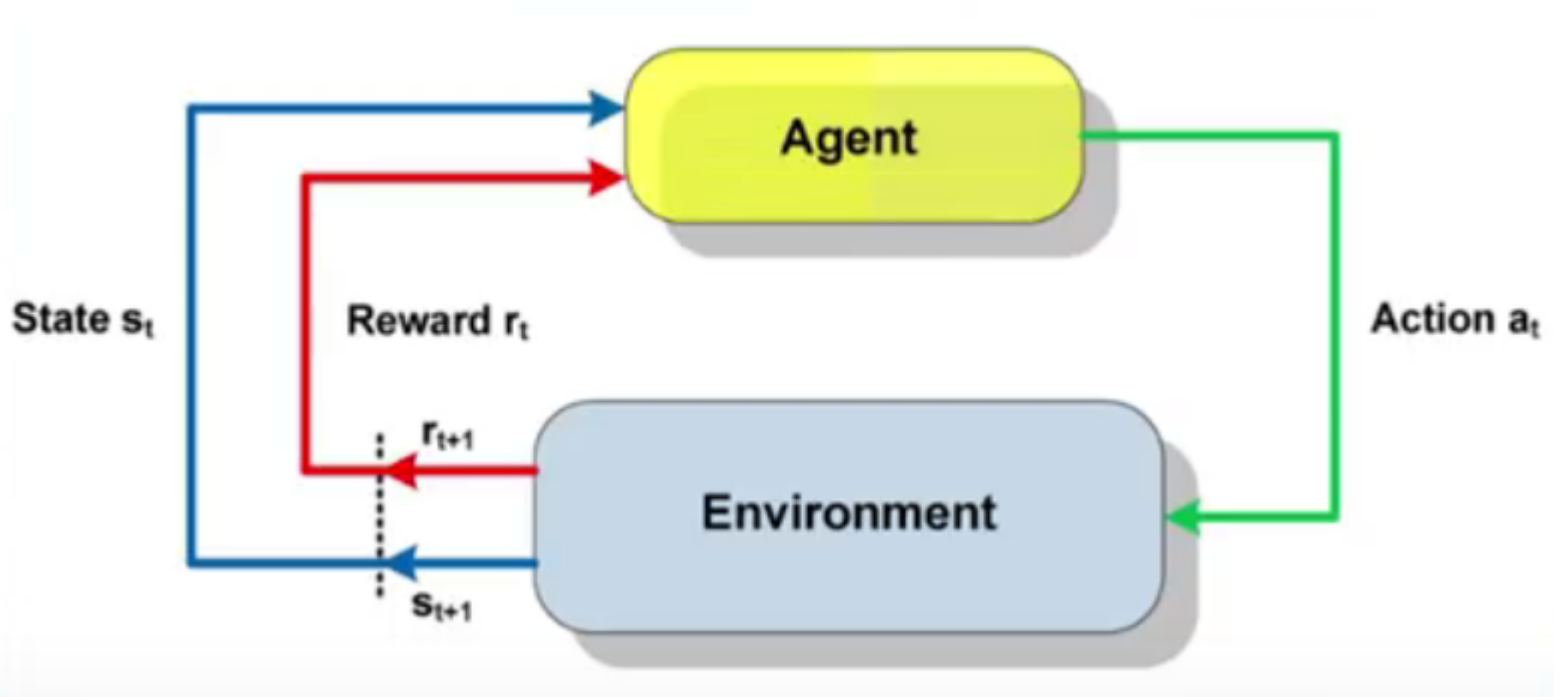
CS294:
Solving system problems with RL

Joey Gonzalez and Ion Stoica

April 1, 2019

Reinforcement Learning

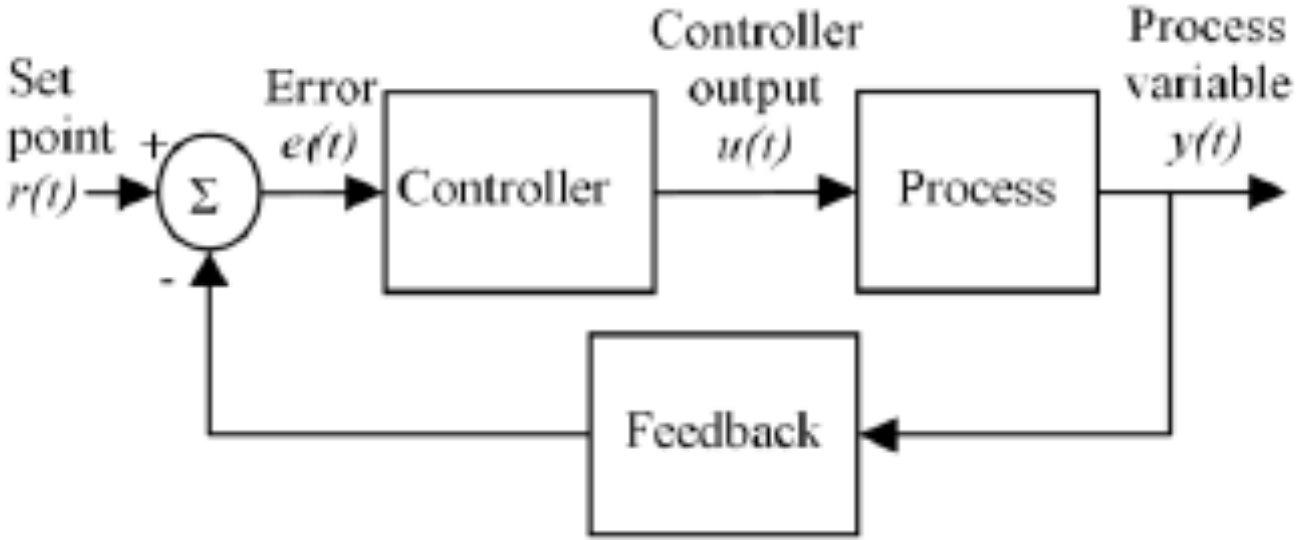


*https://sergioskar.github.io/Reinforcement_learning/

Many systems problems fit this pattern

	TCP	**Video bitrate adaptation**	**Job scheduling**
Objective (reward)	Max throughput	Quality of Experience (QoE)	Max throughput / Min response time
Control	Window size	Bitrate	Next job to schedule
Environment	Losses, etc	throughput, buf size, file size	System utilization, job characteristics, etc

Classic control systems

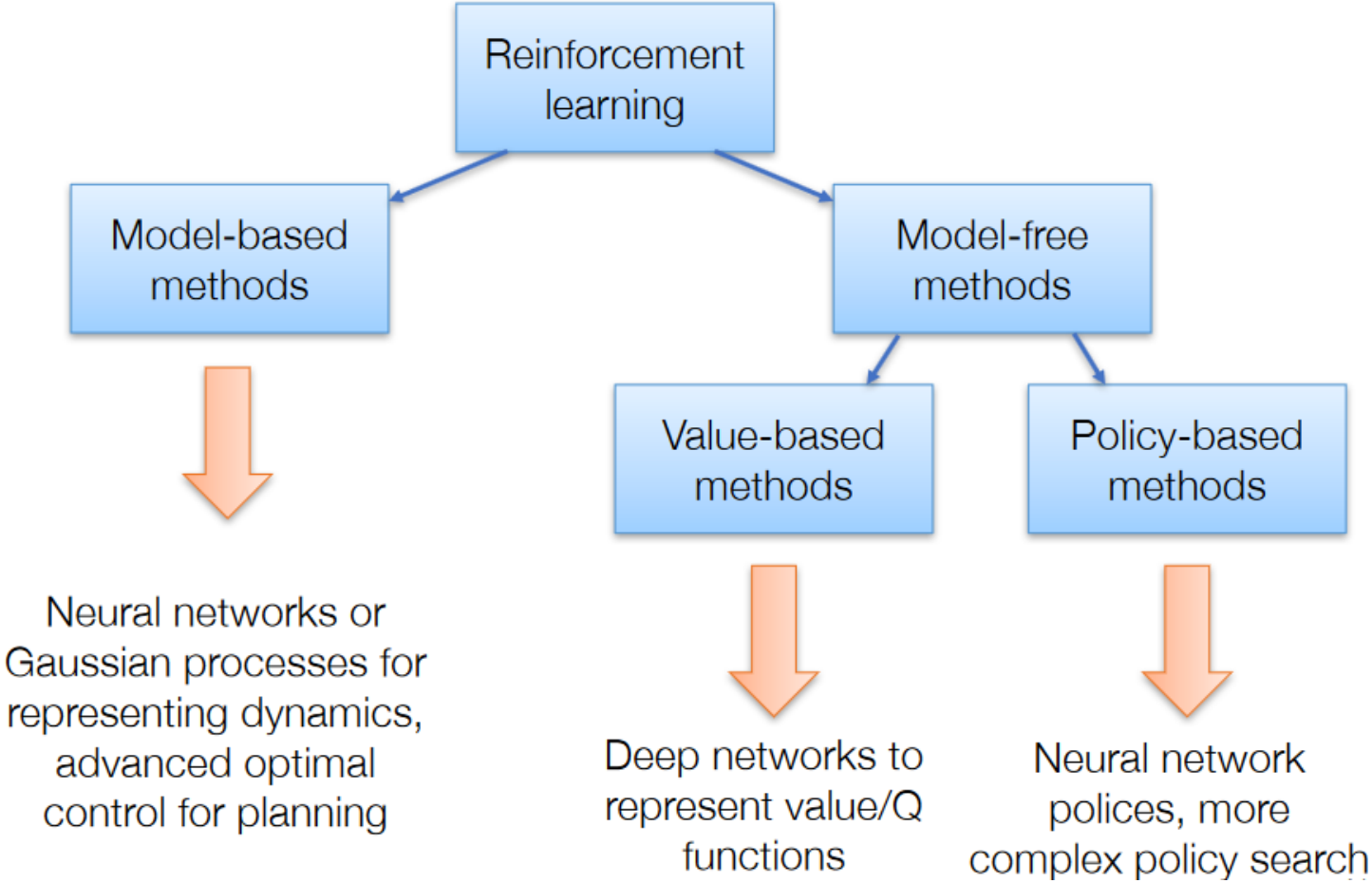


Optimal control vs Deep RL

	Optimal Control*	**RL**
Objective	Cost function	Reward
Control	Set of differential eqs.	Neural Network
Environment	Known model (often expressed as constraints)	Possible unknown model

RL is a form of stochastic optimal controls

*Largely founded by Lev Pontryagin and Richard Bellman


Reinforcement learning
Model-based methods
Model-free methods
Value-based methods
Policy-based methods
Neural networks or Gaussian processes for representing dynamics, advanced optimal control for planning
Deep networks to represent value/Q functions
Neural network polices, more complex policy search