Individual & Collective Intelligence

Xu Chen

Professor, Ph.D. Advisor

School of Data and Computer Science
Sun Yat-sen University, Guangzhou, China

Intelligence

Individual Intelligence

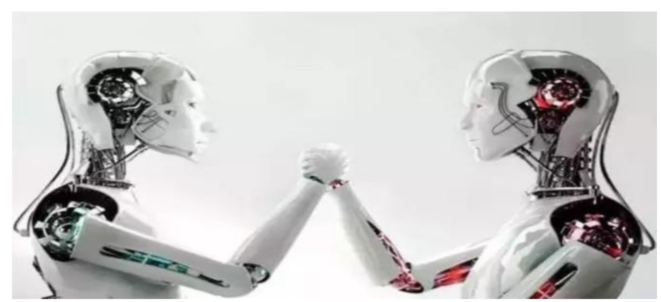
Reinforcement Learning

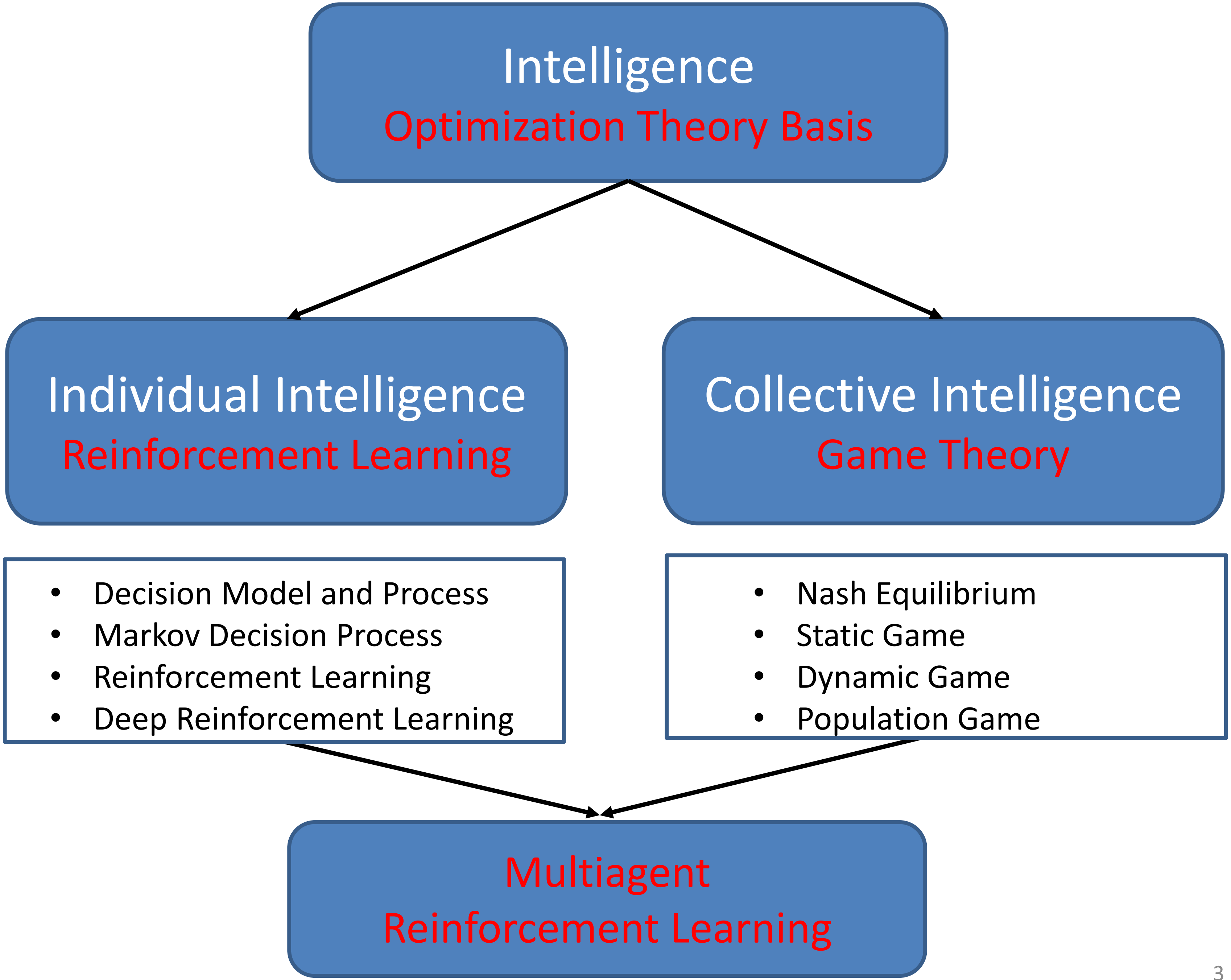
Collective Intelligence

Game Theory

AlphaGo

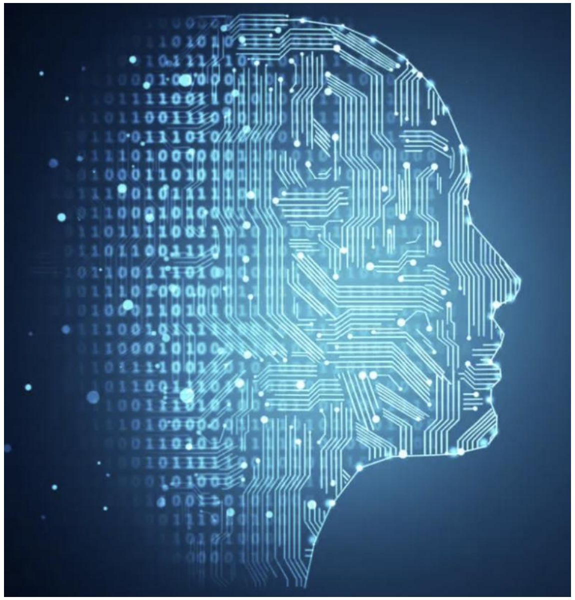
Pluribus




Intelligence
Optimization Theory Basis
Individual Intelligence
Reinforcement Learning
Collective Intelligence
Game Theory
• Decision Model and Process
• Markov Decision Process
• Reinforcement Learning
• Deep Reinforcement Learning
• Nash Equilibrium
• Static Game
• Dynamic Game
• Population Game
Multiagent
Reinforcement Learning

Individual Intelligence

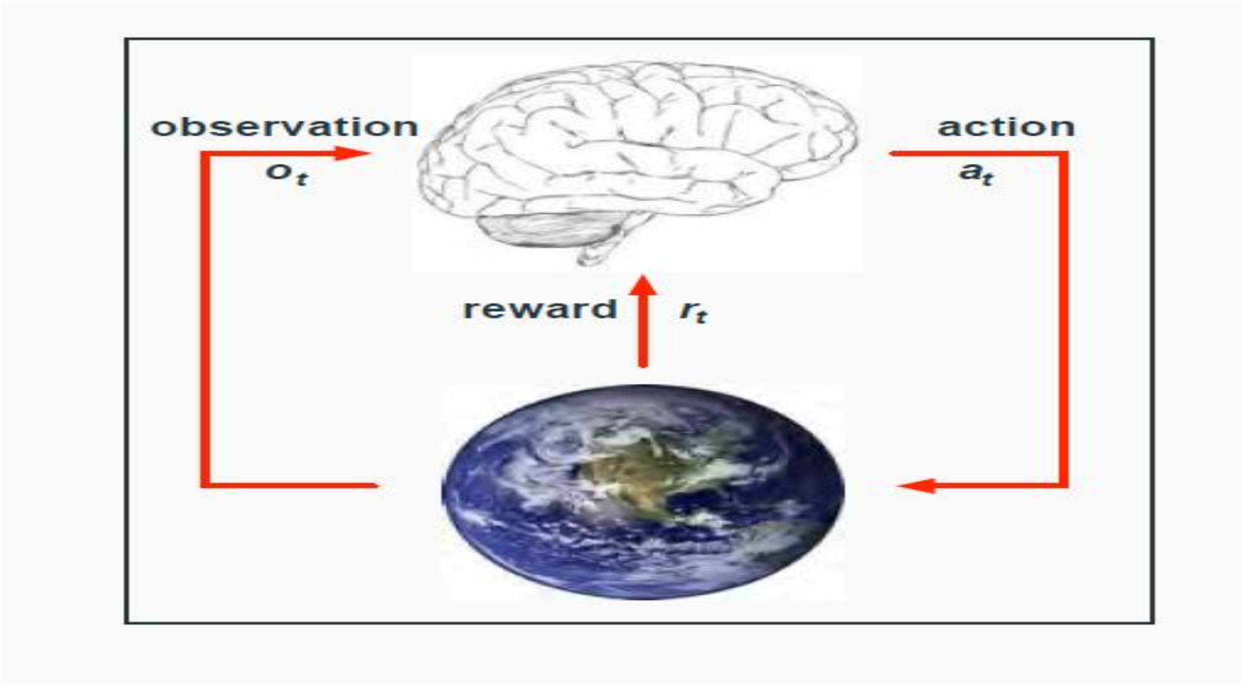
A Brief Introduction to Reinforcement Learning



Fundamental challenge in artificial intelligence is learning to make good decisions under uncertainty

What is reinforcement learning?

A computational approach to learning whereby an agent tries to maximize the total amount of reward it receives while interacting with a complex and uncertain environment.

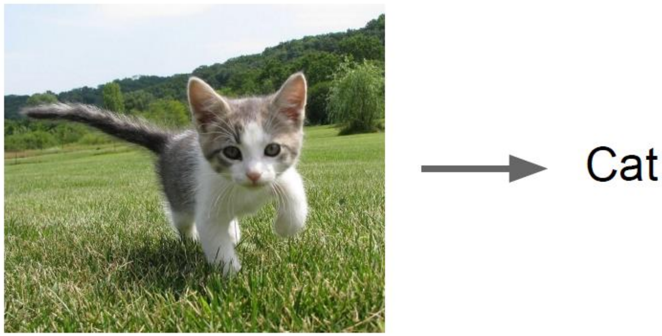


Supervised Learning

Data: (x, y)
x is data, y is label

Goal: Learn a *function* to map x -> y

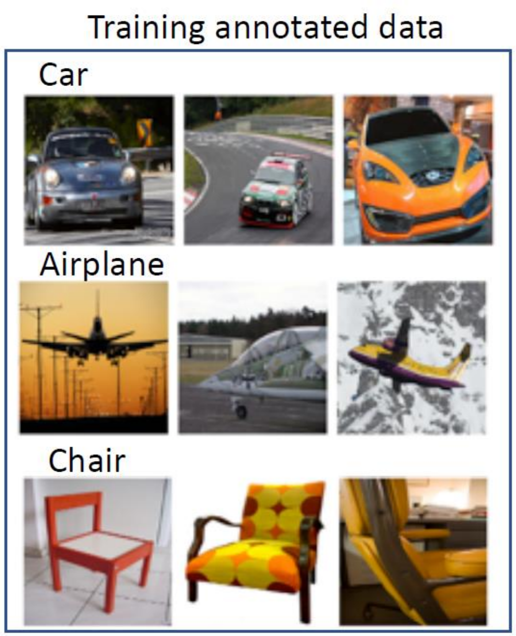
Examples: Classification, regression, object detection, semantic segmentation, image captioning, etc.

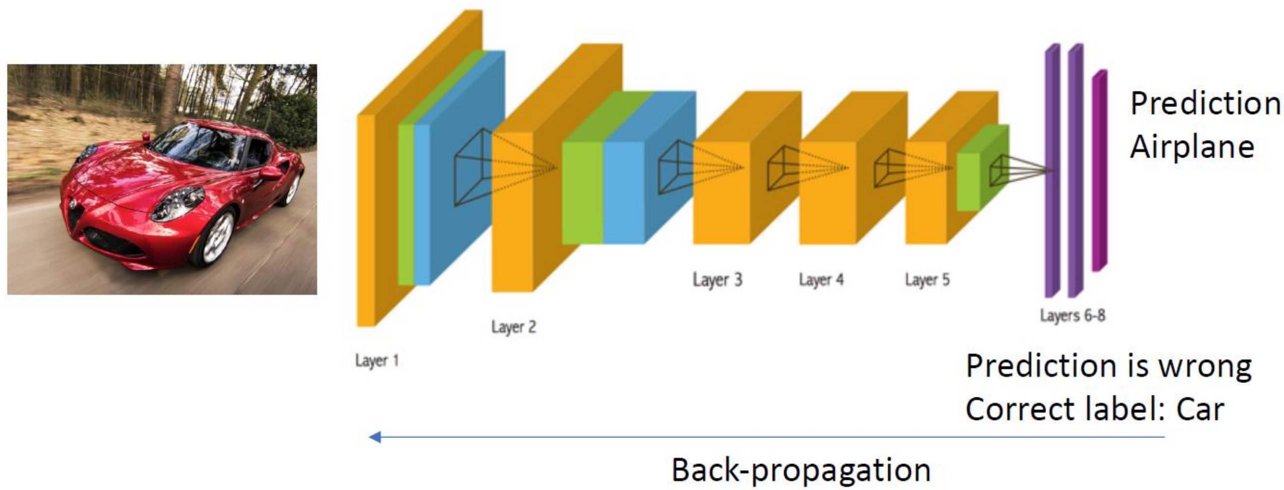


Classification

Supervised Learning: Image Classification

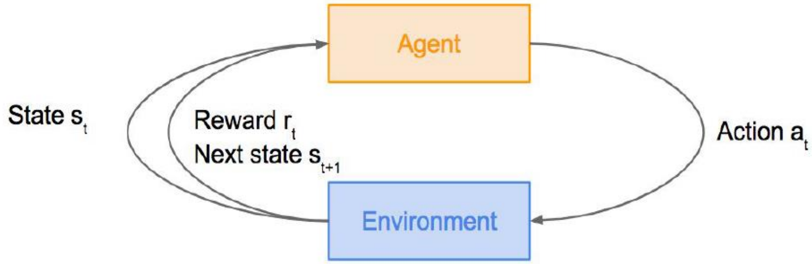
- Annotated images, data follows i.i.d distribution
- Learners are told what the labels are





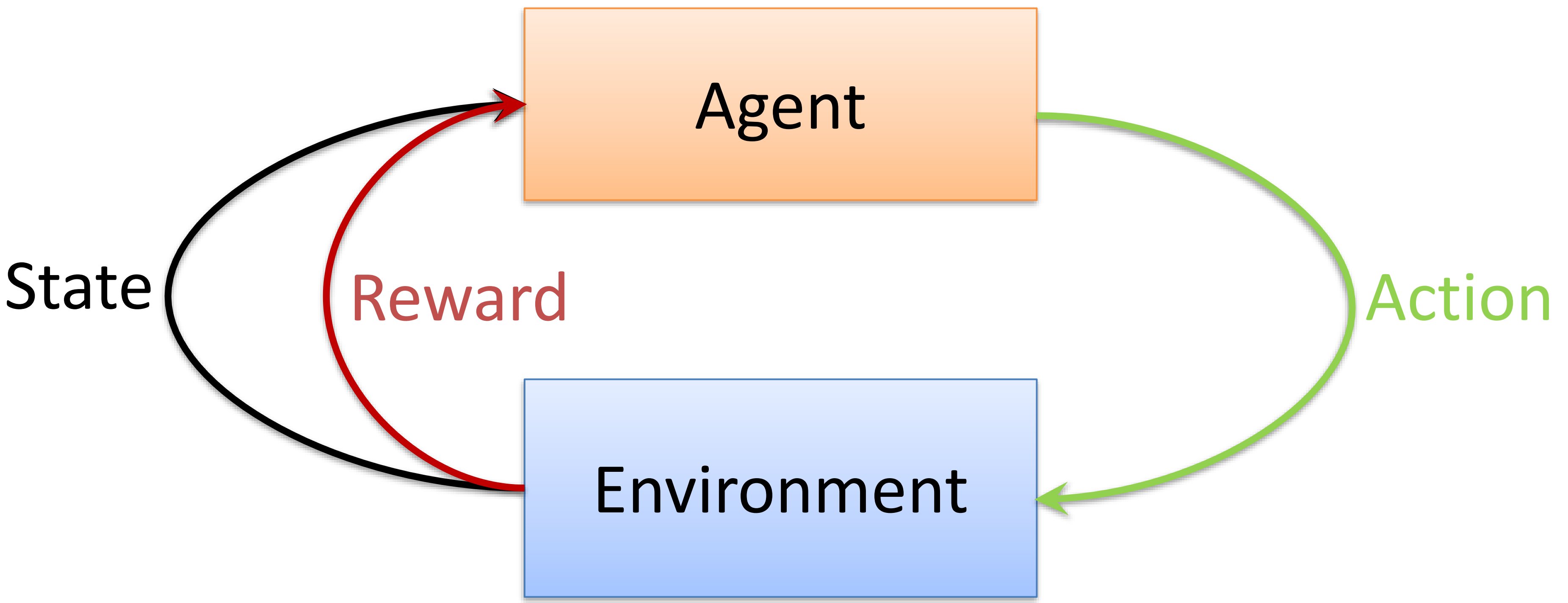
Reinforcement Learning

Problems involving an **agent** interacting with an **environment**, which provides numeric **reward** signals

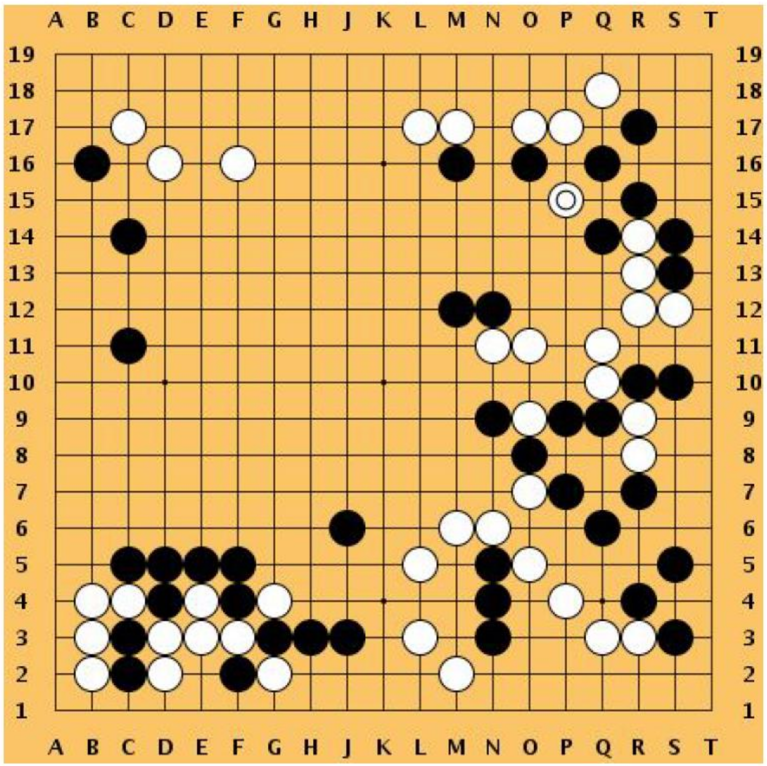


Goal: Learn how to take actions in order to maximize reward

Reinforcement Learning



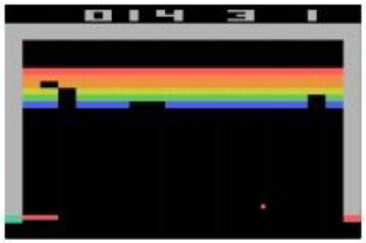
Go



Objective: Win the game!

State: Position of all pieces
Action: Where to put the next piece down
Reward: 1 if win at the end of the game, 0 otherwise

Atari Games

Objective: Complete the game with the highest score

State: Raw pixel inputs of the game state
Action: Game controls e.g. Left, Right, Up, Down
Reward: Score increase/decrease at each time step

Difference between Reinforcement Learning and Supervised Learning

• Sequential data as input (not i.i.d)

• The learner is not told which actions to take, but instead must discover which actions yield the most reward by trying them.

• Trial-and-error exploration (balance between exploration and exploitation)

• There is no supervisor, only a reward signal, which could also be delayed

Big deal: Able to Achieve Superhuman Performance

- Upper bound for supervised learning is human-performance
- Upper bound for reinforcement learning?

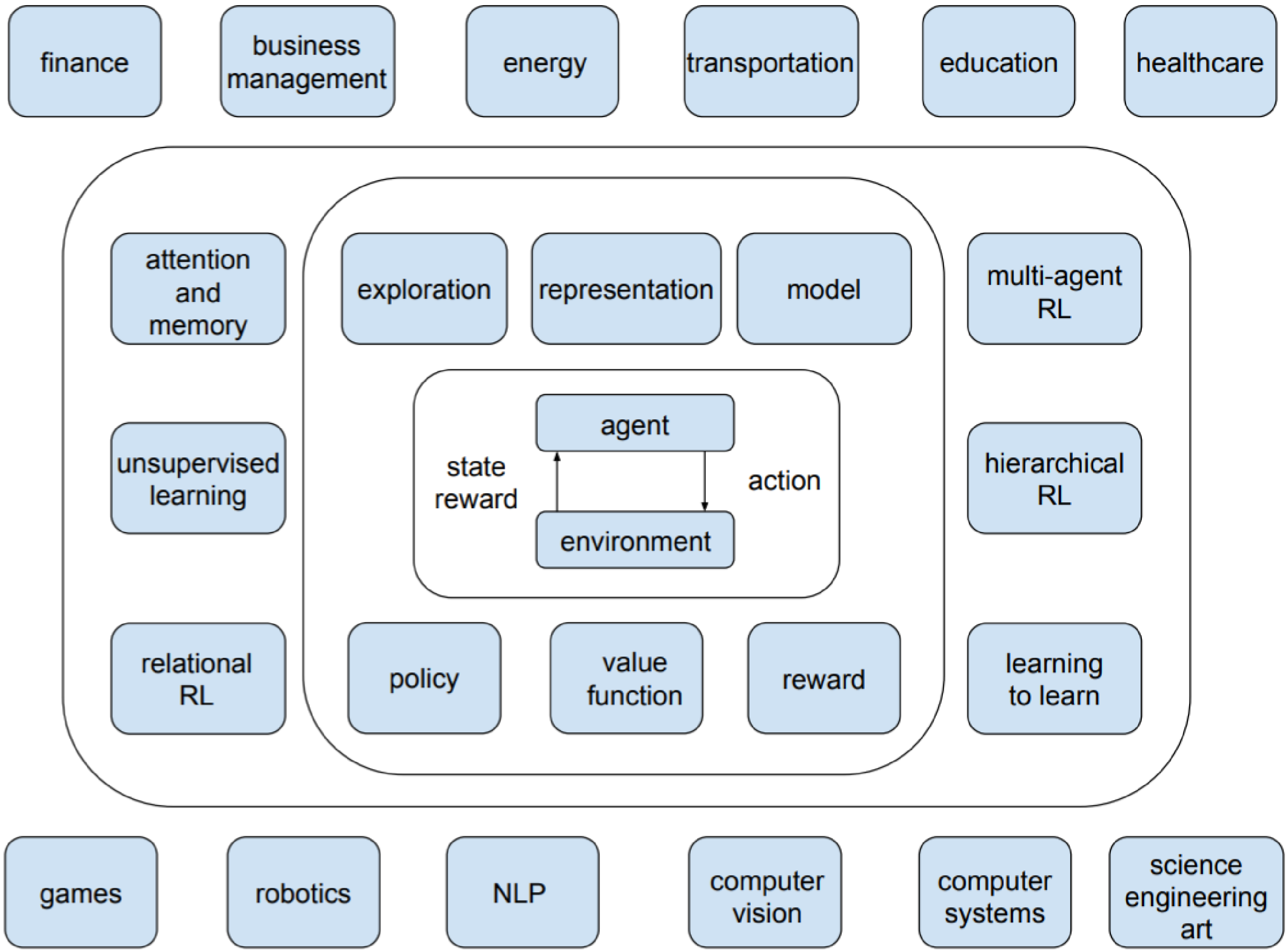
Why Reinforcement Learning Works Now?

- Computation power: many GPUs to do trial-and-error rollout
- Acquire the high degree of proficiency in domains governed by simple and known rules; huge volume of data samples available
- End-to-end deep learning based training, features and policy are jointly optimized toward the end goal

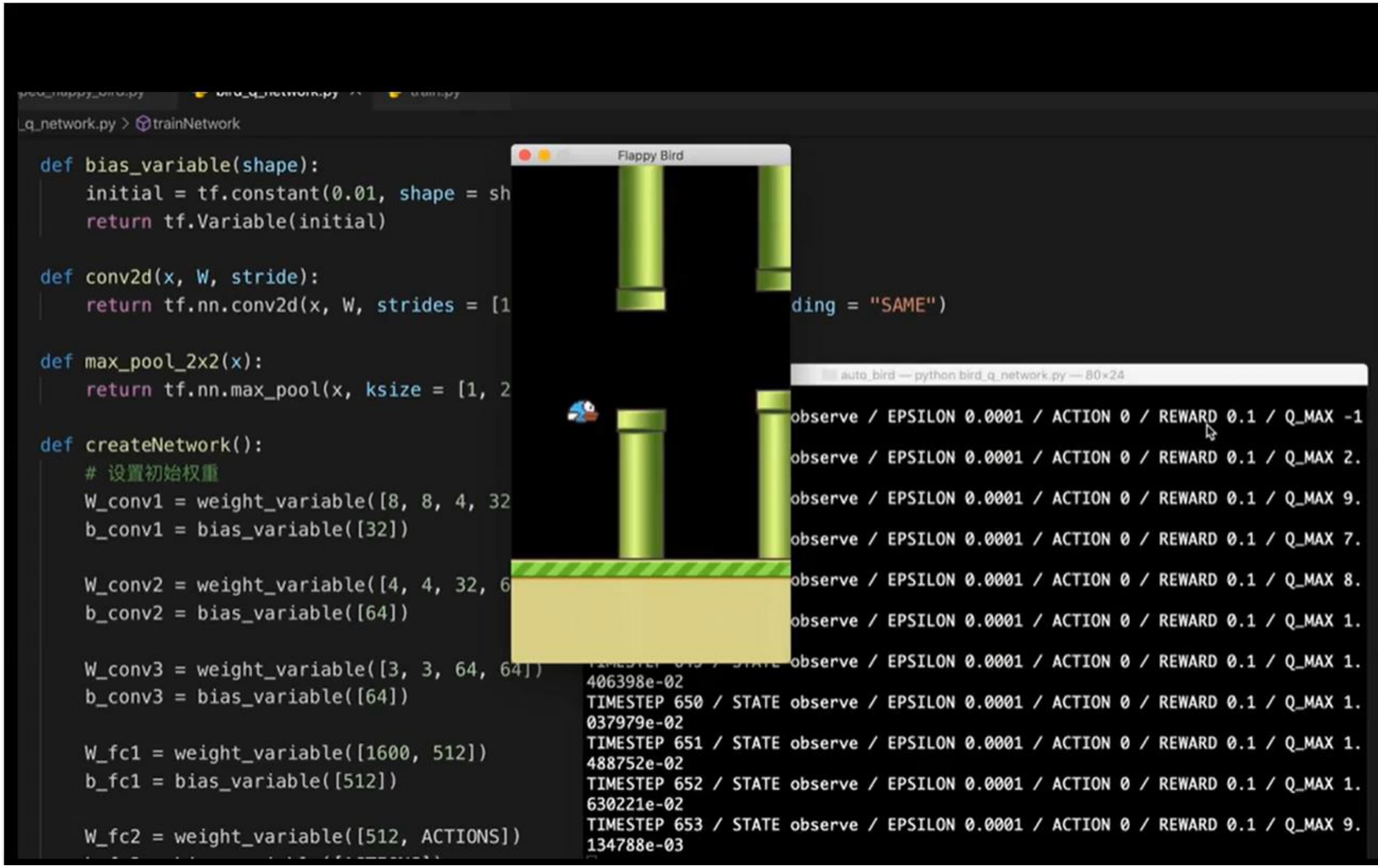
Game playing

Robotics

Beating best human player

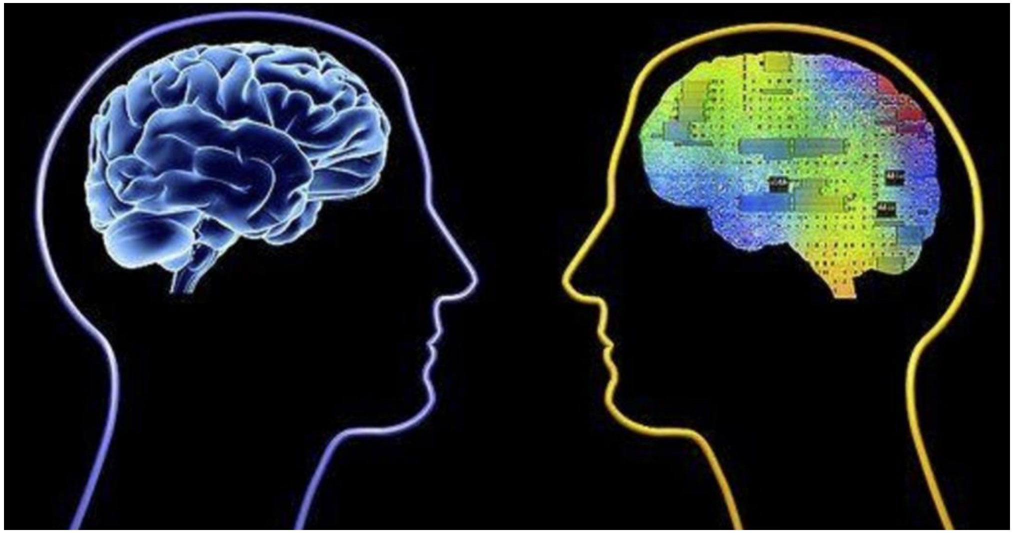

finance
business management
energy
transportation
education
healthcare
attention and memory
exploration
representation
model
multi-agent RL
agent
state
reward
action
environment
unsupervised learning
hierarchical RL
relational RL
policy
value function
reward
learning to learn
games
robotics
NLP
computer vision
computer systems
science engineering art

Reinforcement Learning: Flappy Bird



Collective Intelligence

A Brief Introduction to Game Theory



Game Theory

Rational – user aims to optimize its own objective

Interaction – user needs to take others' decisions into account

"…Game Theory is designed to address situations in which the outcome of a person's decision depends not just on how they choose among several options, but also on the choices made by the people they are interacting with…" --David Easley and Jon Kleinberg

"… Game theory is the study of the ways in which *strategic interactions* among *rational agents* produce *outcomes* with respect to the *utilities* of those agents …." --Stanford Encyclopedia of Philosophy

A Brief History

- 1944: Von Neumann and Oskar Morgenstern

  Theory of Games and Economic Behavior

  Two-player games

- 1950: John Nash

  Nash Equilibrium

  Equilibrium points in n-player games

- After 1950s: widely used in economics, politics, biology…

  Competition between firms

  Auction design

  Role of punishment in law enforcement

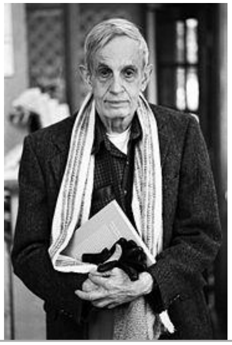
  International policies

  Evolution of species

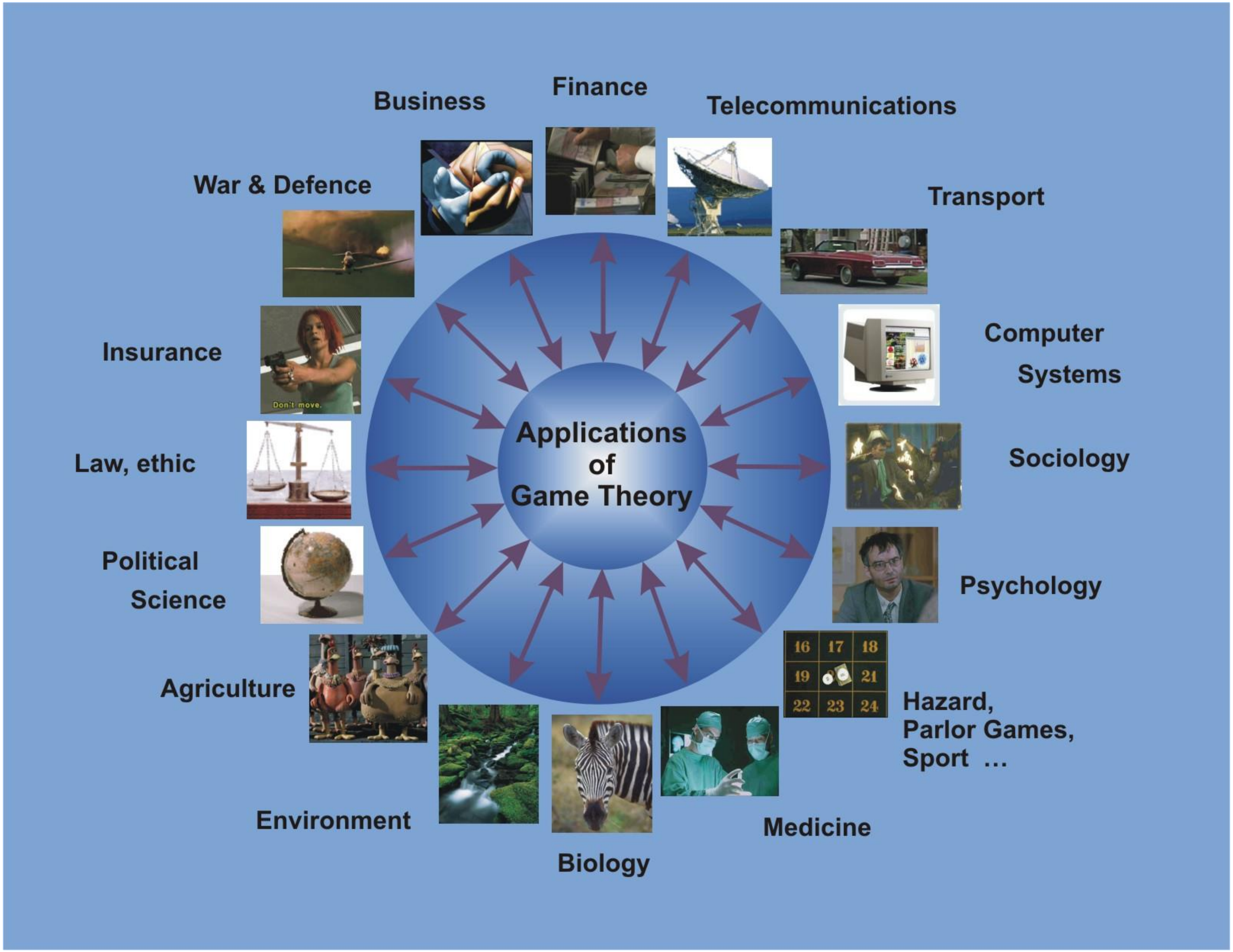
  Artificial Intelligence/machine learning

O. Morgenstern 1902-1977

von Neumann 1903-1957

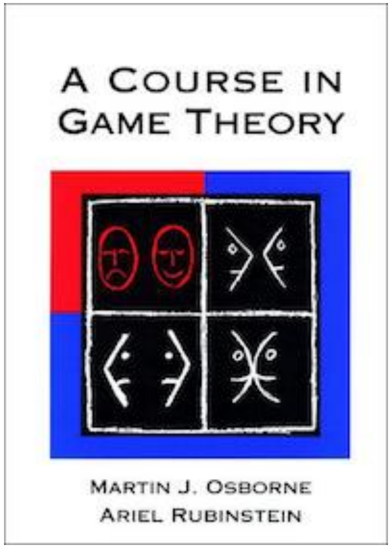


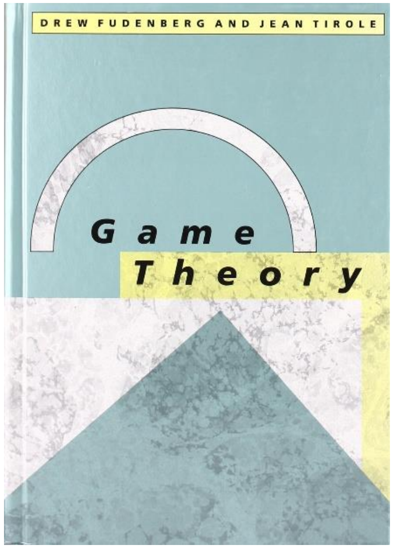
John Nash 1928-2015


Applications
of
Game Theory
Business
Finance
Telecommunications
War & Defence
Transport
Computer
Systems
Insurance
Law, ethic
Sociology
Political
Science
Psychology
Agriculture
Hazard,
Parlor Games,
Sport …
Environment
Medicine
Biology

Relevance to Computing Research

- Economic issues become increasingly important
  - Interactions with/between human users
    e.g., data-driven pricing, resource allocation (Urban/Amazon/DiDi/Taobao)
  - Independent service providers
    e.g., bandwidth trading, peering agreements
- Tool for smart system design
  - Distributed Intelligent algorithms
  - Multi-objective optimization
  - Incentive compatible protocols





Game Theory Basics

- Strategic game form (P, S, U)
  - Players $(P_1, \dots, P_N)$: finite number of decision makers
  - Strategy sets $(S_1, \dots, S_N)$: player P_i has a nonempty set S_i of actions/strategies s_i
  - Payoff function $U_i(s_1, \dots, s_N)$: player's preference/individual utility

- Nash equilibrium (NE)
  - A strategy profile $(s_1^*, \dots, s_i^*, \dots, s_N^*)$ is a NE if for each player i

$$U_i(s_1^*, \dots, s_i^*, \dots, s_N^*) \geq U_i(s_1^*, \dots, s_i, \dots, s_N^*), \forall s_i \in S_i$$

  - No player has incentive to deviate (stable system point)
  - NE is a fixed point of the best response functions

$$s_i^* = \underset{s_i \in S_i}{\operatorname{argmax}} U_i(s_1^*, \dots, s_i, \dots, s_N^*), \forall i$$

- There is no universal rule for finding a Nash equilibrium!

Prisoner’s Dilemma

- Two suspects are arrested
- The police lack sufficient evidence to convict the suspects, unless at least one confesses
- The police hold the suspects in two separate rooms, and tell each of them three possible consequences:
  - If both deny: 1 month in jail each
  - If both confess: 6 months in jail each
  - If one confesses and one denies:
    - The one confesses: walk away free of charge
    - The one denies: serve 12 months in jail

Prisoner’s Dilemma

strategies

		Player 2	
		Deny	Confess
Player 1	Deny	$-1, -1$	$-12, 0$
	Confess	$0, -12$	$-6, -6$

payoffs

Prisoner's Dilemma

- Strictly dominated strategy
  - Player i's strategy s_i' is strictly dominated by player i's strategy s_i if

$$U_i(s_i, s_{-i}) > U_i(s_i', s_{-i}), \forall s_{-i}$$

    where s_{-i} is the strategy profile of all the other players except player i
  - No matter what other people do, by choosing s_i instead of s_i', player $i$will always obtain a better payoff
  - Key principle: Never play a strictly dominated strategy

Prisoner's Dilemma

Player 2's choice

		Player 2	
		Deny	Confess
Player 1	Deny	$-1, -1$	$-12, 0$
	Confess	$0, -12$	$-6, -6$

Player 1's choice

NE of the game

Deny is strictly dominated by Confess!

Finding Nash Equilibrium

- When there are no strictly dominated strategies, we can not easily "simplify" the game
- Nash equilibrium is a state of mutual best responses
- Key principle: derive the best responses

Stag Hunt

- Two hunters decide what to hunt independently
- Each one can hunt a stag (deer) or a hare
- Successful hunt of stag requires cooperation
- Successful hunt of hare can be done individually
- Simultaneous decisions without prior communications

Stag Hunt

		Player 2	
		Stag	Hare
Player 1	Stag	5 , 5	0 , 2
	Hare	2 , 0	2 , 2

There is no strictly dominated strategy

Find out a player's best response given the other player's choice

Stag Hunt

Given Player 2 chooses Stag

Player 1's best response

Player 1 \ Player 2	Stag	Hare
Stag	$\boxed{5}$, 5	0 , 2
Hare	2 , 0	2 , 2

Stag Hunt

Given Player 2 chooses Hare

		Player 2	
		Stag	Hare
Player 1	Stag	$\boxed{5}$, 5	0 , 2
	Hare	2 , 0	$\boxed{2}$, 2

Player 1's best response

Stag Hunt

Player 1	Player 2	Stag	Hare
	Stag	$\boxed{5}$, $\boxed{5}$	0 , 2
	Hare	2 , 0	$\boxed{2}$, $\boxed{2}$

NE of the game

Player 2’s best responses

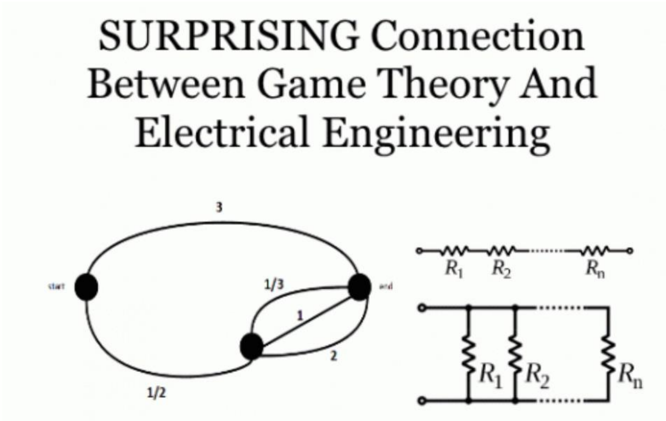
NE is a state of mutual best responses

Stag Hunt

- Two Nash equilibria exist

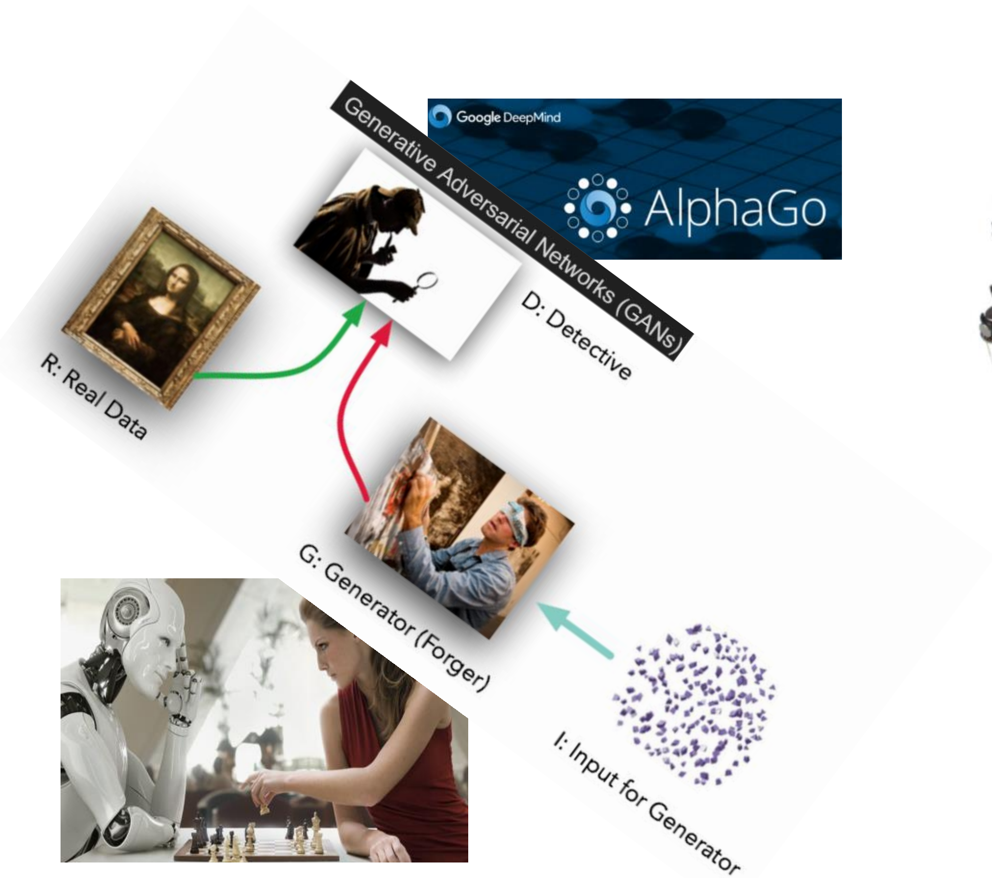
- (Stag, Stag) is payoff dominant
  - Both players get the best payoff possible
  - Require trust among players to achieve coordination

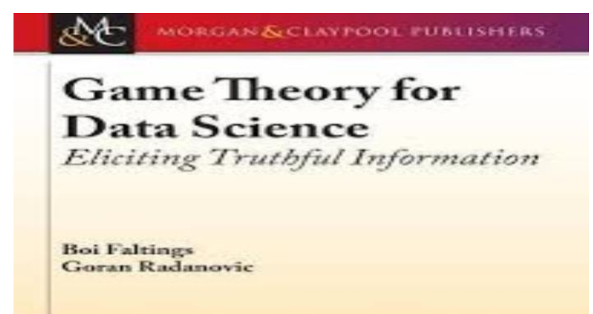
- (Hare, Hare) is risk dominant
  - Minimum risk if player is uncertain of each other’s choice



Using Computational Game Theory To Guide Verification and Security in Hardware Designs

Andrew M. Smith*†, Jackson R. Mayo‡, Vivian Kammler§, Robert C. Armstrong*, and Yevgeniy Vorobeychik¶
*Digital and Quantum Information Systems, Sandia National Laboratories, Livermore, California 94551–0969
Email: amsmit@sandia.gov
†Department of Computer Science, University of California, Davis, CA 95616–8562
‡Scalable Modeling and Analysis Systems, Sandia National Laboratories, Livermore, California 94551–0969
§Embedded Systems Analysis, Sandia National Laboratories, Albuquerque, NM 87185
¶Department of Computer Science, Vanderbilt University, Nashville, TN 37235





Reinforcement Learning + Game Theory = Multiagent Learning