RLearning:
Short guides to reinforcement learning

Introduction

Davud Rostam-Afschar (Uni Mannheim)

What is Reinforcement Learning?

Machine Learning

- Traditional computer science
  - Program computer for every task

Machine Learning

- Traditional computer science
  - Program computer for every task

Machine Learning

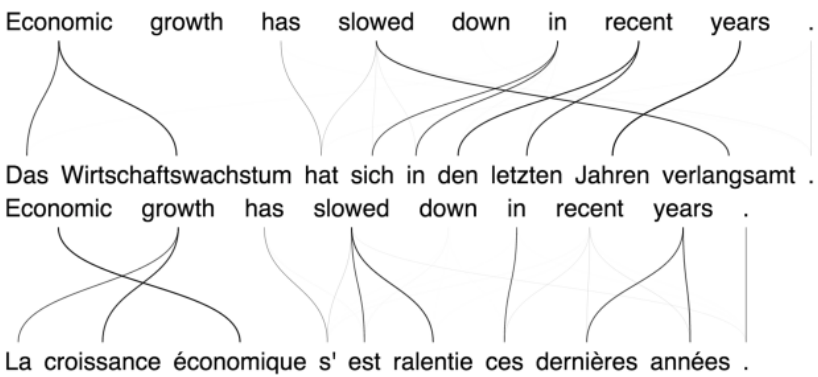
- Traditional computer science
  - Program computer for every task

- New paradigm
  - Provide examples to machine
  - Machine learns to accomplish a task based on the examples

Machine Learning

- Traditional computer science
  - Program computer for every task

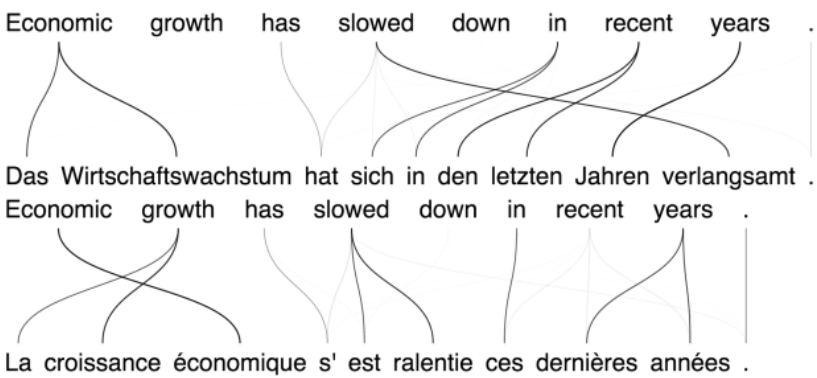
- New paradigm
  - Provide examples to machine
  - Machine learns to accomplish a task based on the examples

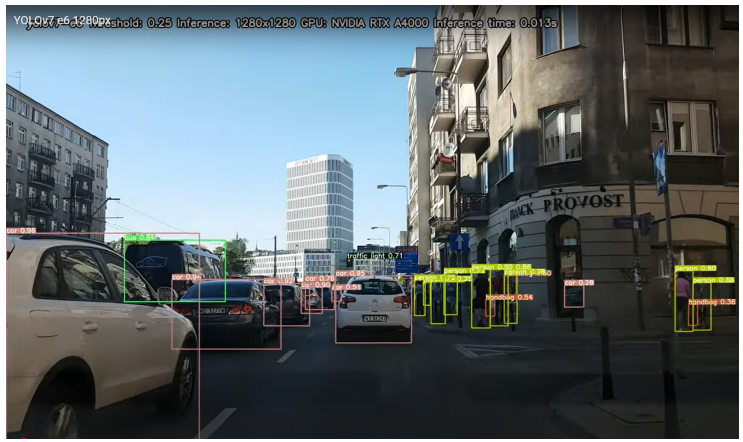


Machine Learning

- Traditional computer science
  - Program computer for every task

- New paradigm
  - Provide examples to machine
  - Machine learns to accomplish a task based on the examples





Machine Learning: Different Approaches

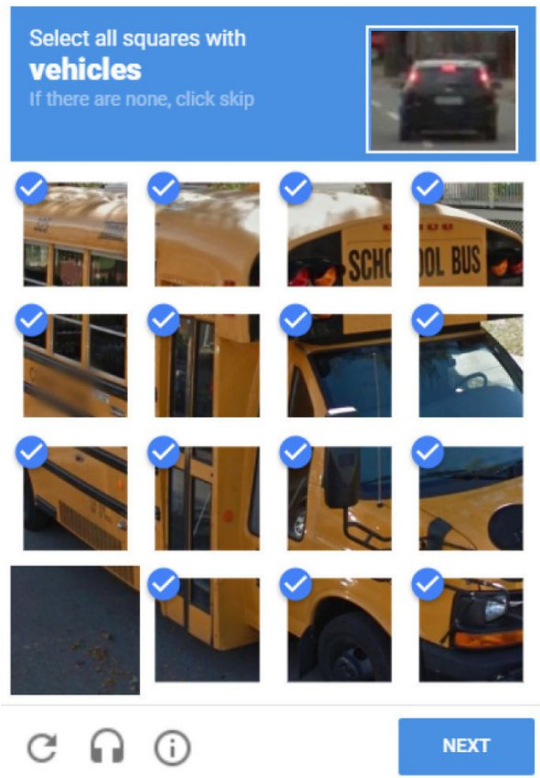
- **Supervised Learning**
  - Most common type of machine learning
  - Learns from labeled examples
    (input + correct output)

Machine Learning: Different Approaches

- **Supervised Learning**
  - Most common type of machine learning
  - Learns from labeled examples
    (input + correct output)
  - **Challenge:** Needs lots of labeled data

Machine Learning: Different Approaches

- **Supervised Learning**
  - Most common type of machine learning
  - Learns from labeled examples (input + correct output)
  - **Challenge:** Needs lots of labeled data



Machine Learning: Different Approaches

- **Supervised Learning**
  - Most common type of machine learning
  - Learns from labeled examples
    (input + correct output)
  - **Challenge:** Needs lots of labeled data

- **Other Approaches**
  - **Unsupervised Learning:**
    Finds patterns without labeled data
  - **Semi-supervised Learning:**
    Mixes labeled and unlabeled data
  - **Reinforcement Learning:**
    Learns by trial and error

Animal Psychology

- Positive reinforcements:
  - Pleasure and food
- Negative reinforcements:
  - Pain and hunger
- Reinforcements used to train animals

Animal Psychology

- Positive reinforcements:
  - Pleasure and food
- Negative reinforcements:
  - Pain and hunger
- Reinforcements used to train animals

- **Let's do the same with computers!**

What is Reinforcement Learning?

- Reinforcement learning is also known as
  - Optimal control
  - Approximate dynamic programming
  - Neuro-dynamic programming

Definition

Reinforcement learning is an area of machine learning inspired by behavioral psychology, concerned with how software **agents** ought to take **actions** in an **environment** so as to maximize some notion of cumulative **reward**.

- **?**, chapter 1
- **?**, chapter 1

What is Reinforcement Learning?

- Reinforcement learning is also known as
  - Optimal control
  - Approximate dynamic programming
  - Neuro-dynamic programming

Definition

Reinforcement learning is an area of machine learning inspired by behavioral psychology, concerned with how software **agents** ought to take **actions** in an **environment** so as to maximize some notion of cumulative **reward**.

- **?**, chapter 1
- **?**, chapter 1

What is Reinforcement Learning?

- Reinforcement learning is also known as
  - Optimal control
  - Approximate dynamic programming
  - Neuro-dynamic programming

Definition

Reinforcement learning is an area of machine learning inspired by behavioral psychology, concerned with how software **agents** ought to take **actions** in an **environment** so as to maximize some notion of cumulative **reward**.

- **?**, chapter 1
- **?**, chapter 1

What is Reinforcement Learning?

- Reinforcement learning is also known as
  - Optimal control
  - Approximate dynamic programming
  - Neuro-dynamic programming

Definition

Reinforcement learning is an area of machine learning inspired by behavioral psychology, concerned with how software **agents** ought to take **actions** in an **environment** so as to maximize some notion of cumulative **reward**.

- **?**, chapter 1
- **?**, chapter 1

What is Reinforcement Learning?

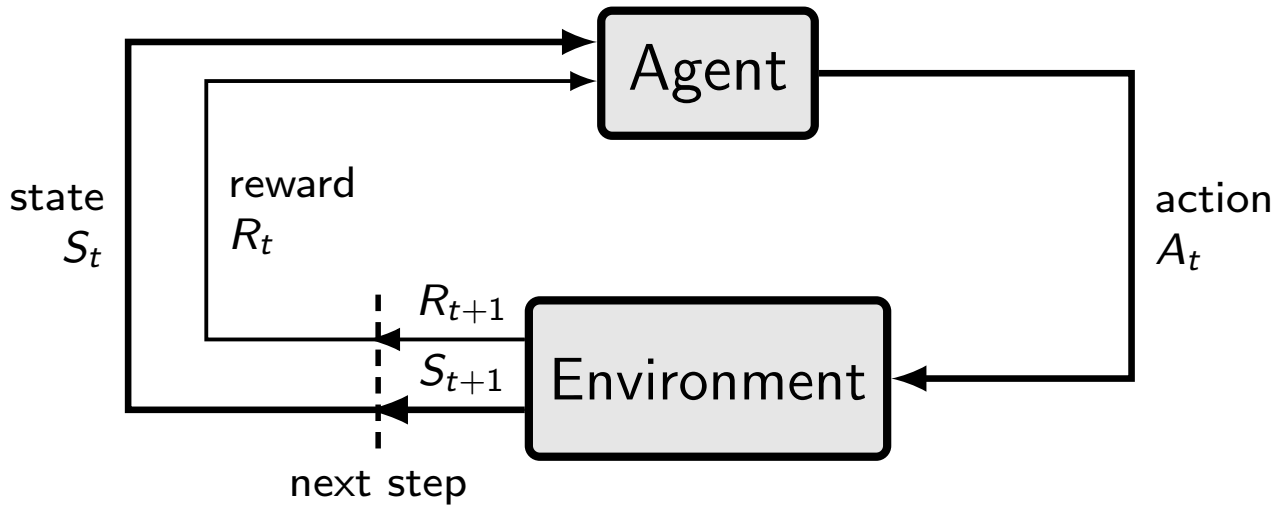
- Reinforcement learning is also known as
  - Optimal control
  - Approximate dynamic programming
  - Neuro-dynamic programming

Definition

Reinforcement learning is an area of machine learning inspired by behavioral psychology, concerned with how software **agents** ought to take **actions** in an **environment** so as to maximize some notion of cumulative **reward**.

- **?**, chapter 1
- **?**, chapter 1

Reinforcement Learning Problem



Goal: Learn to choose actions that maximize rewards

Applications and Examples

RL Examples

- Game playing (go, atari, backgammon)
- Elevator scheduling
- Helicopter control
- Spoken dialog systems
- Data center energy optimization
- Self-managing network systems
- Autonomous vehicles

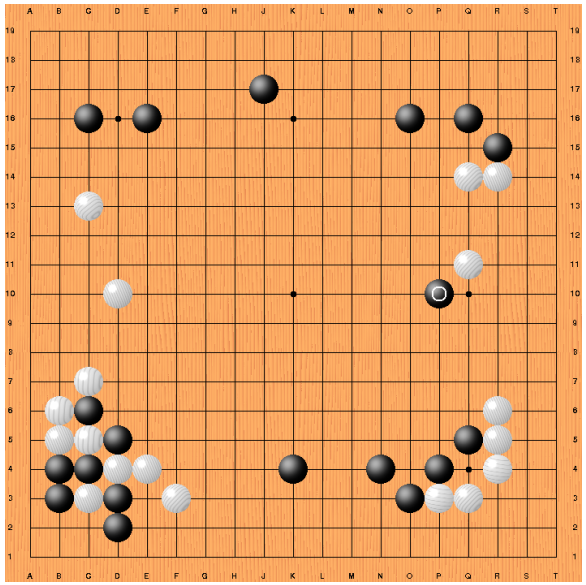
RL Examples in the Social Sciences

- Operations research
  (pricing, vehicle routing)
- Computational finance
  (portfolio optimization, algorithmic trading)

- Example: vehicle routing
- **Agent:** vehicle routing software
- **Environment:** stochastic demand
- **State:** vehicle location, capacity and depot requests
- **Action:** vehicle route
- **Reward:** - travel costs

Game Playing

- Example: Go (one of the oldest and hardest board games)
- **Agent:** player
- **Environment:** opponent
- **State:** board configuration
- **Action:** next stone location
- **Reward:** +1 win / -1 loose



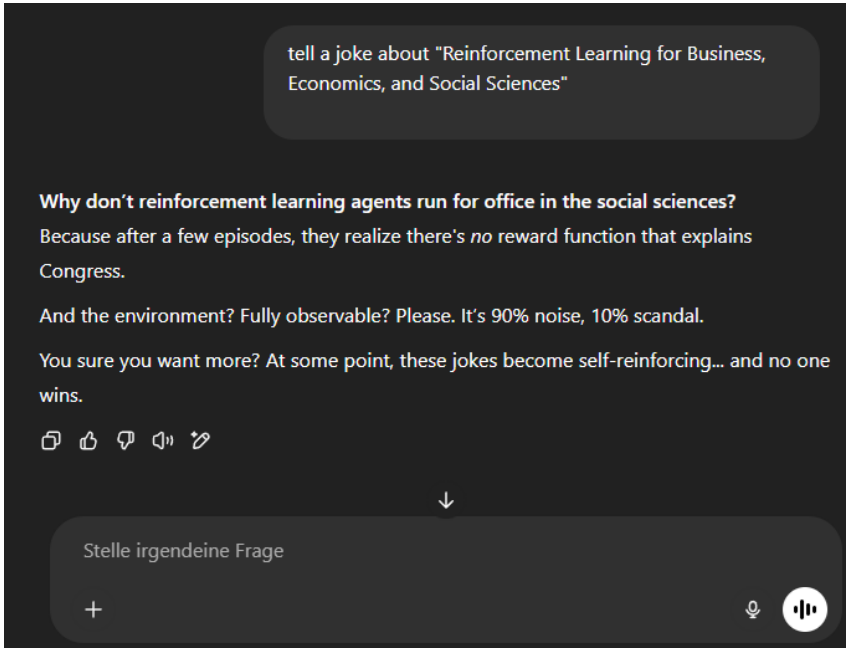
  - 2016: AlphaGo defeats top player Lee Sedol (4-1)
    - Game 2 move 37: AlphaGo plays unexpected move (odds 1/10,000)
      `https://www.youtube.com/watch?v=WXuK6gekU1Y`

Robotic Control

- Example: helicopter control
- **Agent:** controller
- **Environment:** helicopter
- **State:** position, orientation, velocity and angular velocity
- **Action:** collective pitch, cyclic pitch, tail rotor control
- **Reward:** - deviation from desired trajectory
  - 2008 (Andrew Ng): automated helicopter wins acrobatic competition against humans
    `https://www.youtube.com/watch?v=0JL04JJjocc`

Conversational Agent

- Example: Conversational Agent (ChatGPT)
- **Agent:** language model
- **Environment:** user
- **State:** conversation history
- **Action:** next token
- **Reward:** ratings based on task completion, user satisfaction, etc.
- **Today:** active area of research



Computational Finance

- Example: Automated trading
- **Agent:** trading software
- **Environment:** other traders
- **State:** price history
- **Action:** buy/sell/hold
- **Reward:** amount of profit

Example: trading strategies that adapt to real-time market signals

Adaptive lab, field, or survey experiments

- Advertising

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting
- Crime prevention and policing

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting
- Crime prevention and policing
- Political campaign strategy

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting
- Crime prevention and policing
- Political campaign strategy

Explaining Behavior and Assisting Decision Making

- Strategic decision making (game theory)

RL Examples in the Social Sciences

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting
- Crime prevention and policing
- Political campaign strategy

Explaining Behavior and Assisting Decision Making

- Strategic decision making (game theory)
- Households choices (fertility, labor, education, consumption/saving)

Adaptive lab, field, or survey experiments

- Advertising
- Educational learning
- Labor market platforms
- Public health interventions
- Social program targeting
- Crime prevention and policing
- Political campaign strategy

Explaining Behavior and Assisting Decision Making

- Strategic decision making (game theory)
- Households choices (fertility, labor, education, consumption/saving)
- Firms choices (entry, exit, investments, hiring, pricing, output)

Course Overview

Course overview

1. Unit 1: Multi-Armed Bandits
2. Unit 2: Markov Decision Processes
   Assignment 1
3. Unit 3: RL Algorithms
   Assignment 2
4. Unit 4: Deep RL

References I

Charpentier, A., R. Élie, and C. Remlinger (2023): "Reinforcement learning in economics and finance," *Computational Economics*, 62(1), 425–462.

Goodfellow, I. (2016): *Deep learning*, vol. 196. MIT press, Available at `http://deeplearningbook.org/`.

Puterman, M. L. (2014): *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Russell, S. J., and P. Norvig (2016): *Artificial intelligence: a modern approach*. Pearson.

Schulz, L., and R. Bhui (2024): "Political reinforcement learners," *Trends in Cognitive Sciences*, 28(3), 210–222.

Sutton, R. S., and A. G. Barto (2018): "Reinforcement learning: An introduction," *A Bradford Book*, Available at `http://incompleteideas.net/book/the-book-2nd.html`.

Szepesvári, C. (2022): *Algorithms for reinforcement learning*. Springer nature, Available at `https://sites.ualberta.ca/~szepesva/RLBook.html`.

Takeaways

What is Reinforcement Learning?

- Comprehensive, but challenging form of machine learning
  - Interdependent sequence of decisions
  - Incomplete model
  - Stochastic environment
  - No supervision
  - Partial and delayed feedback
- **Long term goal:** autonomous agents—without needing explicit supervision (according to ChatGPT)