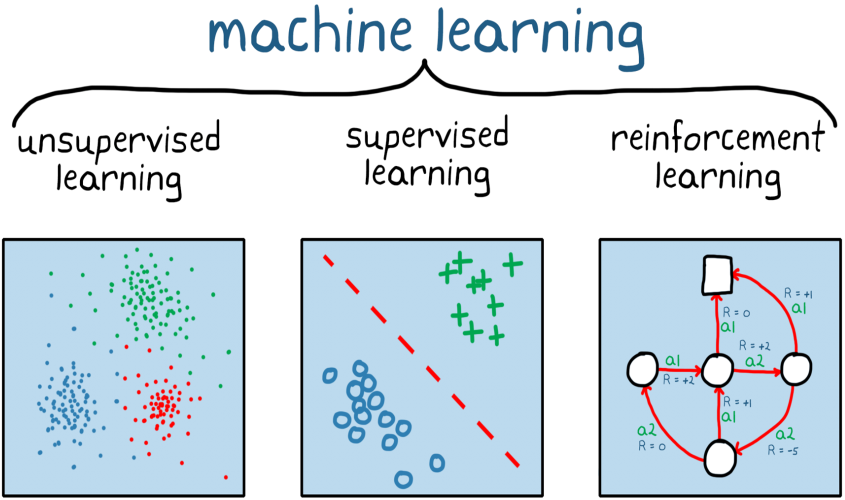
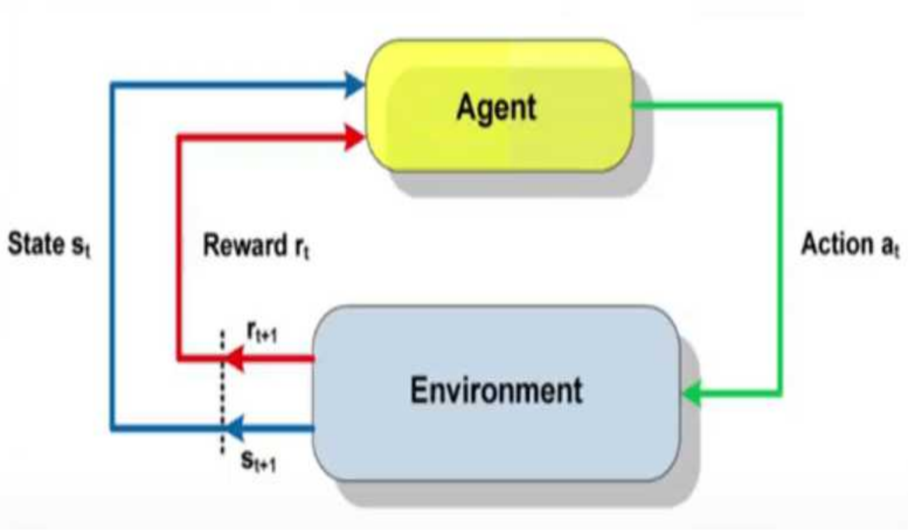
# RL in Finance and Retail Industries

Ashwin Rao

Stanford University


machine learning
unsupervised learning
supervised learning
reinforcement learning
R = +1
a1
R = 0
a1
R = +2
a2
a1
R = +2
R = +1
a1
a2
R = 0
a2
R = -5

# RL Process - *Actions* maximizing *Returns* in each *State*



# Many real-world problems fit this framework

- Self-driving vehicle (speed/steering to optimize safety/time)
- Game of Chess (Boolean *Reward* at end of game)
- Complex Logistical Operations (eg: movements in a Warehouse)
- Make a humanoid robot walk/run on difficult terrains
- Manage an investment portfolio
- Control a power station
- Optimal decisions during a football game
- Strategy to win an election (highly complex)

# Richard Bellman's Equation (from 75 years ago)

$$V^*(s) = \max_a \sum_{s'} \mathcal{P}(s, a, s') \cdot (\mathcal{R}(s, a, s') + \gamma \cdot V^*(s'))$$

- Recursive formulation of the Optimal Expected Return
- Gave us elegant Dynamic Programming algorithms
- Assuming we have a model of transition probabilities
- Impractical, doesn't scale (Curse of Dimensionality/Modeling)
- To resolve both curses effectively, we need RL

# Reinforcement Learning (RL)

- RL is a "trial-and-error" approach linking *Actions* to *Returns*
- Try different actions & learn what works, what doesn't
- RL incrementally learns from a stream of data through interactions
- Typically interacting with a *simulated* environment (*sampling model*)
- **Core Sauce**: Sampling & Function Approximation come together
- Deep Neural Networks for function approx. was RL's game-changer

# Future of RL

- **Promise of modern A.I. is based on success of RL algorithms**
- Agents that efficiently adapt to changing circumstances
- Sequentially optimizing to achieve business goals
- Possibilities in Finance and Retail industries are endless

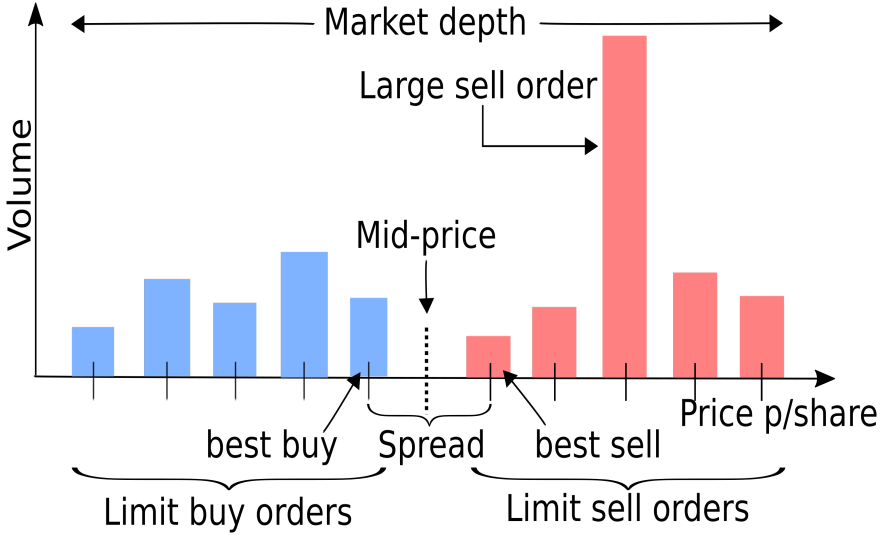
# Dynamic Asset-Allocation and Consumption

- The broad topic is Investment Management
- Applies to Corporations as well as Individuals
- The two considerations are:
  - How to allocate money across assets in one's investment portfolio
  - How much to consume for one's needs/operations/pleasures
- Asset-Allocation and Consumption decisions at each time step
- Asset-Allocation decisions typically deal with Risk-Reward tradeoffs
- Consumption decisions are about spending now or later

# Consider the simple example of Personal Finance

- Broadly speaking, Personal Finance involves the following aspects:
    - Receiving Money: Salary, Bonus, Rental income, Asset Liquidation etc.
    - Consuming Money: Food, Clothes, Rent/Mortgage, Car, Vacations etc.
    - Investing Money: Savings account, Stocks, Real-estate, Gold etc.
- Goal: Maximize lifetime-aggregated Expected Utility of Consumption
- We can model this in the RL framework as follows:
- *State:* Age, Asset Holdings, Asset Valuation, Career situation etc.
- *Action:* Changes in Asset Holdings, Optional Consumption
- *Reward:* Utility of Consumption of Money
- *Model:* Career uncertainties, Asset market uncertainties

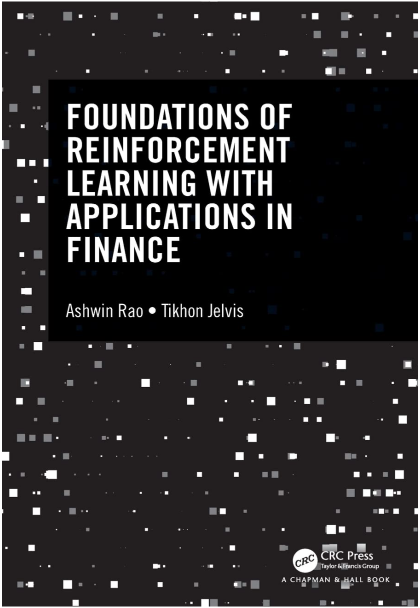
# Trading Order Book - various applications of RL



# Some more applications I've been excited about

- Financial Derivatives Pricing and Hedging
- Inventory Control in Supply-Chain
- Promotional Pricing and Clearance Pricing

# Book


FOUNDATIONS OF REINFORCEMENT LEARNING WITH APPLICATIONS IN FINANCE
Ashwin Rao • Tikhon Jelvis
CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK