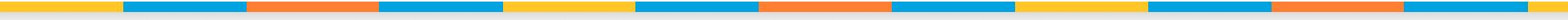

Reinforcement Learning

Heni Ben Amor, Ph.D.
Assistant Professor
Arizona State University



Introduction to Reinforcement Learning



- | **Reinforcement learning (RL)**
- | **Type of machine learning**
- | **Focuses on agent autonomy, trial-and-error**
- | **Teacher only provides feedback: “good vs. bad”**

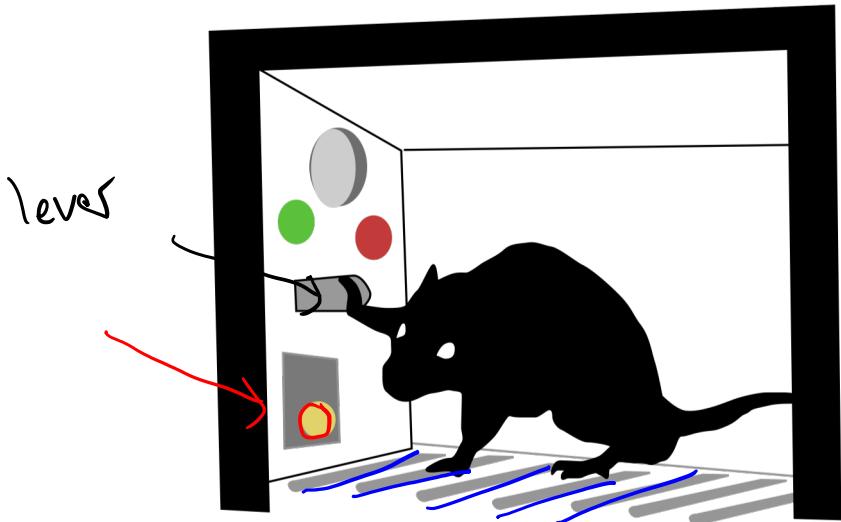
AI Successes with RL



- | Backgammon, Tesauro 1992
- | Atari games, DeepMind 2013
- | Playing Dota 2 at competitive levels,
OpenAI 2017/2018
- | Defeating world champion in Go,
DeepMind 2016

Conditioning

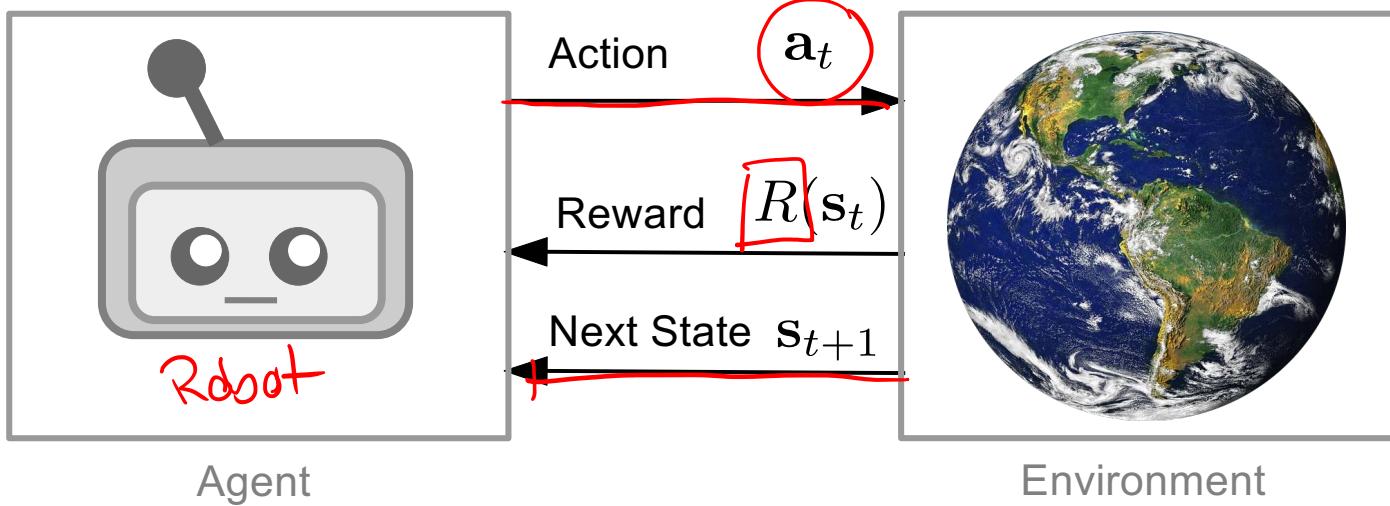
- | Learning in animals and humans
- | Operant Conditioning according to Skinner
- | Learning through reward and punishment
- | Relationship between behavior and consequence
- | Behaviors that are reinforced, are repeated



Skinner Box:
Rat is rewarded by food pellet when lever is pressed.

Reinforcement Learning

- | Learning through trial and error
- | Agent repeatedly interacts with environment
- | In each state s_t the agent chooses an action a_t
- | Reward is provided by a reward function R



Learning a Policy

- | Goal in reinforcement learning
- | Policy that takes states and produces actions
- | Learn a policy that maximizes sum of rewards
- | Reward improves as learning progresses
- | Ideally, learn with few trials
- | Sample-efficiency

$$a_t = \pi(s_t)$$



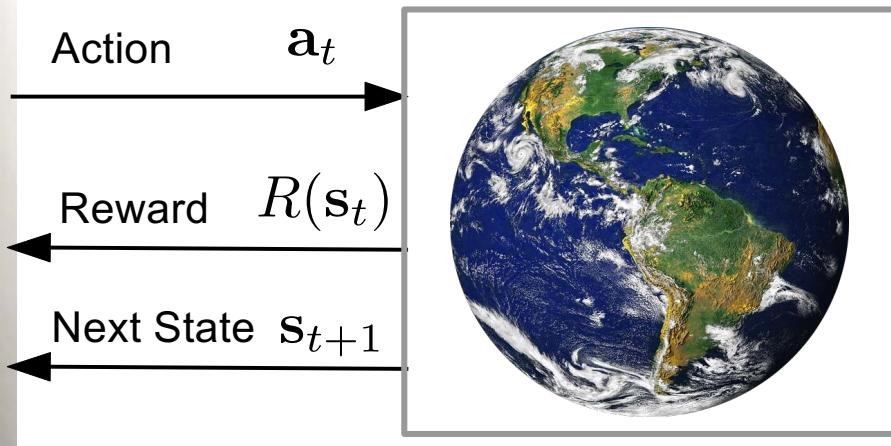
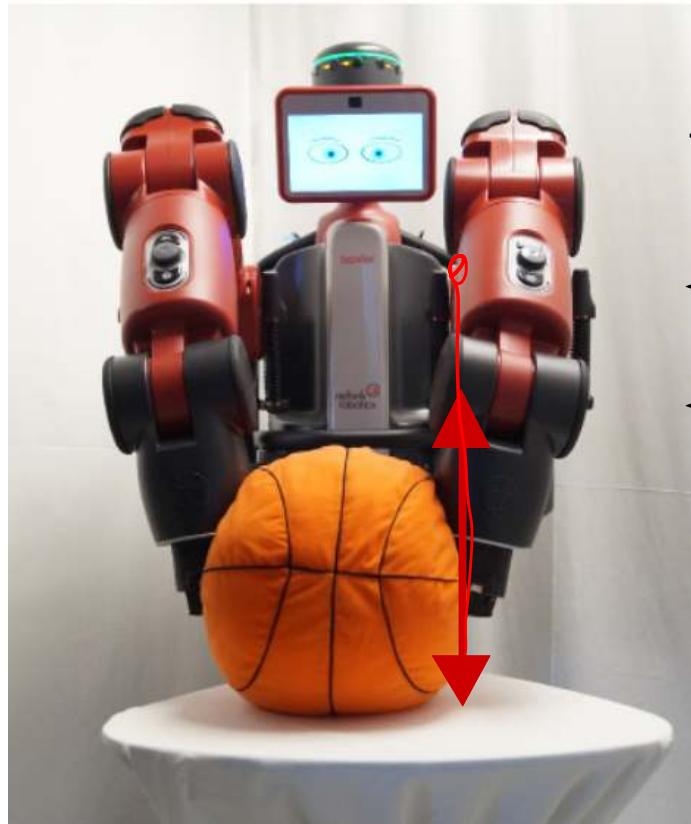
Agent

Reward Functions



- | Evaluates the current situation $R(s_t)$
- | High values if good state, low values if bad
- | Guides learning toward good policies
- | Typically manually specified by the user
- | Not always easy to do

Example: Learning to Lift an Object



Environment

Reward Function:

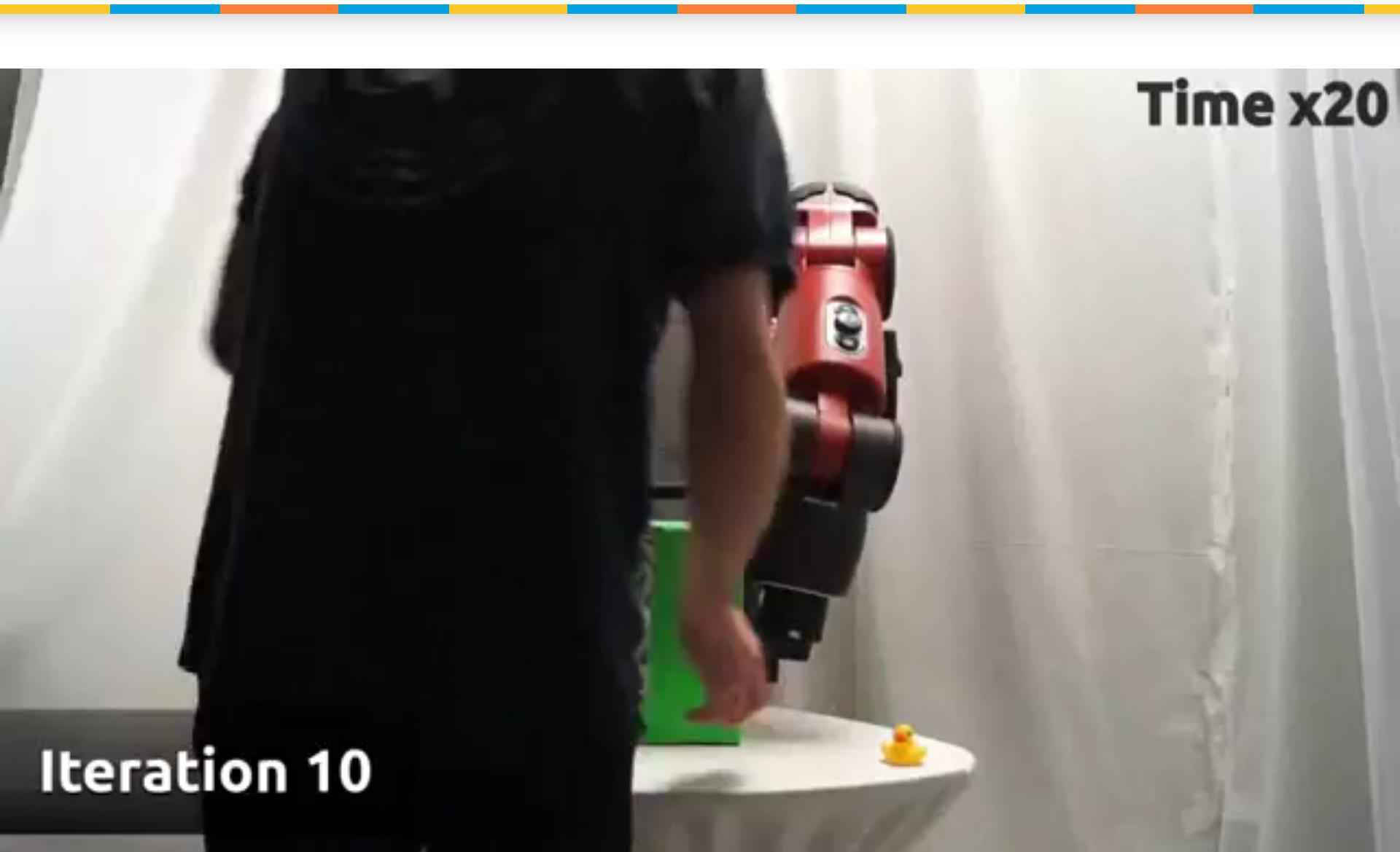
Reward function aims at maximizing the height of the object. The higher, the better.

Example: Learning to Lift an Object

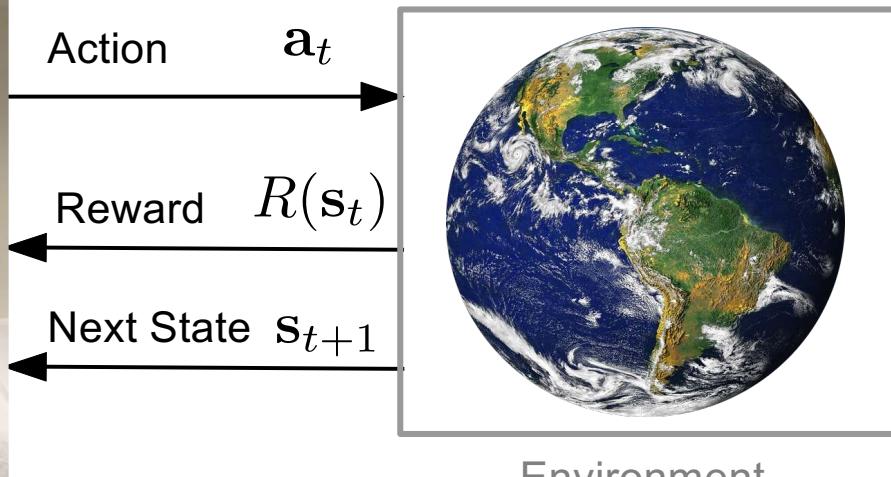
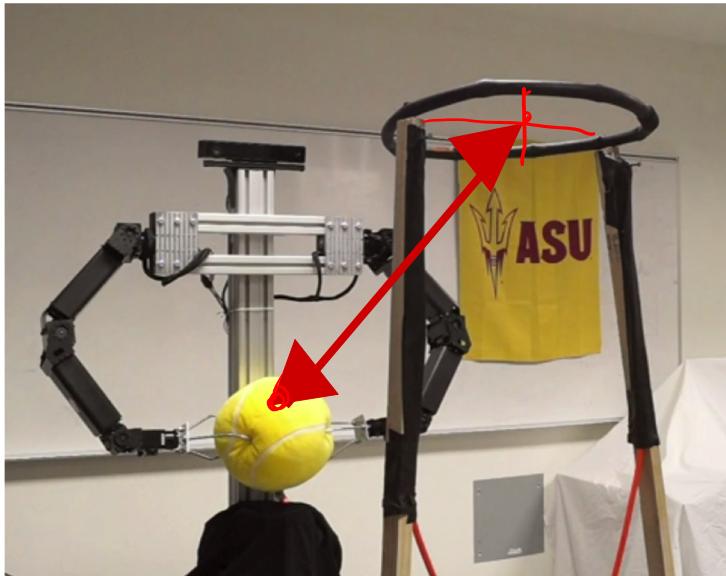


Iteration 1

Example: Learning to Lift an Object



Example 2: Throwing a Ball into a Hoop



Reward Function:

Distance to the center of the hoop.

Example 2: Throwing a Ball into a Hoop

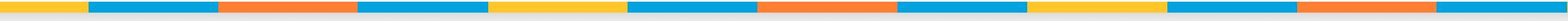
Iteration 1 (Speed 2x) 



Example 2: Throwing a Ball into a Hoop



Challenges in Reinforcement Learning



- | How to learn from few trials
- | How to specify reward functions
- | How to learn in a safe way and avoid harm
- | How to learn to solve multiple tasks
- | How to understand a learned policy