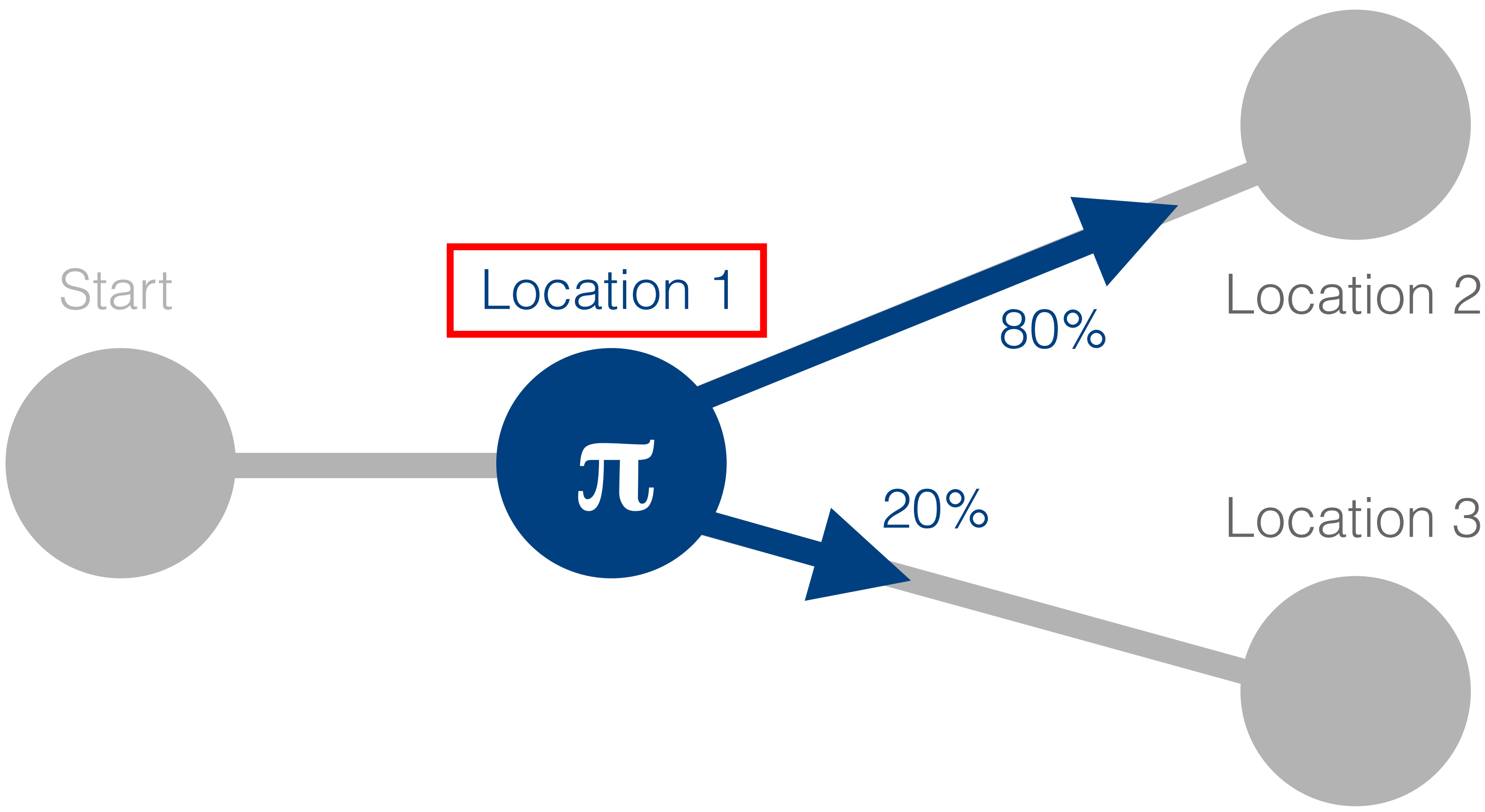
Using a learned model or values . . .
(that can be learned directly from experience)

Decision Making



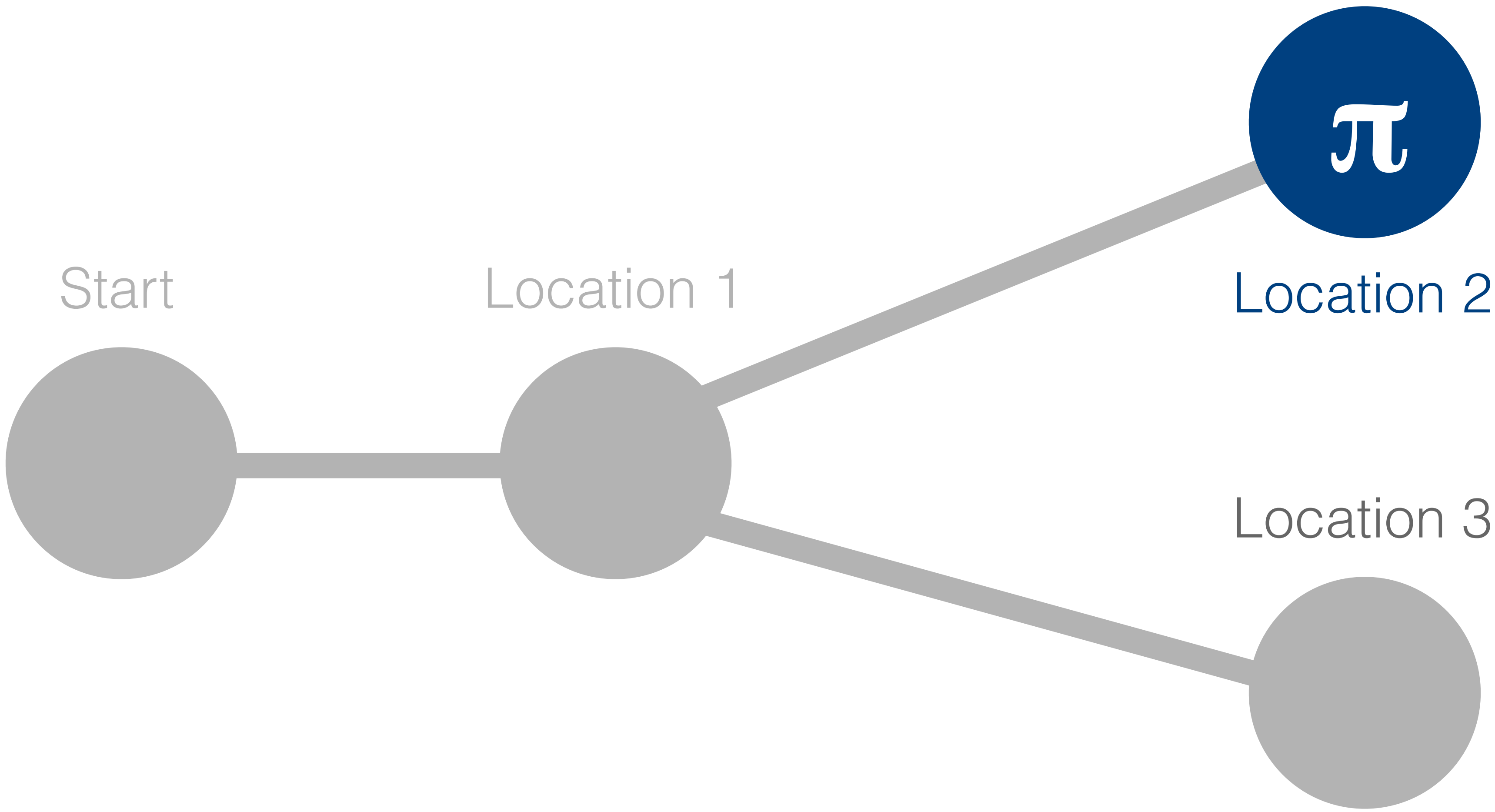
By using reward to alter and improve a policy.

Decision Making

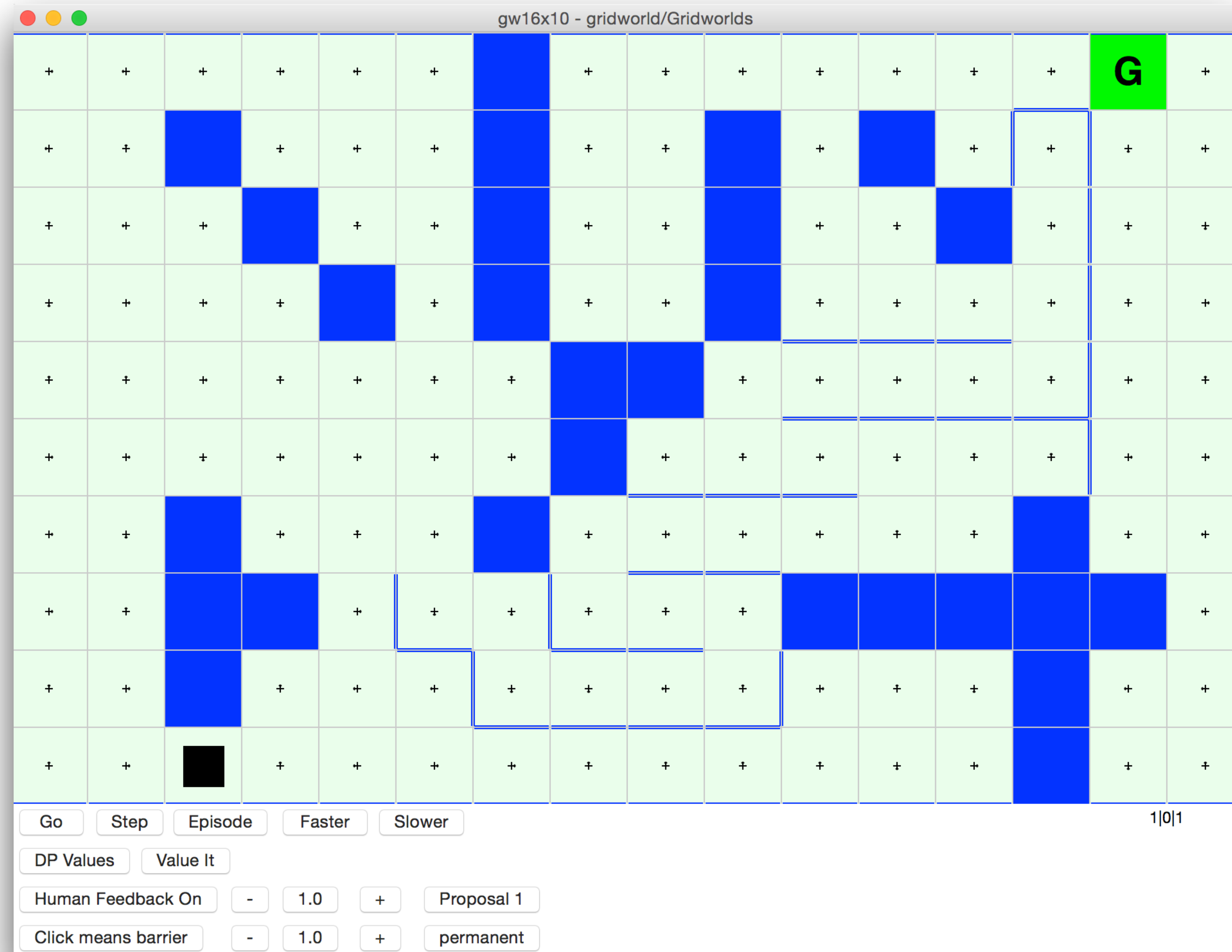


By using reward to alter and improve a policy.

Decision Making



By using reward to alter and improve a policy.



Demonstration

Exercise 1: On-policy Prediction Learning

How do we learn values of states and actions?

Exercise 2: Control Learning

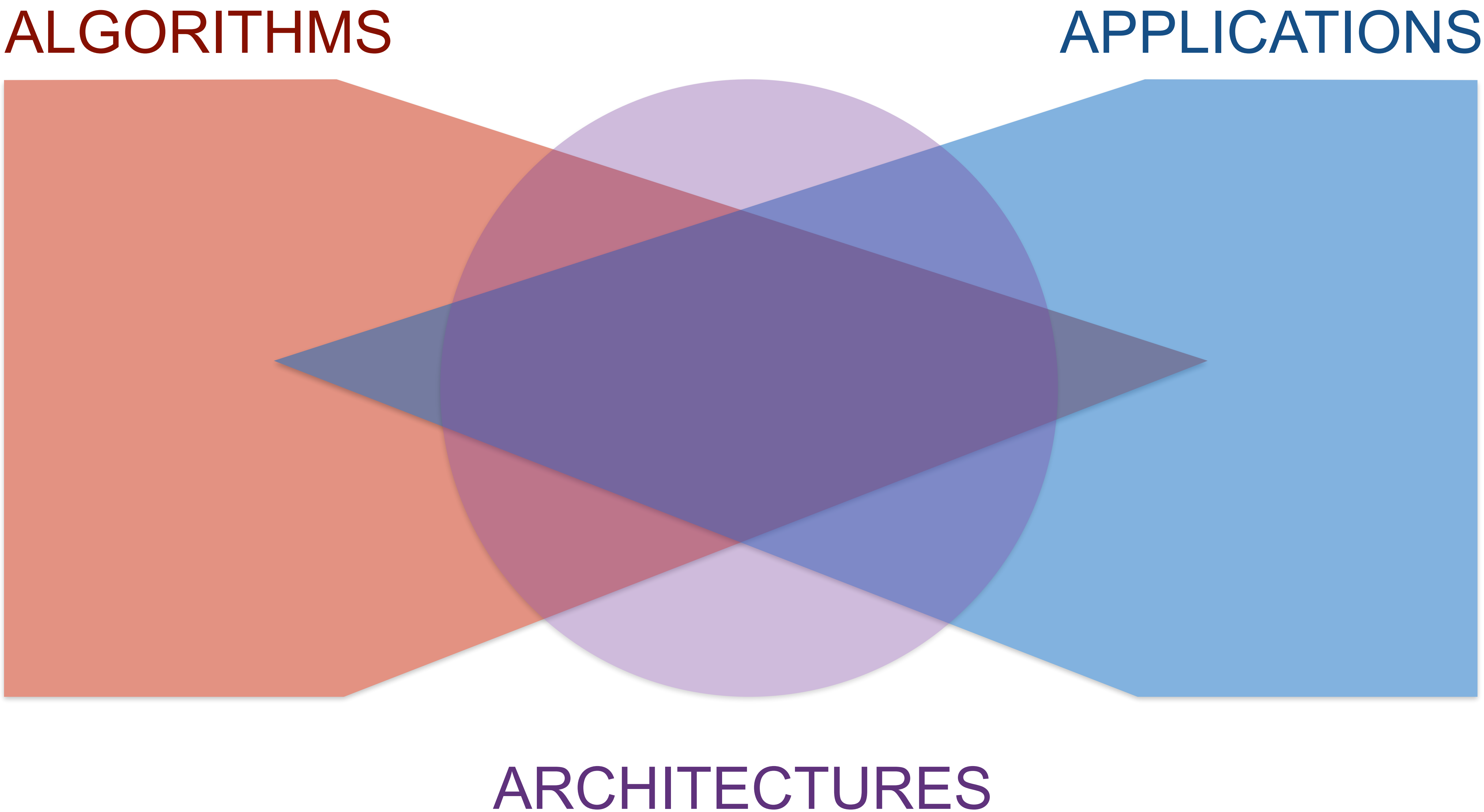
How can reward change how we behave?

Exercise 3:
Control Learning from Good and Bad

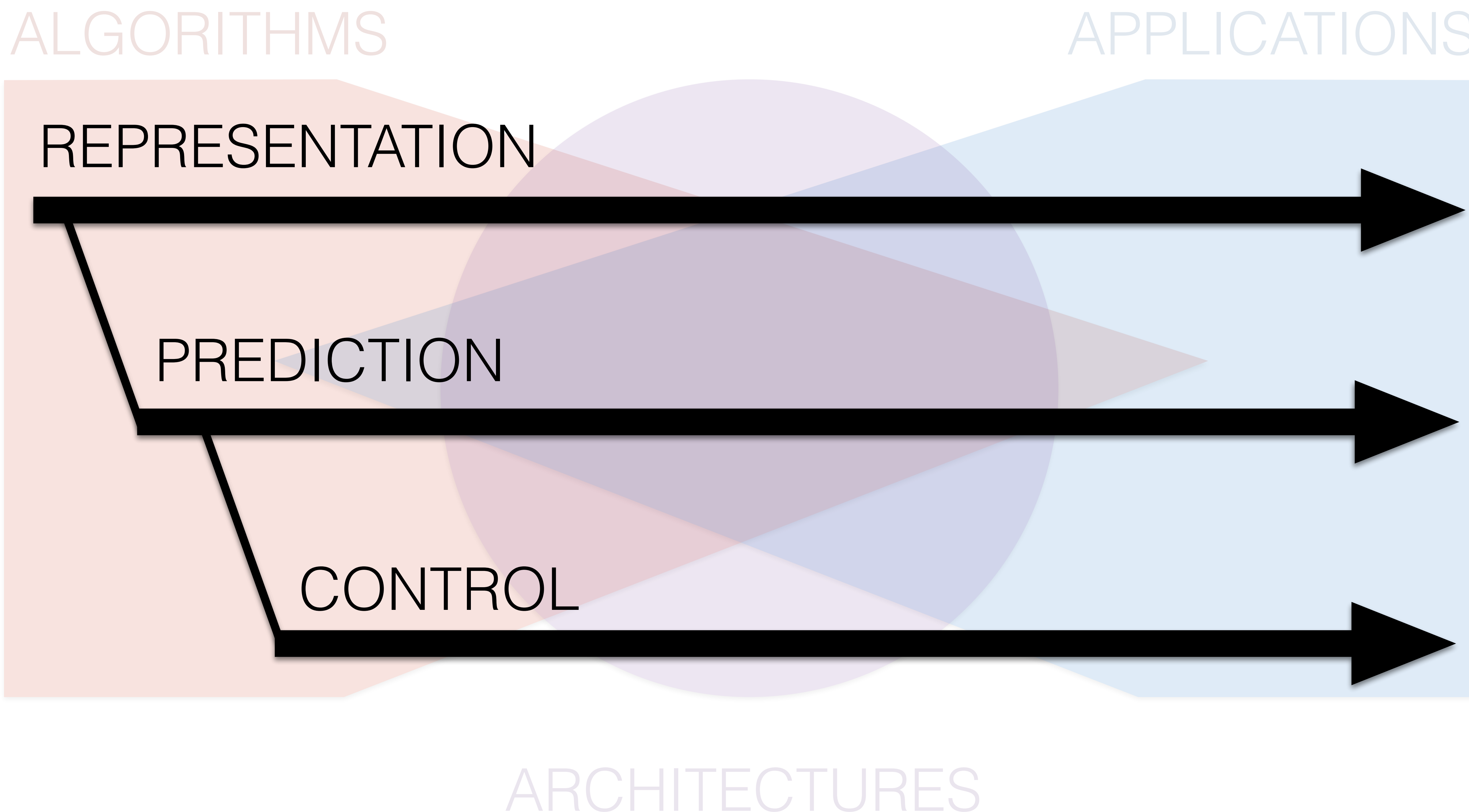
Dolphin training!
(Thanks Kory!)

Exercise 4:
Predictions + Reward

Predictive representations of state, a.k.a. you throw stuff at me.

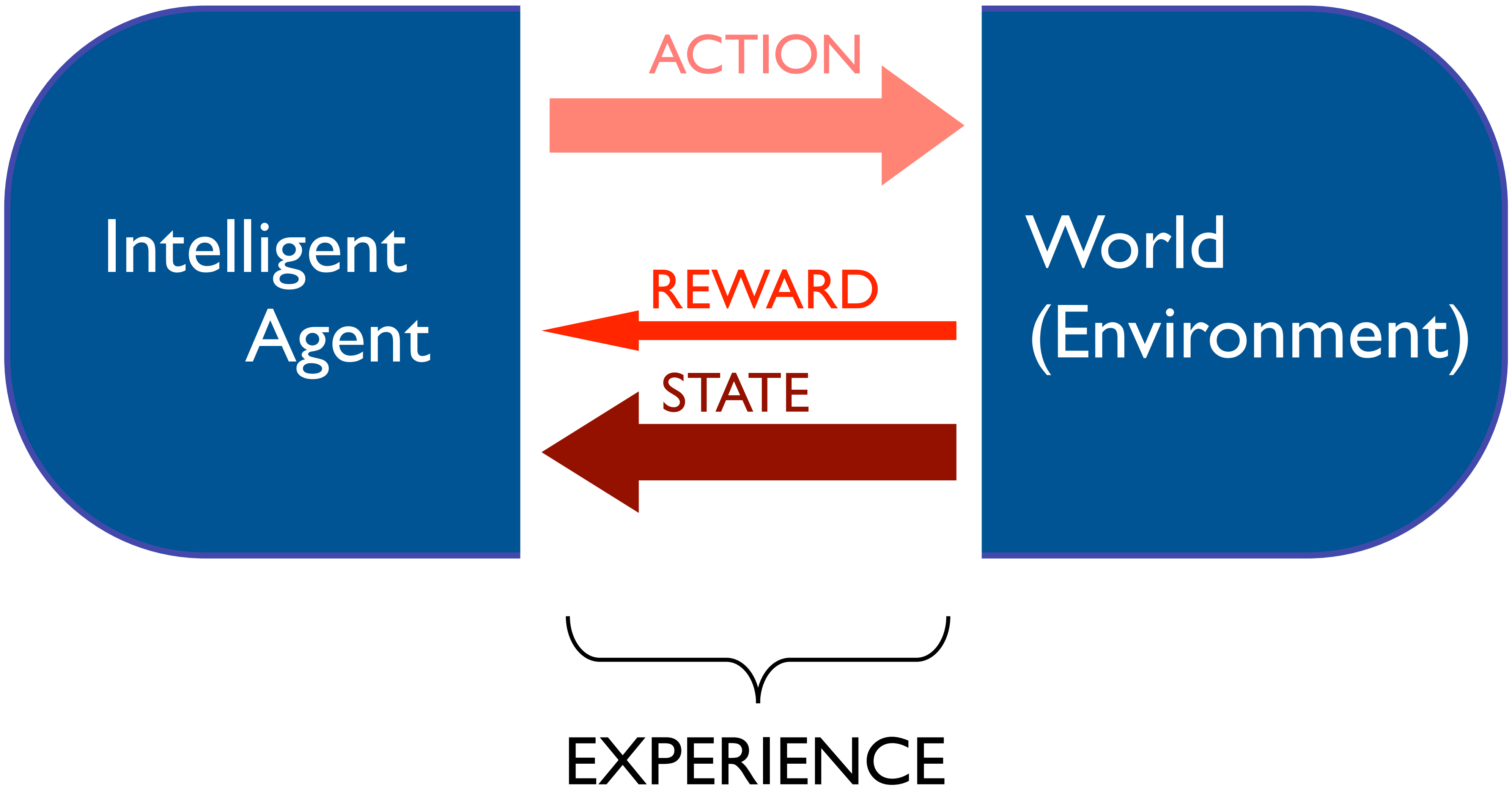


Machines need to make **good decisions** in unknown, uncertain, or changing environments.



These are all forms of **knowledge**
and leverage a process of **continual learning**

Reinforcement Learning



Questions

... and thank you very much!

pilarski@ualberta.ca

http://www.ualberta.ca/~pilarski/