Artificial Intelligence: Reinforcement Learning in Python

By: The LazyProgrammer

Reinforcement Learning

- When people talk about AI, they don't usually think about supervised/unsupervised learning
- These seem trivial compared to what we want an AI to do: play Chess/Go, drive cars, beat video games at superhuman level
- RL does all that and more
- Like Deep Learning, a lot of theory developed in 70s/80s, but not popular until now
- Now AIs can play games like Doom and Super Mario
- Law of accelerating returns: this improvement will grow exponentially!

Reinforcement Learning

- Supervised/Unsupervised Learning is not easy! I have 16+ courses on them already!
- Yet reinforcement learning opens up a whole new world
- RL paradigm is more different from Supervised/Unsupervised than they are from each other

What's in this course?

- In RL: huge gap between theory and code
- Theory is very abstract, important to continually think about how it applies to AI tasks like controlling a robot / playing a game
- So we start the course with some games:
- Slot machines
- Picks up where my course on Bayesian ML (A/B testing) left off
- Explore-exploit dilemma
- Epsilon-greedy, UCB1, Bayesian methods
- We can go beyond A/B testing, e-commerce stores, ads → how to behave in the real world

What's in this course?

- Tic-tac-toe
- How would you code an agent to play tic-tac-toe as a first-year comp-sci student?
- We'll introduce the RL way of doing things

What's in this course?

- Markov Decision Process (MDP)
- 3 techniques for solving MDPs:
  - Dynamic Programming (DP)
  - Monte Carlo (MC)
  - Temporal Difference Learning (TD)
- They all look very similar. The hard part is differentiating these 3!

What's in this course?

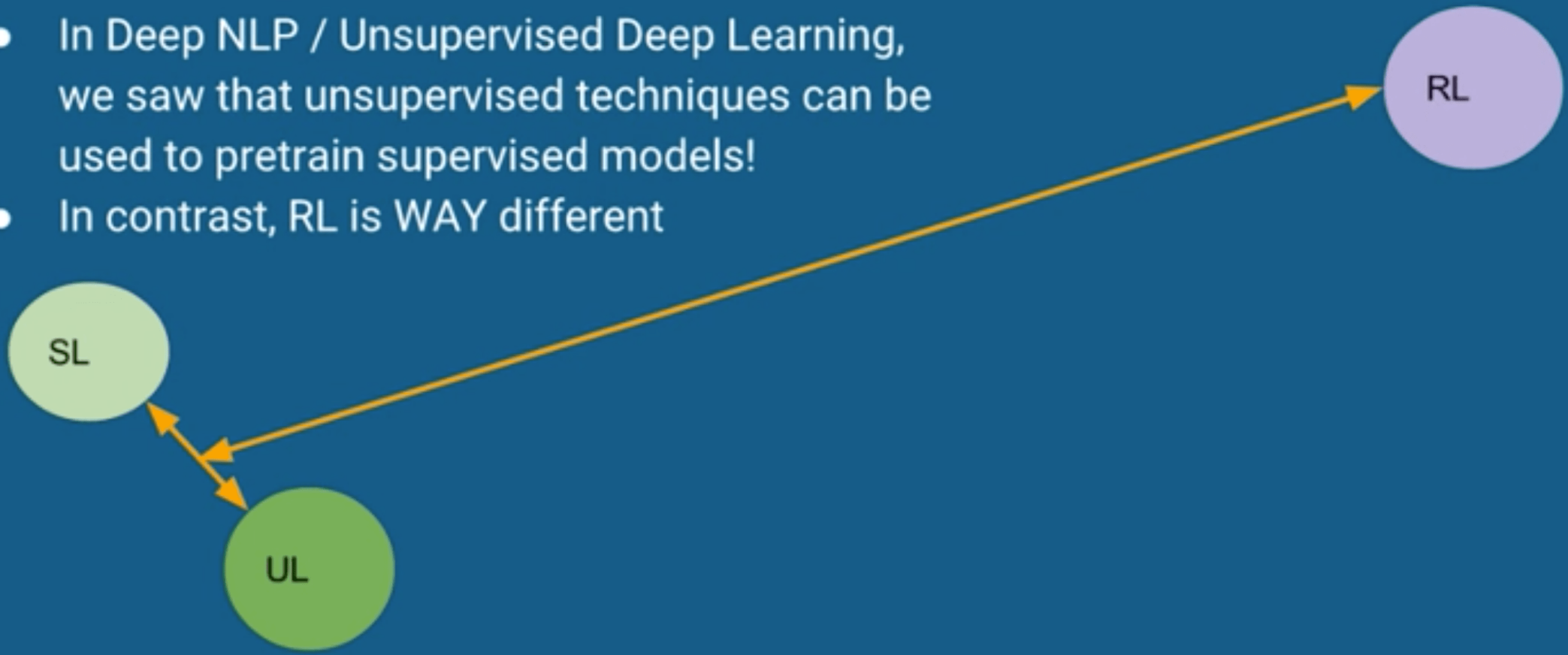
- Last section: Approximation methods
- Is a teaser into Deep Reinforcement Learning
- 3 solutions in this course won't scale to large problems, we need to modify them
- What's good for function approximation? Deep learning!
- This course will show you where you can "plug-in" a deep neural network, or any other approximator
- Without this course, Deep RL won't make much sense
- My original idea was to include these prerequisites into Deep RL, but there was just too much info!

What is reinforcement learning?

- What is reinforcement learning?
- How is it different from supervised and unsupervised learning?
- What are its applications?

What is reinforcement learning?

- In Deep NLP / Unsupervised Deep Learning, we saw that unsupervised techniques can be used to pretrain supervised models!
- In contrast, RL is WAY different



What is reinforcement learning?

- In supervised/unsupervised ML, we always imagine the same interface:

```
class SupervisedModel:
  def fit(X, Y): …
  def predict(X): …

class UnsupervisedModel:
  def fit(X): …
  def transform(X): … // PCA, Autoencoders, RBMs
                      // K-Means, GMM, …don't really transform data
```

What is reinforcement learning?

- Common theme with Unsupervised/Supervised Learning:
- The interface is the training data (matrices of numbers)
- In supervised learning we can make predictions on unseen/future data
- Useful by itself:
- Ex. app where you can take a photo at the grocery store, it classifies the item and shows its nutritional information

What is reinforcement learning?

- RL is different
- RL can guide an agent on how to act in the real world
- Interface is much more broad than training vectors, it's the entire environment
- Environment can be real world or simulated world (video game)
- Ex. real world: robot that vacuums your house, robot that learns to walk
- Military is interested in this tech: RL agents can replace soldiers. Not just walk, but fight, diffuse bombs, make important decisions

What is reinforcement learning?

- RL agents also train in a completely different way
- Many references to psychology
- Agents are used to model animal behavior
- Objective is a goal
- Supervised objective: maximize accuracy or likelihood / minimize cost
- RL agents get feedback as they interact with environment
- Feedback signals (rewards) are automatically given to the agent by the environment
- vs. Supervised learning: labels need to be made by humans (time-consuming and costly)

What is reinforcement learning?

- Phrasing objective in terms of goal allows us to solve a wide variety of problems
- AlphaGo's goal is to win Go
- Goal of video game AI: win the game / get highest score
- What about animals and humans?
- "Selfish Gene" theory (Richard Dawkins)
- AlphaGo found unique ways of winning: unusual and unexpected by observers
- Does that apply to animals/humans?

What is reinforcement learning?

- Why do we do anything we do?
- Selfish Gene theory says everything we do is designed to serve our genes' desire to multiply
- Ex. Why do people want to be rich?
- Perhaps led to better healthcare or social status, led to genes maximizing their goal
- Richness has no physical relationship to genes, yet it's a novel solution to the problem

What is reinforcement learning?

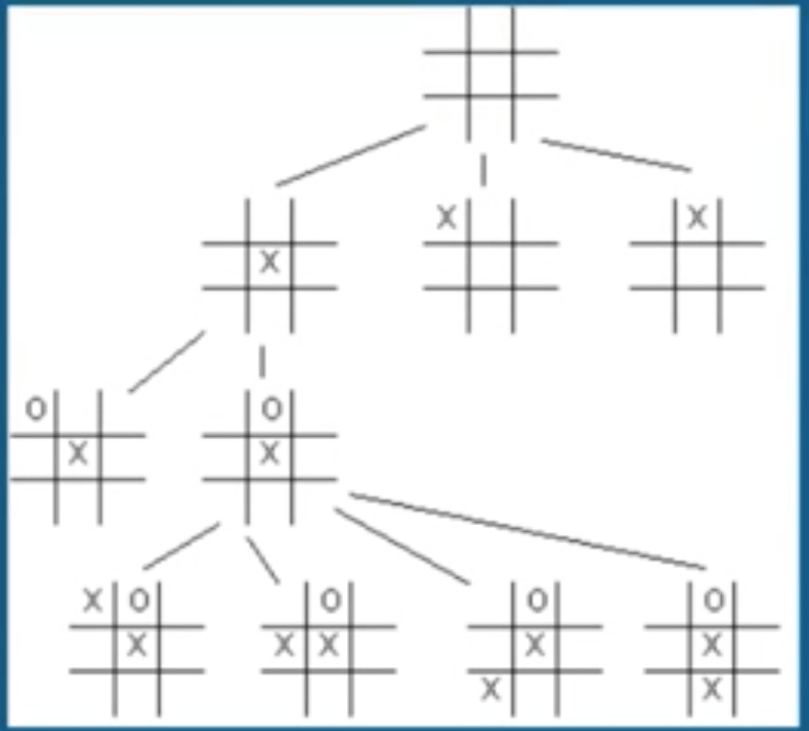
- “Desiring money” just a random example
- Replace with anything you like: “being healthy and strong”, “having good analytic skills”
- Studying those factors is a social scientist’s job
- For us, it’s more interesting that there is just one main objective to maximize but various novel ways to achieve it
- At one point in history, seeking as much sugar as possible would give you energy (hence improving survival). Today, it kills you!
- Genes use mutation and natural selection
- AI uses reinforcement learning

What is reinforcement learning?

- Humans / AIs alike never sense the entire world/universe at once
- We have sensors which feed signals to our brain from the environment
- We don’t even know everything that’s going on in a room
- Thus the sensors limit the amount of information we get
- The measurements we get from these sensors (e.g. sight, sound, touch) make up a “state”
- We’ll only discuss finite state spaces
- State spaces with an infinite number of states are possible too

What is reinforcement learning?

- If you don't know how tic-tac-toe works, do a google search, you can play right in the browser!
- What's the # of states?
  [if we simplify the problem so that
  we can keep adding x's and o's
  even after a player gets 3 in a row]



What is reinforcement learning?

- Each location on the board has 3 possibilities: empty, X, O
- 9 locations on the board
- # states = 3x3x...x3 = 3^9

Recap so far

- 3 important terms (I may have snuck in a few more)

Agent: thing that senses the environment, thing we're trying to code intelligence/learning into.

Environment: Real world or simulated world that the agent lives in.

State: Different configurations of the environment that the agent can sense.