# SSM Course

# MARL Systems

Mirco Musolesi

mircomusolesi@acm.org

# Definition of Multiagent Systems

- Several possible definitions:
  - *Multiagent systems are distributed systems of independent actors called agents that are each independently controlled but that interact with one another in the same environment.* (see: Wooldridge, “Introduction to Multiagent Systems”, 2002 and Tulys and Stone, “Multiagent Learning Paradigms”, 2018).
  - *Multiagent systems are systems that include multiple autonomous entities with (possibly) diverging information* (see Shoham and Leyton-Brown, “Multiagent systems”, 2009).

# Definition of Multiagent Learning

- We will use the following definition of multiagent learning:
    - *"The study of multiagent systems in which one or more of the autonomous entities improves automatically through experience*".

# Characteristics of Multiagent Learning

- Different scale:
  - A city or an ant colony or a football team.
- Different degree of complexity:
  - A human, a machine, a mammal or an insect.
- Different types of interaction:
  - Frequent interactions (or not), interactions with a limited number of individuals, etc.

# Presence of Regularity

- It is fundamental that there is a certain degree of regularity in the system otherwise prediction of behaviour is not possible.
- Assumption: past experience is somehow predictive of future expectations.
- Dealing with non-stationarity is a key problem.
  - It is the usual problem of reinforcement learning at the end.

# Potential Paradigms

- We will consider 5 paradigms:
  - Online Multi-agent Reinforcement Learning towards individual utility
  - Online Multi-agent Reinforcement Learning towards social welfare
  - Co-evolutionary learning
  - Swarm intelligence
  - Adaptive mechanism design

# Online Reinforcement Learning towards Individual Utility

- One of the most-studied scenarios in multiagent learning is that in which multiple independent agents take actions in the same environment and learn online to maximise their own individual utility functions (i.e., expected returns).
- From a formal point view (game-theory point of view), this can be considered a repeated normal form game.
  - A *repeated game* is a game that is based of a certain number of repetitions.
  - *Normal form games* are games that are presented using a matrix.
    - As aside, an *extensive form game* is a game for which an explicit representation of the sequence of the players' possible moves, their choices at every decision point and the information about other player's move and relative payoffs.

# Example: Prisoner's Dilemma

- The Prisoner's Dilemma is a classic 2-player game.
- Description of the "game": two prisoners committed a crime together and are being interrogated separately.
- If neither of them confesses to the crime (they both "cooperate"), then they will both get a small punishment (corresponding to a payoff of 5).
- If one of them confesses (or "defects"), but the other does not, then the one that confesses gets off for free (payoff of 10), but the other gets the worst punishment possible (payoff of 0).
- If they both defect, they get a worst punishment (payoff of 1)

# Prisoners' Dilemma

| | Defect | Cooperate |
|---|---|---|
| Defect | (1, 1) | (10, 0) |
| Cooperate | (0, 10) | (5, 5) |

# Example: Prisoner's Dilemma

- Normal games were initially introduced as one-shot game.
- The players know each other's full utility (reward) functions and play the game only once.
- In this setting, the concept of Nash equilibrium was introduced: a set of actions such that no player would be better off deviating given that the other player's actions are fixed.
- Games can have one or multiple Nash equilibria.
- In the Prisoner's Dilemma, the only Nash Equilibrium is for both agents to defect.

[13] Whitehead, J. H. C., "Simple Homotopy Types." If $W = 1$, Theorem 5 follows from (17:3) on p. 155 of S. Lefschetz, *Algebraic Topology*, (New York, 1942) and arguments in §6 of J. H. C. Whitehead, "On Simply Connected 4-Dimensional Polyhedra" (*Comm. Math. Helv.*, 22, 48–92 (1949)). However this proof cannot be generalized to the case $W \neq 1$.

---

# *EQUILIBRIUM POINTS IN N-PERSON GAMES*

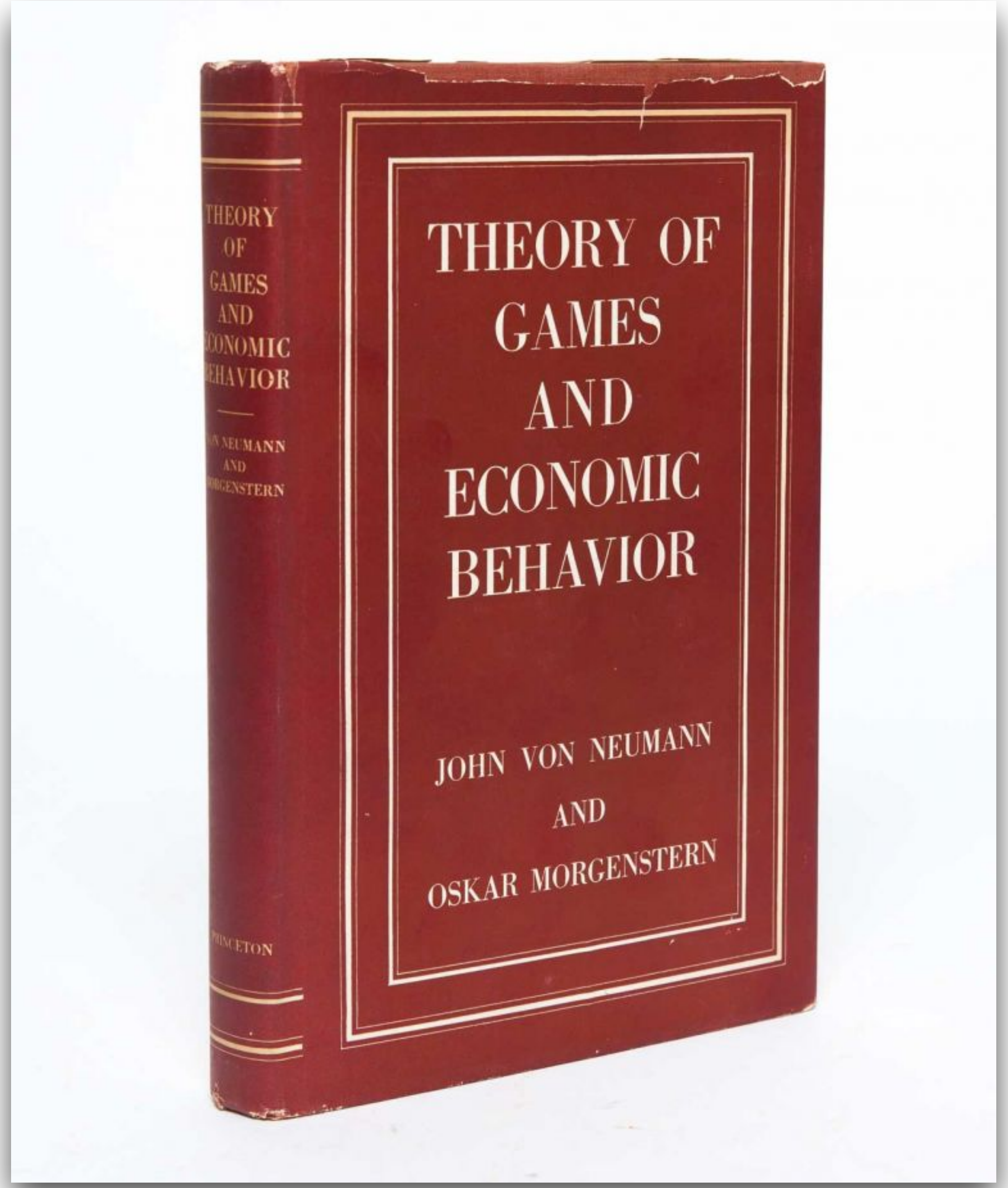

BY JOHN F. NASH, JR.*

PRINCETON UNIVERSITY



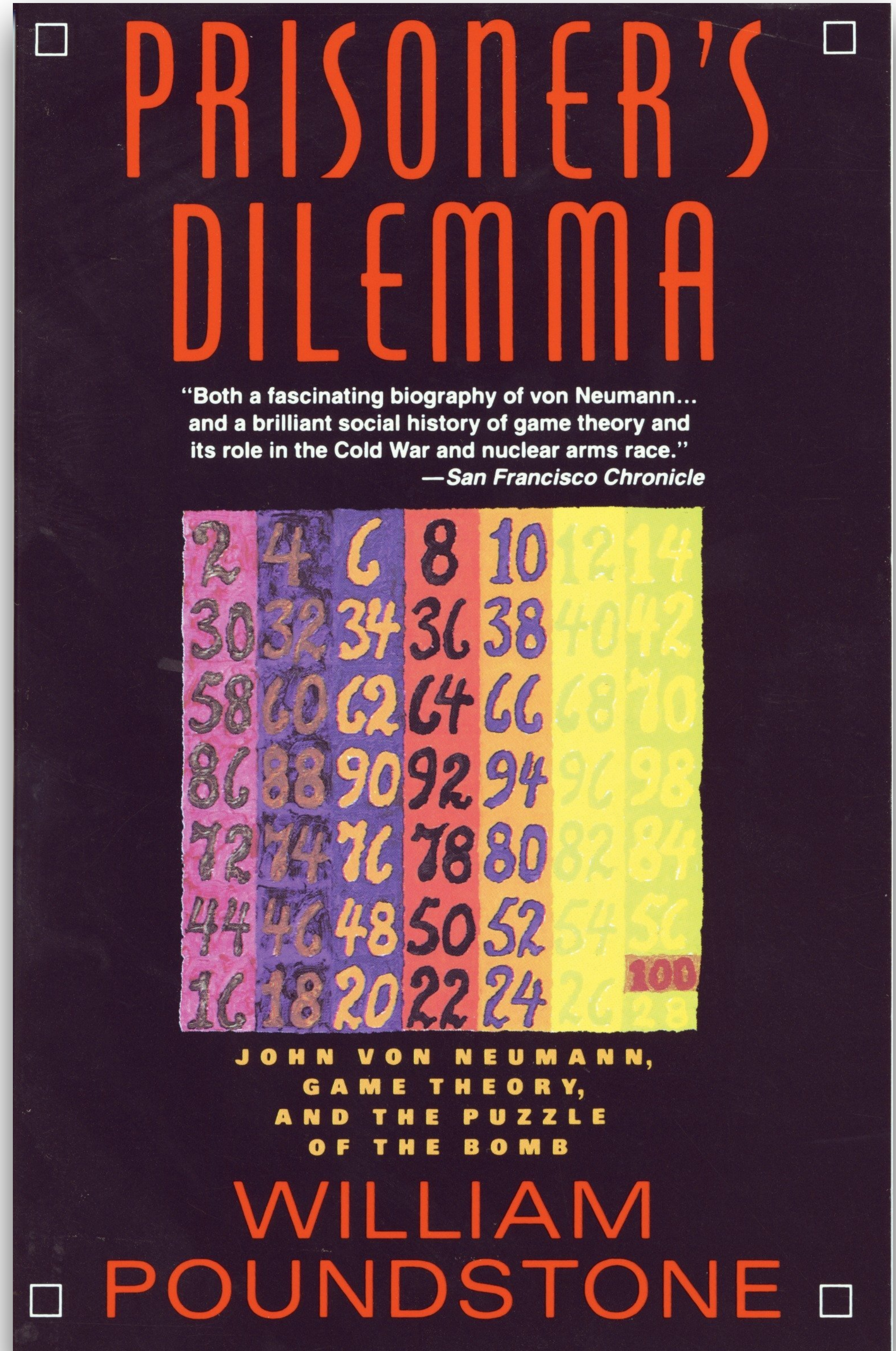
Communicated by S. Lefschetz, November 16, 1949

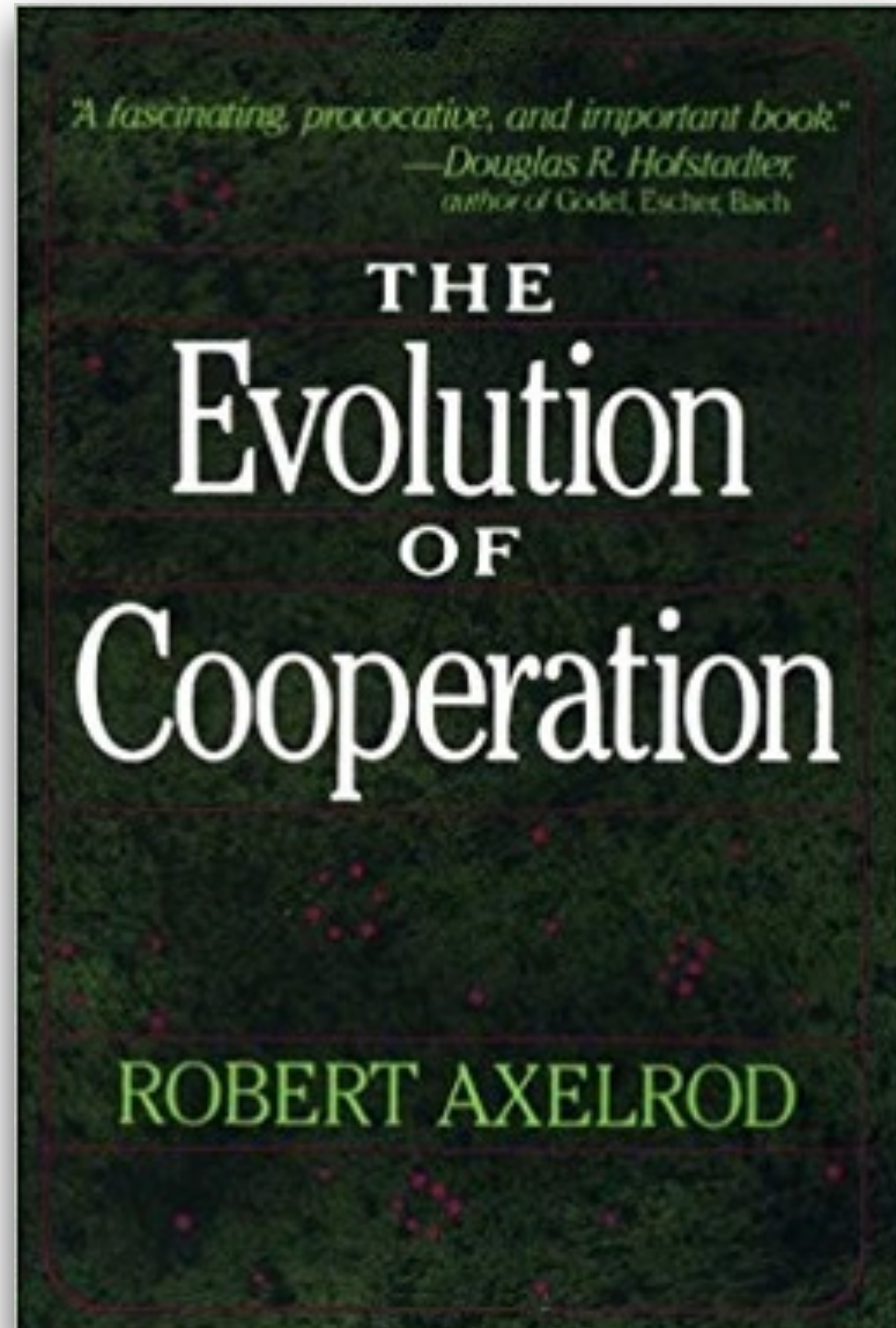

One may define a concept of an $n$-person game in which each player has a finite set of pure strategies and in which a definite set of payments to the $n$ players corresponds to each $n$-tuple of pure strategies, one strategy being taken for each player. For mixed strategies, which are probability

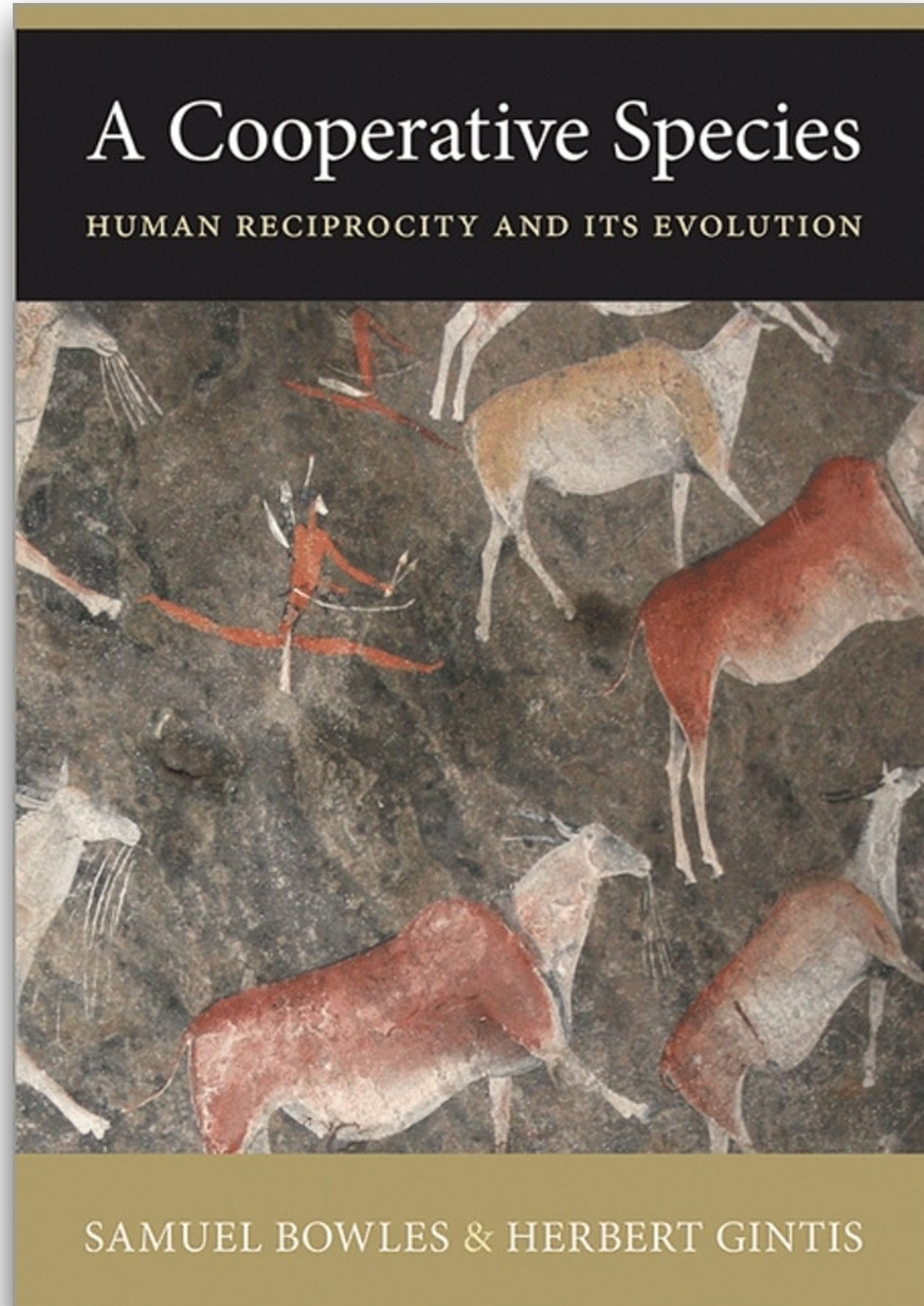

THEORY OF
GAMES
AND
ECONOMIC
BEHAVIOR
JOHN VON NEUMANN
AND
OSKAR MORGENSTERN
PRINCETON

# Repeated Normal Form Games

- In repeated normal form games, players interact with one another multiple times with the objectives of maximising their sum utilities (i.e., expected returns) over time.
- As you can imagine, Reinforcement Learning and possibly Deep Reinforcement Learning is well suited for this type of problems.
- Reinforcement Learning can also be used to understand the problem of the “evolution of cooperation” and the presence of altruism: why do we humans cooperate even if in presence of maximisation of personal reward function?


PRISONER'S
DILEMMA
"Both a fascinating biography of von Neumann…
and a brilliant social history of game theory and
its role in the Cold War and nuclear arms race."
—San Francisco Chronicle
JOHN VON NEUMANN,
GAME THEORY,
AND THE PUZZLE
OF THE BOMB
WILLIAM
POUNDSTONE


"A fascinating, provocative, and important book."
—Douglas R. Hofstadter,
author of Godel, Escher, Bach
THE
Evolution
OF
Cooperation
ROBERT AXELROD


A Cooperative Species
HUMAN RECIPROCITY AND ITS EVOLUTION
SAMUEL BOWLES & HERBERT GINTIS

# Online Reinforcement Learning towards Social Welfare

- An alternative paradigm is the one in which multiple independent agents take actions in the same environment and learn online to maximise a global utility function.

- These are also called coordination games, where different players coordinate to achieve a given objective (i.e., global expected return).

# Multiagent Cooperation and Competition with Deep Reinforcement Learning

**Ardi Tampuu*** **Tambet Matiisen***

**Dorian Kodelja** **Ilya Kuzovkin** **Kristjan Korjus**

**Juhan Aru**[†] **Jaan Aru** **Raul Vicente**[✉]

Computational Neuroscience Lab, Institute of Computer Science, University of Tartu
[†] Department of Mathematics, ETH Zurich
ardi.tampuu@ut.ee, tambet.matiisen@ut.ee, raul.vicente.zafra@ut.ee

** these authors contributed equally to this work*

## Abstract

Multiagent systems appear in most social, economical, and political situations. In the present work we extend the Deep Q-Learning Network architecture proposed by Google DeepMind to multiagent environments and investigate how two agents controlled by independent Deep Q-Networks interact in the classic videogame *Pong*. By manipulating the classical rewarding scheme of Pong we demonstrate how competitive and collaborative behaviors emerge. Competitive agents learn to play and score efficiently. Agents trained under collaborative rewarding schemes find an optimal strategy to keep the ball in the game as long as possible. We also describe the progression from competitive to collaborative behavior. The present work demonstrates that Deep Q-Networks can become a practical tool for studying the decentralized learning of multiagent systems living in highly complex environments.

DeepMind

2020-12-17

# Open Problems in Cooperative AI

**Allan Dafoe[1], Edward Hughes[2], Yoram Bachrach[2], Tantum Collins[2], Kevin R. McKee[2], Joel Z. Leibo[2], Kate Larson[2, 3] and Thore Graepel[2]**

[1]Centre for the Governance of AI, Future of Humanity Institute, University of Oxford, [2]DeepMind, [3]University of Waterloo

**Problems of cooperation—in which agents seek ways to jointly improve their welfare—are ubiquitous and important. They can be found at scales ranging from our daily routines—such as driving on highways, scheduling meetings, and working collaboratively—to our global challenges—such as peace, commerce, and pandemic preparedness. Arguably, the success of the human species is rooted in our ability to cooperate. Since machines powered by artificial intelligence are playing an ever greater role in our lives, it will be important to equip them with the capabilities necessary to cooperate and to foster cooperation.**

# References


- Stefano V. Albrecht and Peter Stone. Autonomous Agents Modelling Other Agents: A Comprehensive Survey and Open Problems. Artificial Intelligence. Volume 258. 2018.
- Lucian Busoniu, Robert Babuska, Bart De Schutter. A Comprehensive Survey of Multiagent Reinforcement Learning. In IEEE Transactions on Systems, Man and Cybernetics. Volume 38. Issue 2. March 2008.

# References


- Kevin Leyton-Brown and Yoav Shoham. Multiagent Systems, Game-theoretic and Logical Foundations. Cambridge University Press. 2009.

- Karl Tuyls and Gerhard Weiss. Multiagent Learning: Basics, Challenges and Prospects. AI Magazine. Volume 33. Issue 3. 2012.

# References


- Karl Tulys and Peter Stone. Multiagent Learning Paradigms. In Francesco Belardinelli and Estefania Argente, editors, Multi-agent Systems and Agreement Technologies. Lecture Notes in Artificial Intelligence. Pages 3-21. Springer 2018.

- Micheal Wooldridge. An Introduction to MultiAgent Systems. Second Edition. Wiley. 2009.

# References


- Kaiqing Zhang, Zhuoran Yang and Tamer Basar. Multi-agent Reinforcement Learning: A Selective Overview of Theories and Algorithms. arXiv:1911.10635v2. 2021.
- Sven Gronauer and Klaus Diepold. Multi-agent Reinforcement Learning: A Survey. Artificial Intelligence Review. 55:895-943. Springer. 2022.