Reinforcement Learning: Overview

Prabuchandran K.J.

Assistant Professor, IIT Dharwad

5 March 2020

Outline

- Faces of Reinforcement Learning (RL)

Outline

- Faces of Reinforcement Learning (RL)
  - Computer Science $\rightarrow$ Machine Learning $\rightarrow$ Reinforcement Learning

Outline

- Faces of Reinforcement Learning (RL)
  - Computer Science $\rightarrow$ Machine Learning $\rightarrow$ Reinforcement Learning
- Sequential Decision Making (SDM) problems

Outline

- Faces of Reinforcement Learning (RL)
  - Computer Science $\rightarrow$ Machine Learning $\rightarrow$ Reinforcement Learning
- Sequential Decision Making (SDM) problems
  - How to solve SDM Problems?

Outline

- Faces of Reinforcement Learning (RL)
  - Computer Science $\rightarrow$ Machine Learning $\rightarrow$ Reinforcement Learning
- Sequential Decision Making (SDM) problems
  - How to solve SDM Problems?

Reinforcement Learning Methods

  - Examples and successful RL solutions

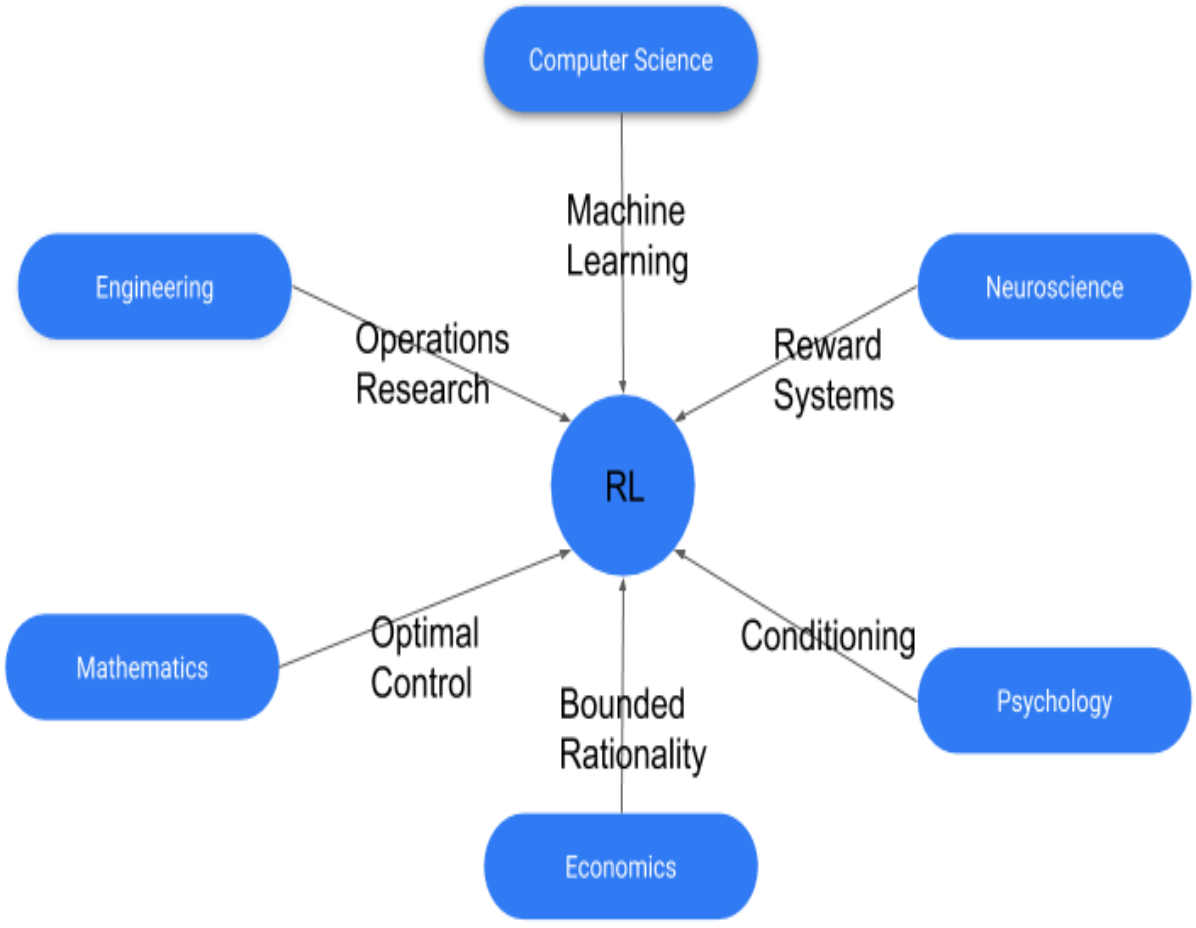
Outline

- Faces of Reinforcement Learning (RL)
  - Computer Science $\rightarrow$ Machine Learning $\rightarrow$ Reinforcement Learning
- Sequential Decision Making (SDM) problems
  - How to solve SDM Problems?

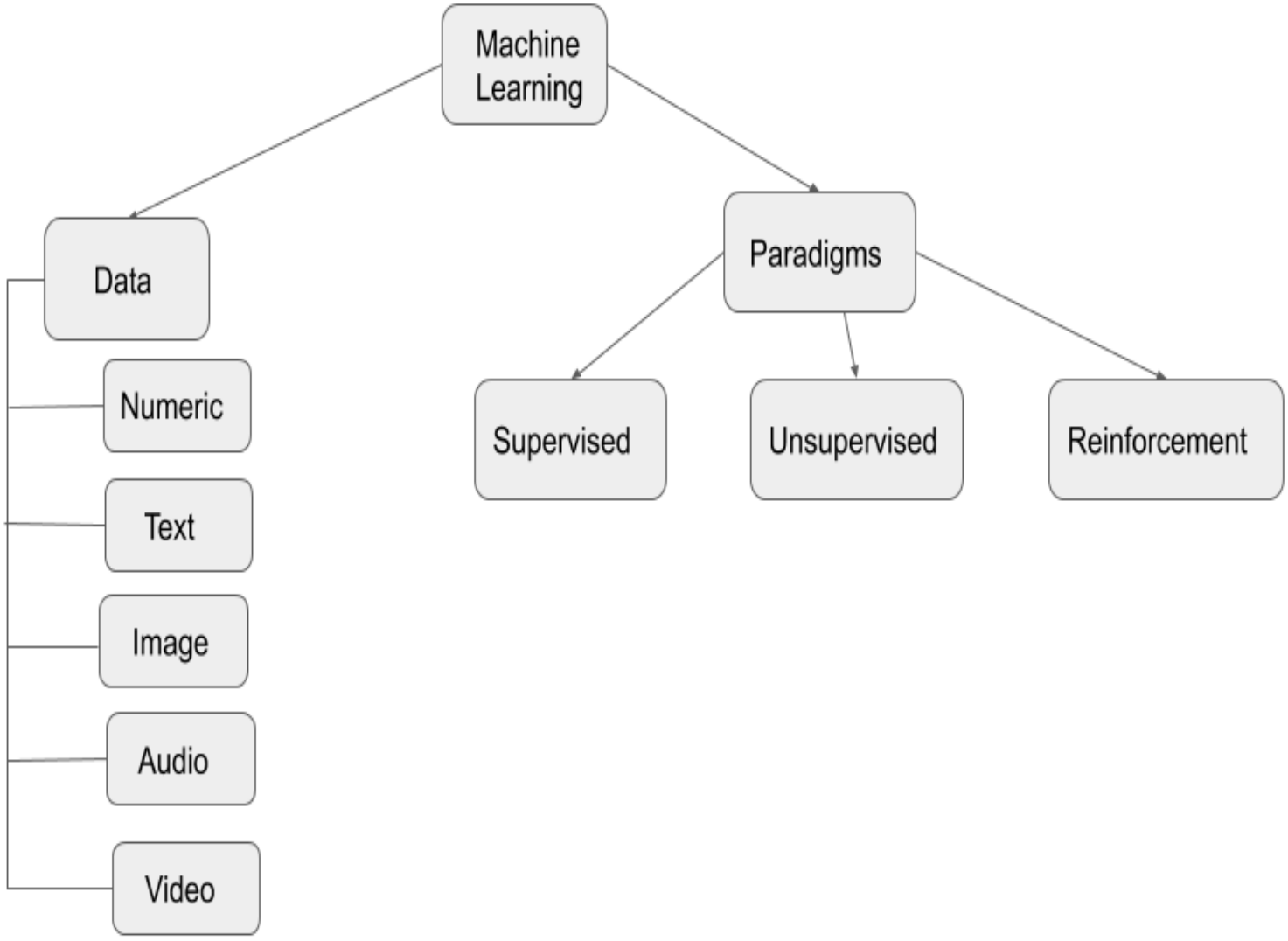
Reinforcement Learning Methods

  - Examples and successful RL solutions
- Mathematical frameworks for studying SDM Problems

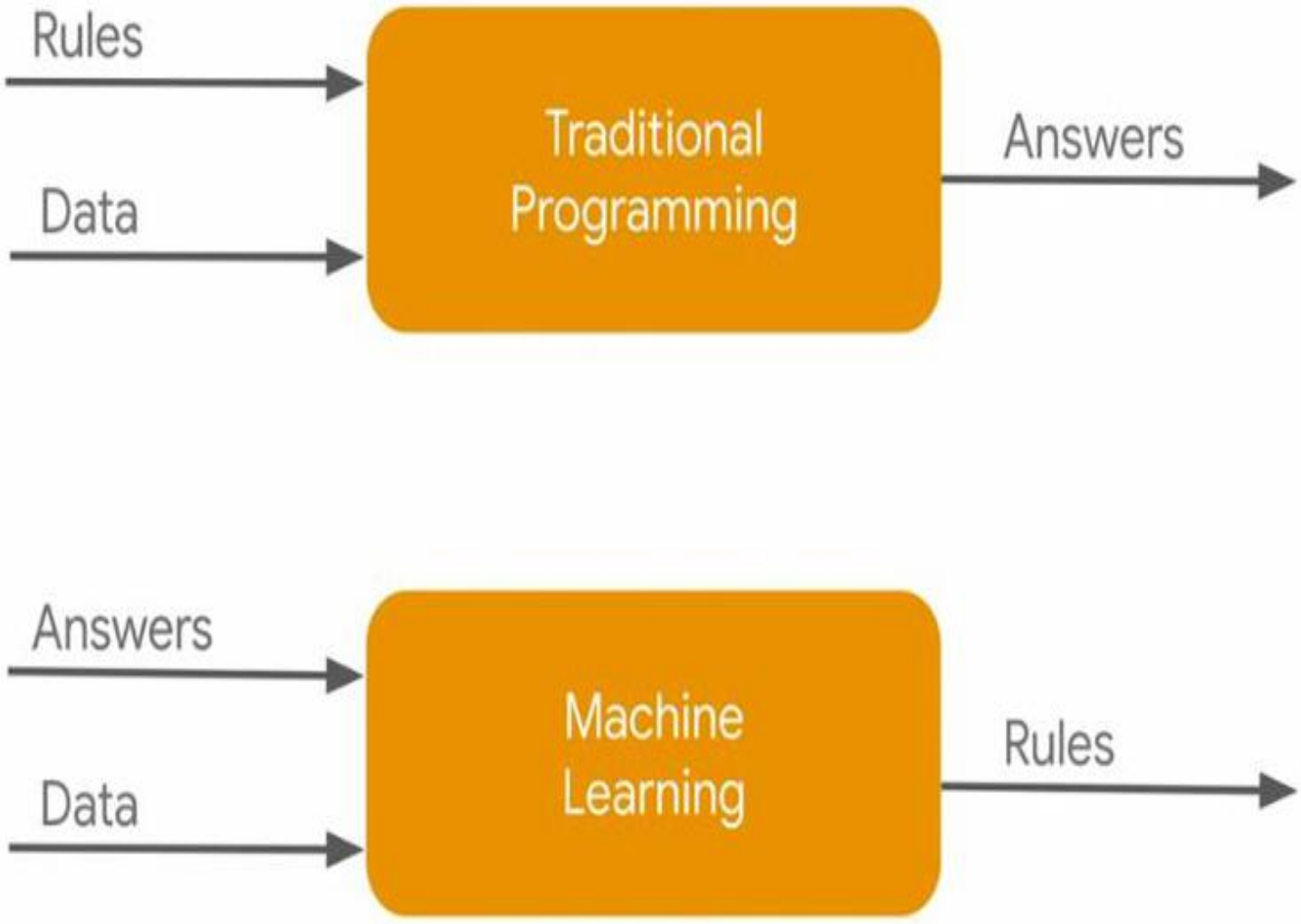
Faces of Reinforcement Learning



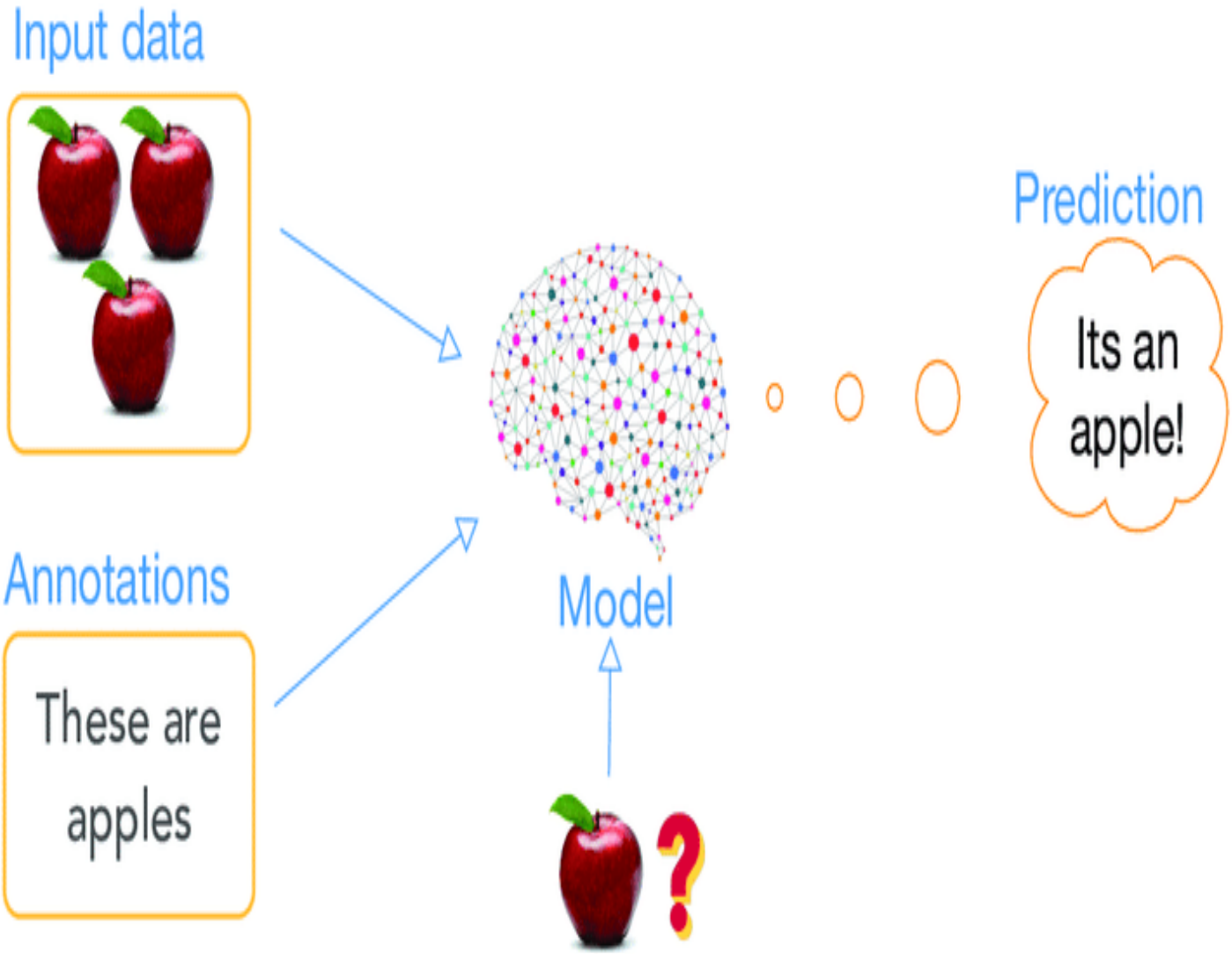
Machine Learning



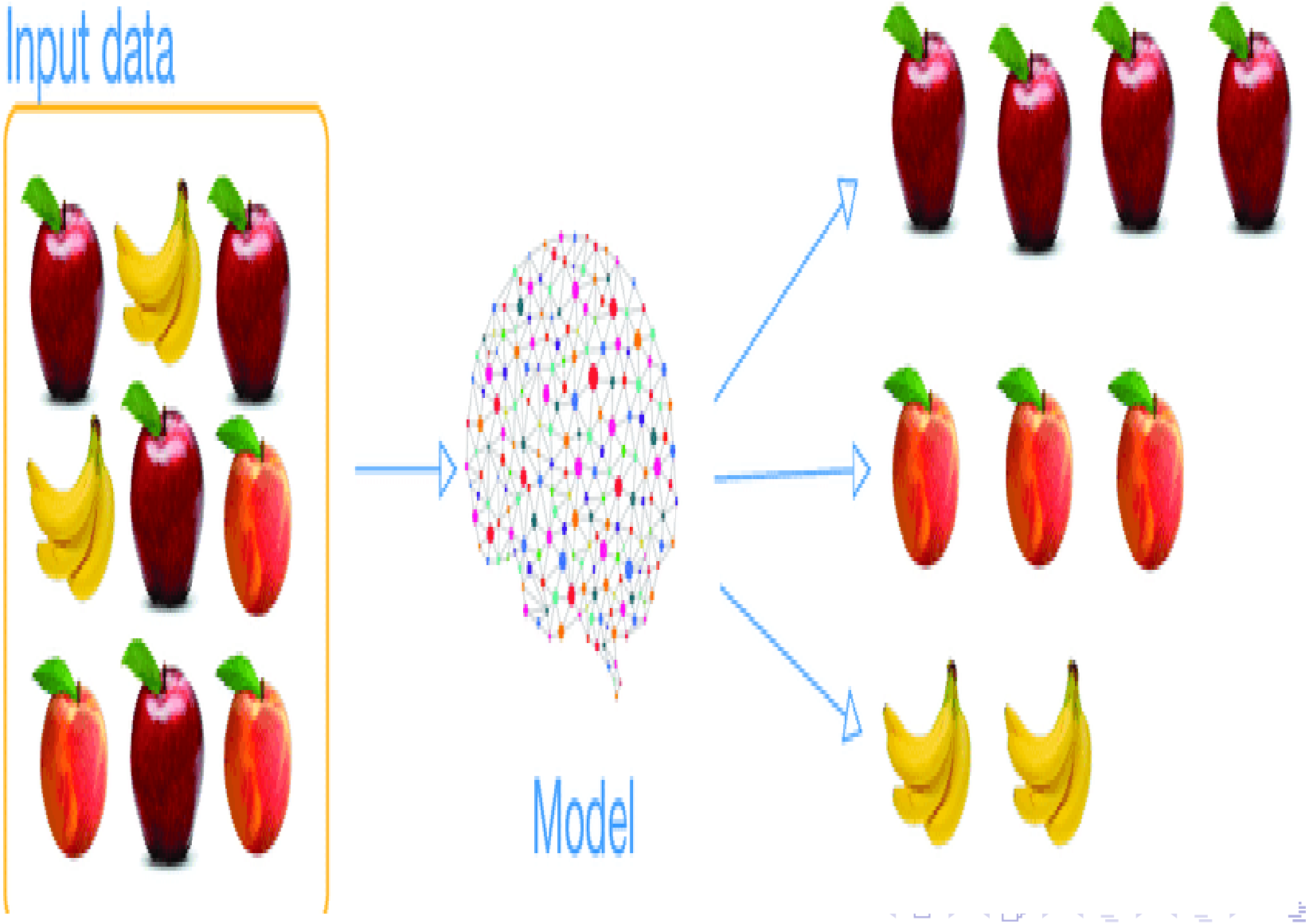
Traditional Programming vs Machine Learning



Supervised Learning

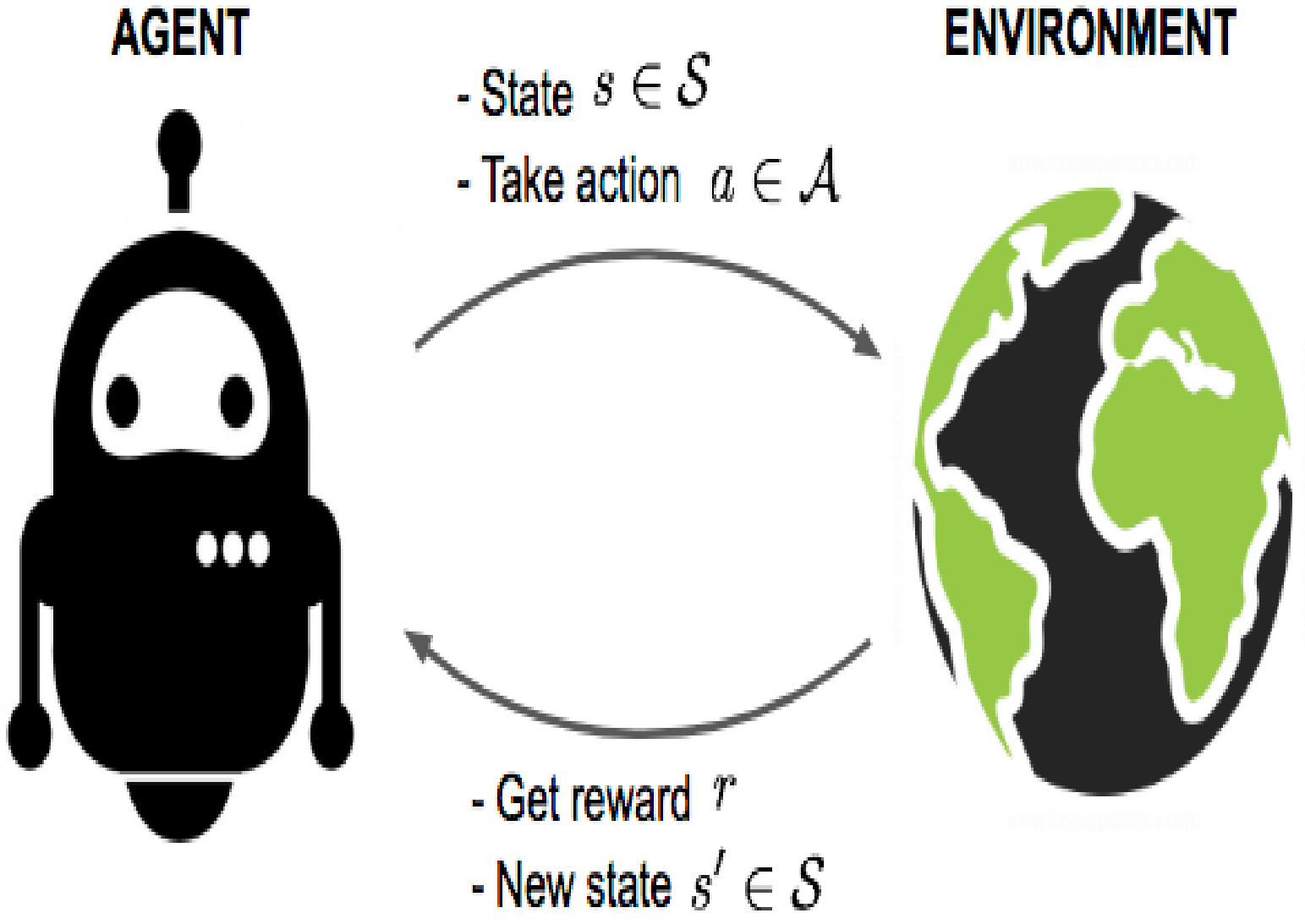


Unsupervised Learning



Reinforcement Learning

Reinforcement Learning Problem



Supervised vs Reinforcement Learning

Supervised Learning

- Examples with target
- (images,labels)
- Instructive Feedback

Reinforcement Learning

- Examples without target
- (states,actions,rewards)
- Evaluative Feedback

Characteristics of Reinforcement Learning

- Decision Making in the face of uncertainty

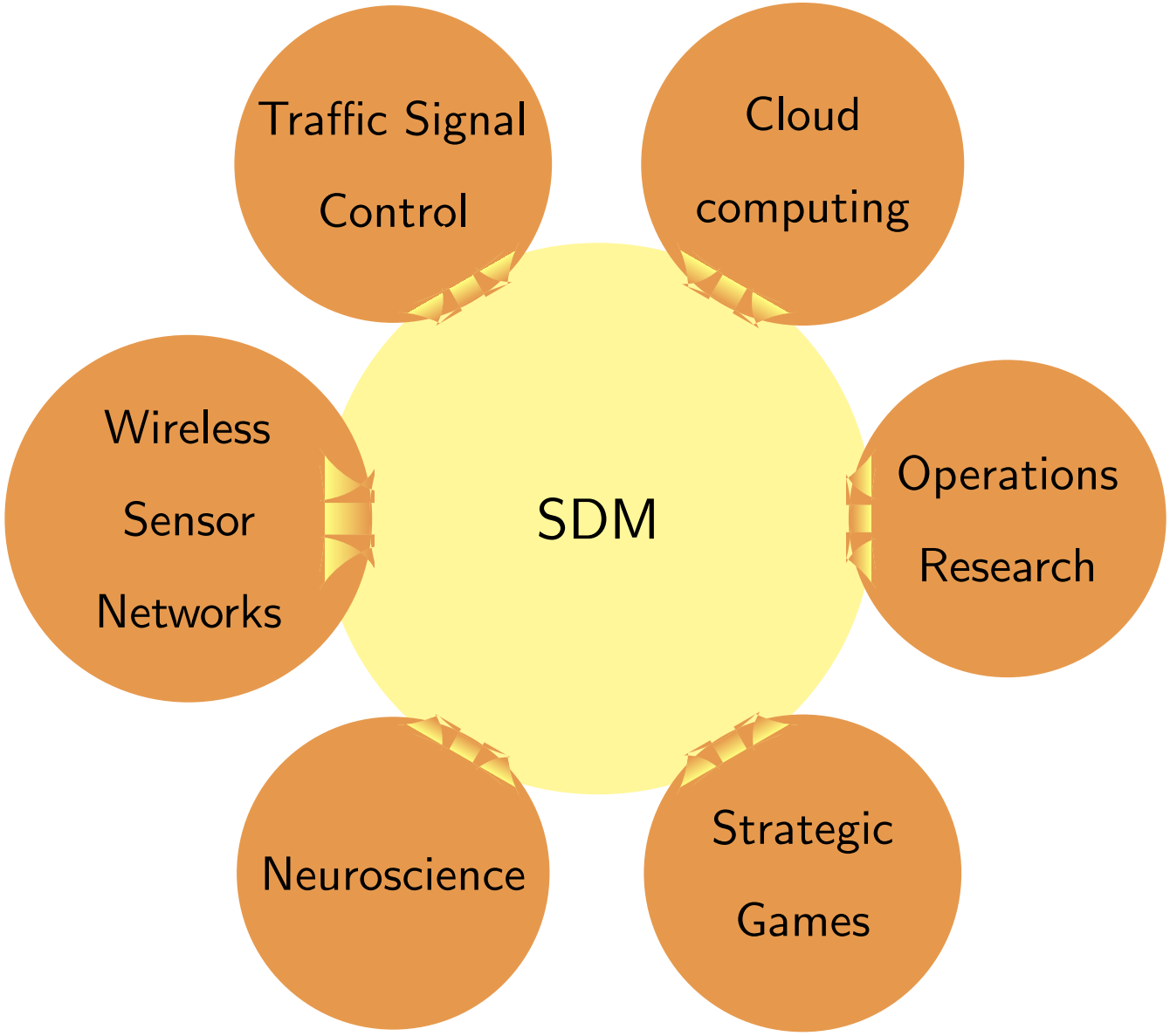
Characteristics of Reinforcement Learning

- Decision Making in the face of uncertainty
- Delayed rewards

Characteristics of Reinforcement Learning

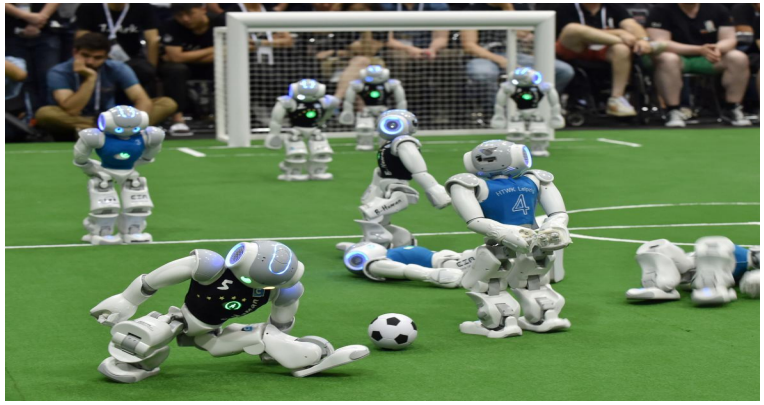
- Decision Making in the face of uncertainty
- Delayed rewards
- Credit assignment Problem

Sequential Decision Making (SDM) Problems

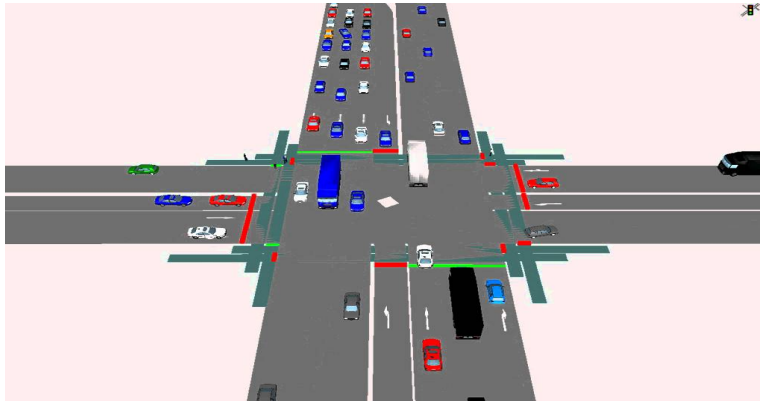


Examples of SDM problems

Strategic Games



Robo Soccer



Traffic Signal Control

Inventory Management

Common Features of SDM Problems

- Long-term goal that needs to be achieved

Common Features of SDM Problems

- Long-term goal that needs to be achieved
- Uncertainity in the evolution of configuration (state)

Common Features of SDM Problems

- Long-term goal that needs to be achieved
- Uncertainity in the evolution of configuration (state)
- Decisions (actions) need to be taken in stages

Common Features of SDM Problems

- Long-term goal that needs to be achieved
- Uncertainity in the evolution of configuration (state)
- Decisions (actions) need to be taken in stages
- Simple feedback signal (reward/cost) - how good is the action for the given state

Common Features of SDM Problems

- Long-term goal that needs to be achieved
- Uncertainity in the evolution of configuration (state)
- Decisions (actions) need to be taken in stages
- Simple feedback signal (reward/cost) - how good is the action for the given state
- Available information or experience - state, action and reward

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

Solving SDM Problems

How do we model?

Bandits/ Markov Decision Process

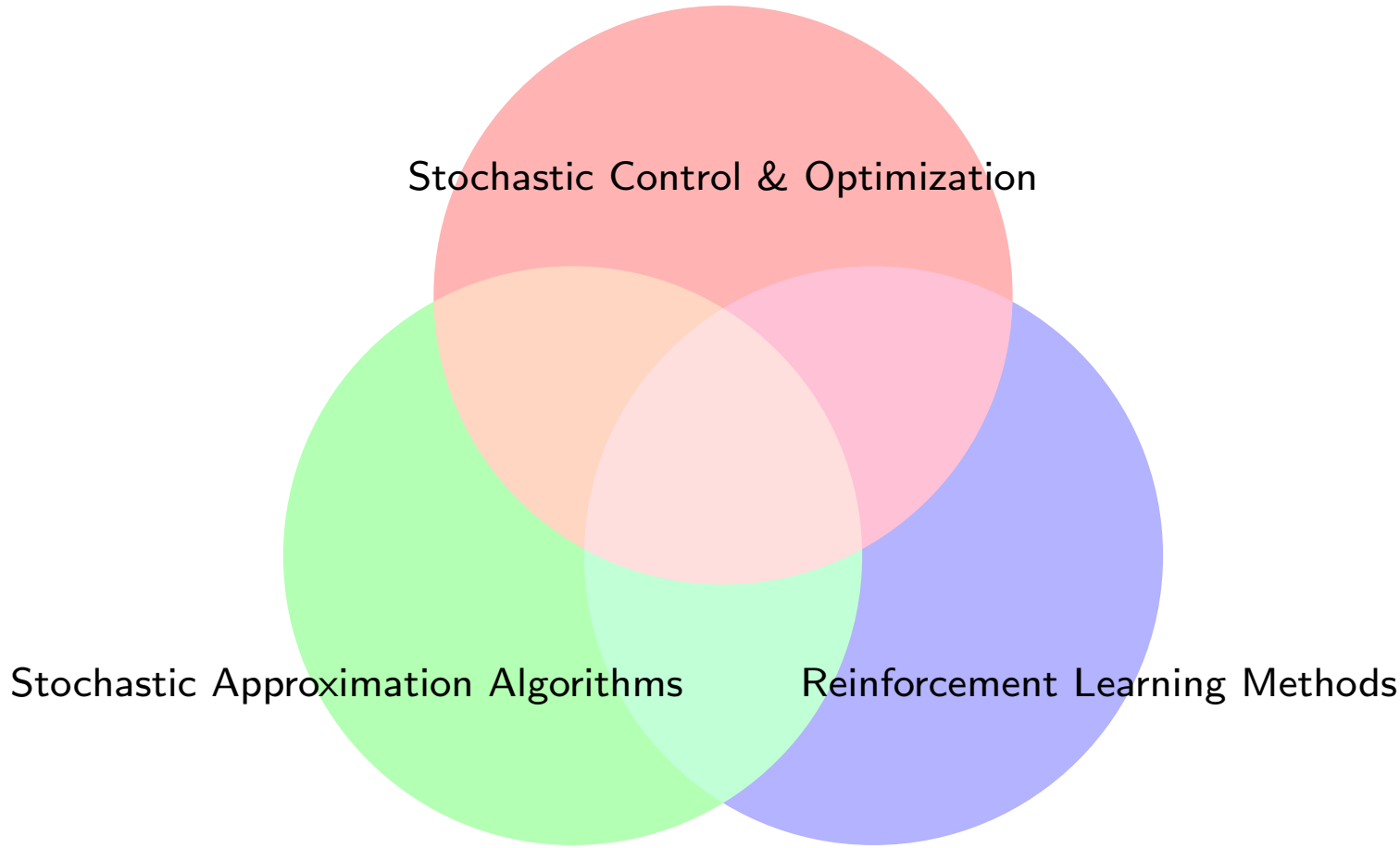
How to learn from experience?

Reinforcement Learning and Stochastic Optimization algorithms

How to analyse these algorithms?

Stochastic Approximation framework

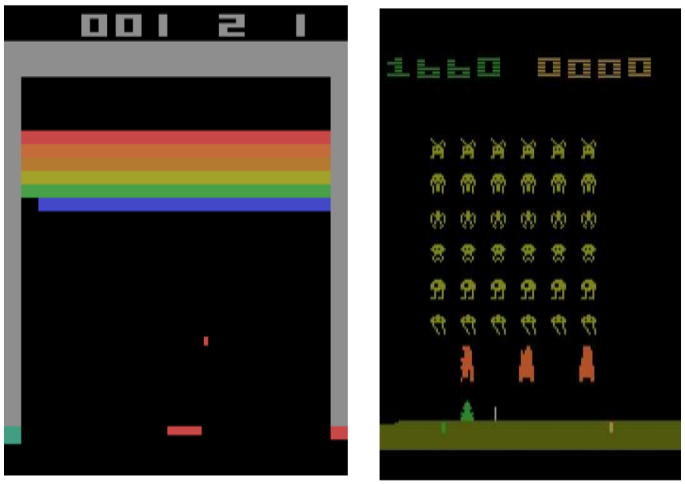
Solution to SDM problems



Successful Reinforcement Learning (RL) solutions

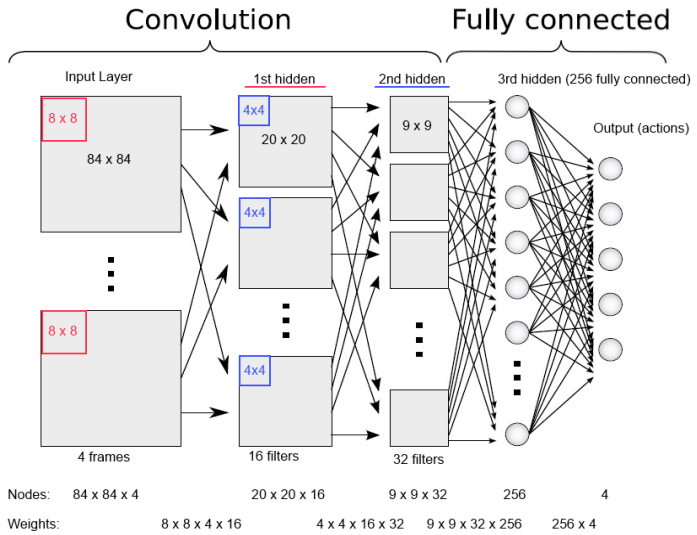
- Q-learning / Deep Q-Networks (DQN)
- Actor-Critic methods (AC) / Policy gradient methods
- Upper Confidence Tree (UCT) / Monte-Carlo Tree Search algorithm

Deep Q-Networks



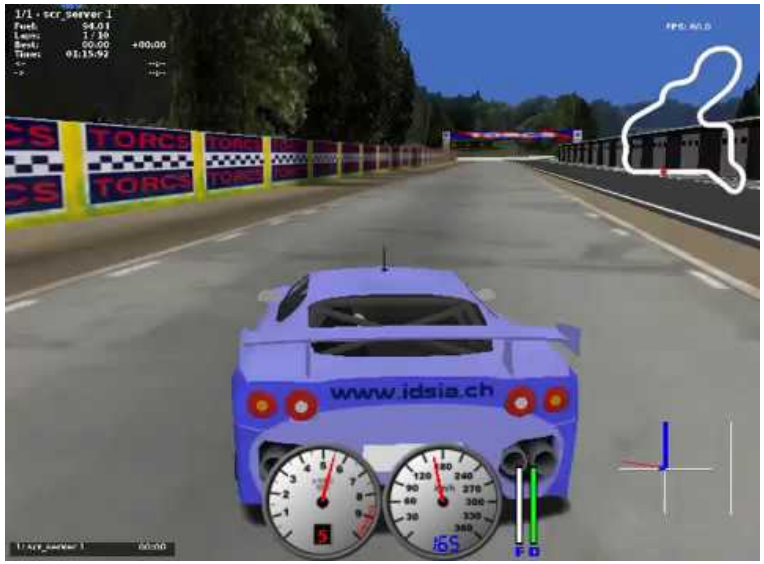
Breakout and Space Invaders, 2 of the 49 Atari games used in the paper

Break out and Space Invaders

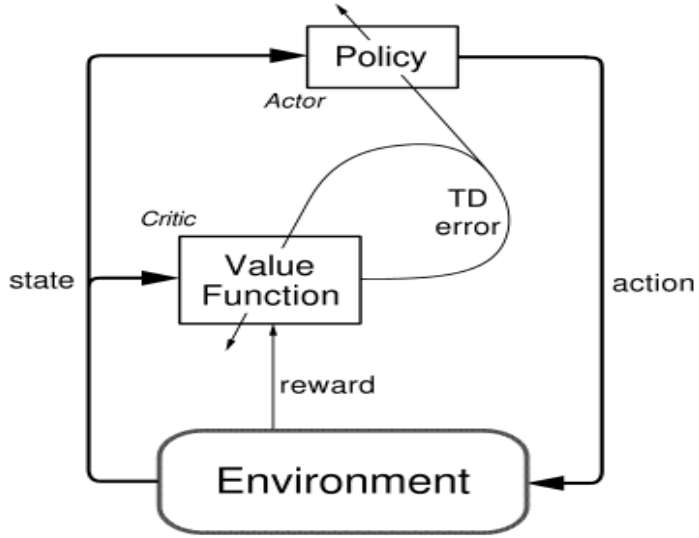


Deep Q-Network

Deep Deterministic Policy Gradient

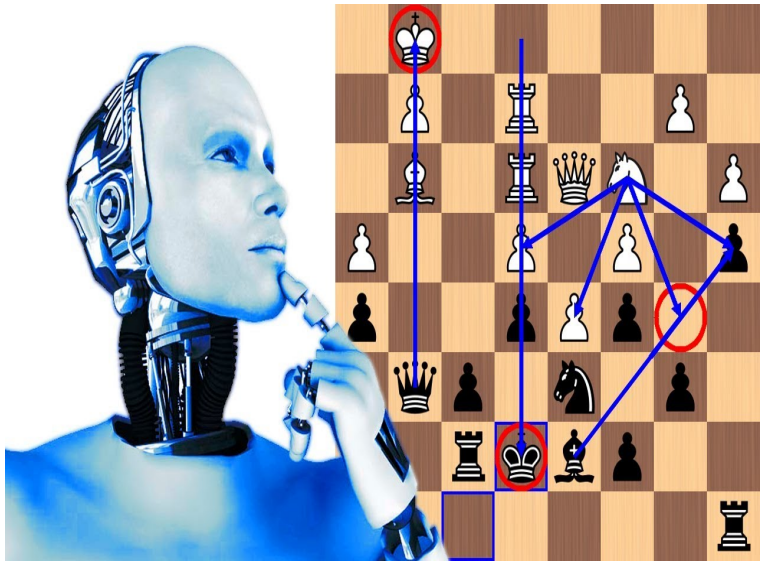


TORCS car simulation

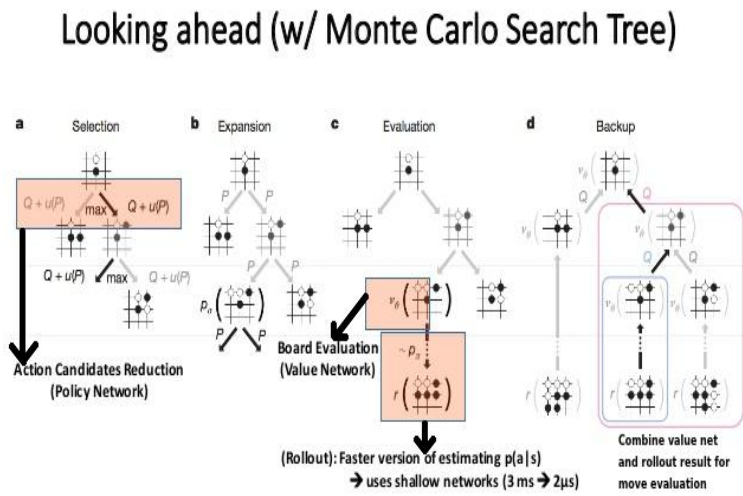


Actor-Critic Network

Upper Confidence Tree (UCT)



Alpha Zero



UCT algorithm

Reason for Success

Rule Based Methods

Feature Based Methods

Automatic Feature Based Methods

Reason for Success

Rule Based Methods

Feature Based Methods

Automatic Feature Based Methods

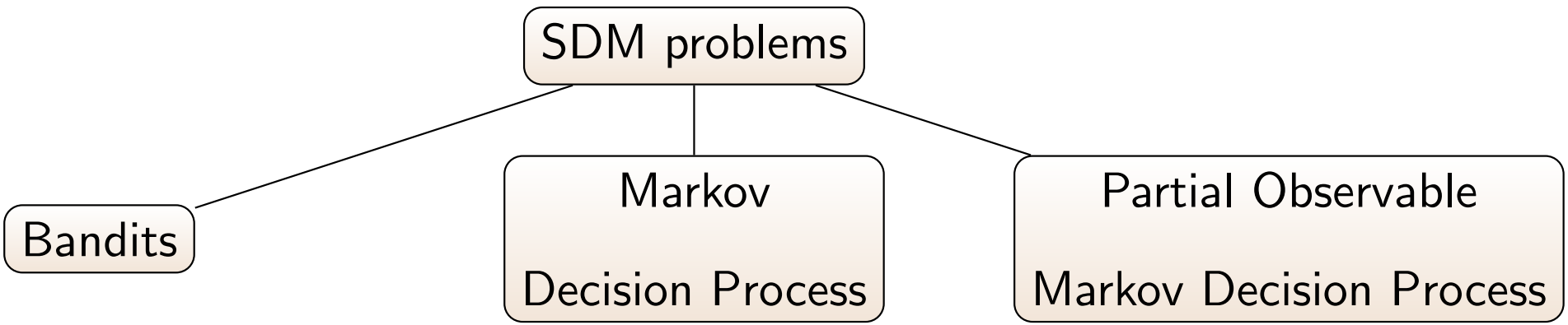
Reason for Success

Rule Based Methods

Feature Based Methods

Automatic Feature Based Methods

Mathematical Framework for solving SDM Problems



Questions

Thank you !