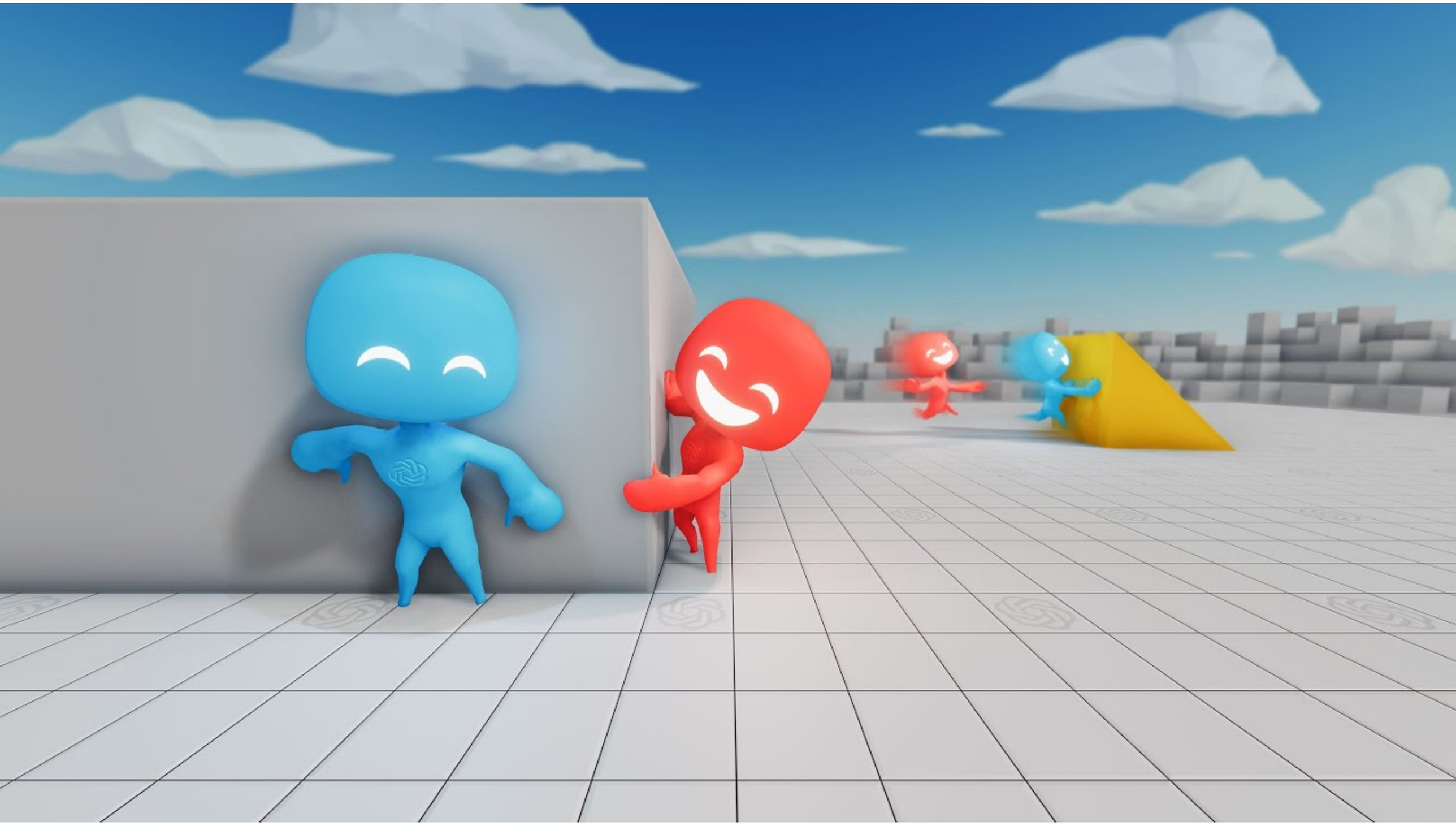
Multi-agent Reinforcement Learning

Cooperative AI

Mirco Musolesi

mircomusolesi@acm.org



EMERGENT TOOL USE FROM MULTI-AGENT AUTOCURRICULA

Bowen Baker*
OpenAI
bowen@openai.com

Ingmar Kanitscheider*
OpenAI
ingmar@openai.com

Todor Markov*
OpenAI
todor@openai.com

Yi Wu*
OpenAI
jxwuyi@openai.com

Glenn Powell*
OpenAI
glenn@openai.com

Bob McGrew*
OpenAI
bmcgrew@openai.com

Igor Mordatch*†
Google Brain
imordatch@google.com

ABSTRACT

Through multi-agent competition, the simple objective of *hide-and-seek*, and standard reinforcement learning algorithms at scale, we find that agents create a self-supervised autocurriculum inducing multiple distinct rounds of emergent strategy, many of which require sophisticated tool use and coordination. We find clear evidence of six emergent phases in agent strategy in our environment, each of which creates a new pressure for the opposing team to adapt; for instance, agents learn to build multi-object shelters using moveable boxes which in turn leads to agents discovering that they can overcome obstacles using ramps. We further provide evidence that multi-agent competition may scale better with increasing environment complexity and leads to behavior that centers around far more human-relevant skills than other self-supervised reinforcement learning methods such as intrinsic

Cooperative AI

- We will focus on scenario in which agents try to improve their global degree of cooperation.

- We will consider multiple aspects of the problem, considering the potential applications to a series of practical situations.

Common and Conflicting Interests

- **Games of pure common interest**: any increase in one agent's payoff always corresponds with an increase in payoff of others.

- **Games with mixed motives**: agents' interests are, to varying extents, sometimes aligned and sometimes in conflict.

- **Games of pure conflicting interests**: an increase in one agent's payoff is always associated to a decrease in the payoff of others.

Common and Conflicting Interests

- Opportunities for cooperation exist, at least in principle, in situations of pure common interest and mixed motives.
- Only with pure conflicting interest, cooperation is impossible.
- Games of pure conflicting interests are typically confined to two-player games.
  - Typically, the introduction of a third player offer the opportunity of cooperating at least to one dyad.

Credit: Jason May

Types of Cooperative Agents

- We can consider different types of agents:
  - Machines;
  - Humans;
  - Organisations (of humans and/or machines)
    - States
    - Firms
    - …
- Very complex (and interesting!) situations arise in mixed societies of machines and humans.

AI Alignment

- AI Alignment is a related field.
- AI Alignment concerns with the problem of aligning the machine agent so that its preferences are as the humans intend.
- In Cooperative AI, the preferences of agents are given.
- When alignment succeeds, the human-machine dyad possesses pure common interest, which is a sort of limited case of cooperation problems.
- In AI alignment, human (principal) has normative priority over the AI agent. This is not the case in cooperative AI.

AI Alignment

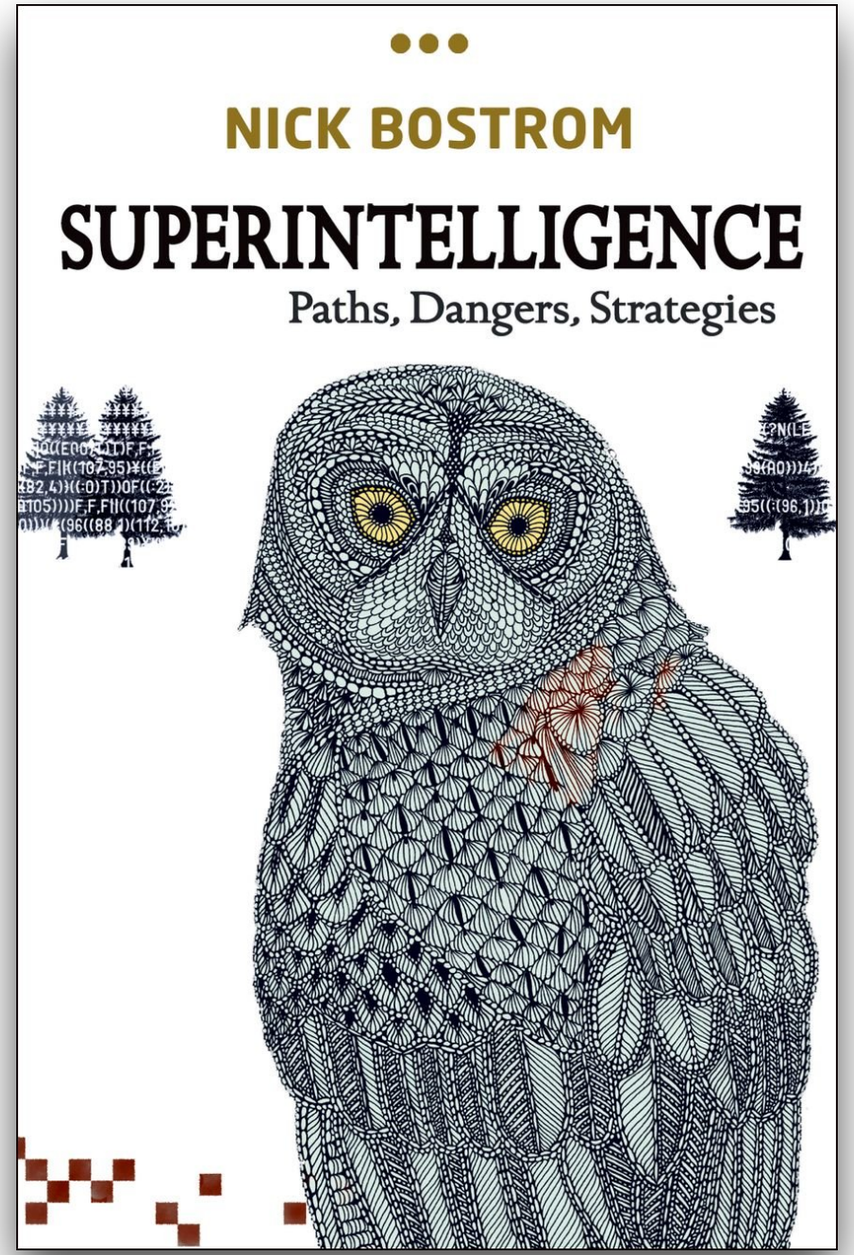
- AI Alignment is mostly about *control* (especially in case it is difficult to change the preferences of the AI agent).

- This is also related to the so-called singularity problems (and AI containment.)

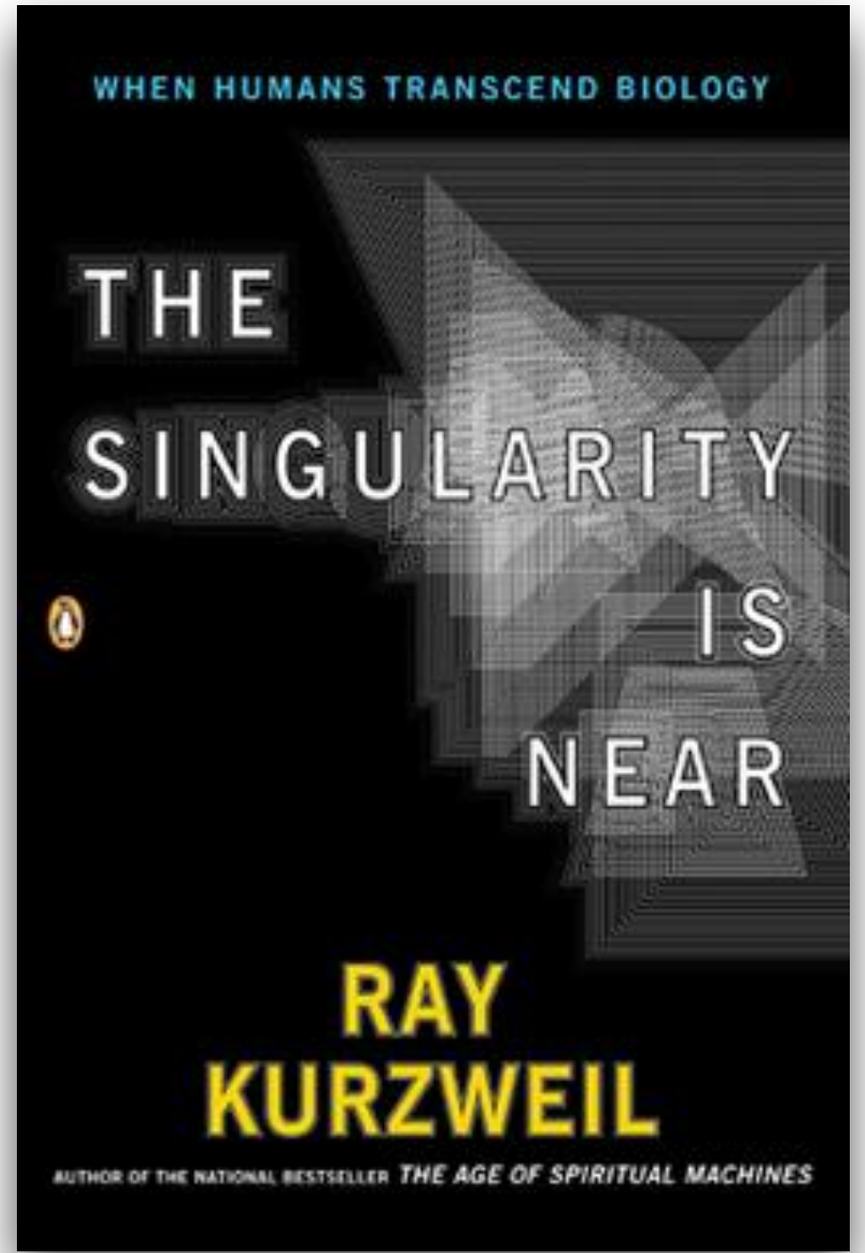
  - Let's say asymptotically… but there are variety of existing exciting problems right now.

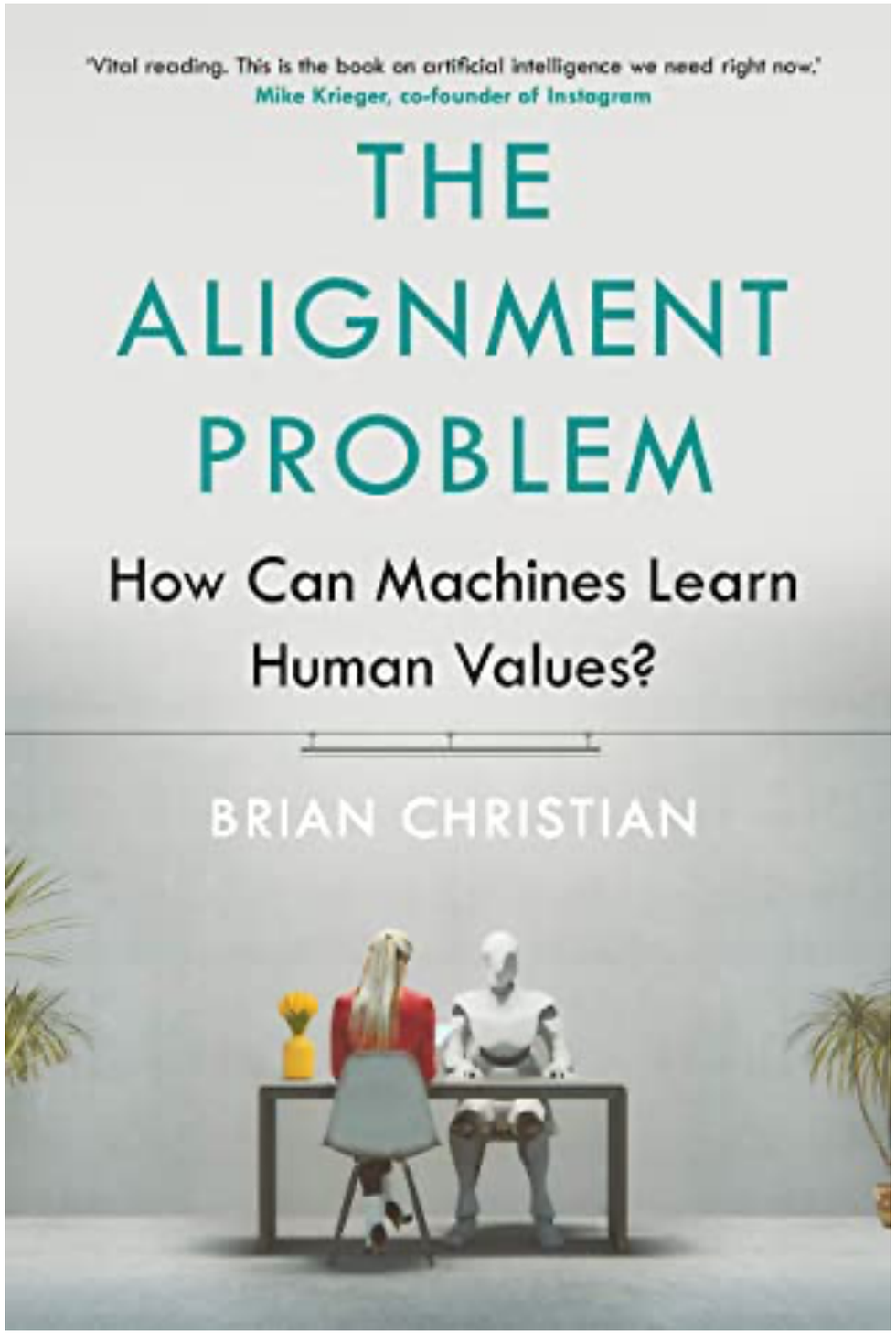
- Cooperative AI focusses on problems such as bargaining, credible communication and commitment.

Superintelligence



The Singularity is Near




'Vital reading. This is the book on artificial intelligence we need right now.'
Mike Krieger, co-founder of Instagram
THE ALIGNMENT PROBLEM
How Can Machines Learn Human Values?
BRIAN CHRISTIAN

Institutions

- *Institutions* are social structures, which involves a system of beliefs and norms, which define the “rules” of the games played by individuals and organisations composing a collective, shaping the actions that can be taken by individuals and the outcomes determined by these actions, resulting in stable patterns of behaviour.

- *Cooperative institutions* are those that support cooperative dynamics:
  - Mainly by resolving coordination problems and aligning incentives to solve social dilemmas.
    - Social dilemmas are situations that create a conflict between the individual’s interests and the collective’s interests.

Conventions and Norms

- In *games of common interest*, conventions are patterns of expectations and behaviour that promote coordination.

- In *mixed-motives games*, norms are used to reinforce patterns of expectations and behaviour through social reward and sanctions.

  - Norms are broadly understood to be informal rules that guide the behaviour of a group or a society.

Decentralised Institutions

- *Institutions* are formal entities that allocate roles, responsibilities, power and resources in a way designed to reproduce patterns of desired interactions.

- Institutions can be:
  - Emergent vs designed;
  - Informal vs formal;
  - Decentralised vs centralised.

Decentralised Institutions

- In *decentralised institutions*, there is no single central trusted authority, which can enforce decisions on behalf of a group of agents.
- Institutional structures can emerge from the interactions of agents over time.
  - Agents might act in a way that incentivise the others to behave in a certain way (for example, through *social punishment*).
- One of the prominent mechanisms at the basis of decentralised institutions are norms.

Norms

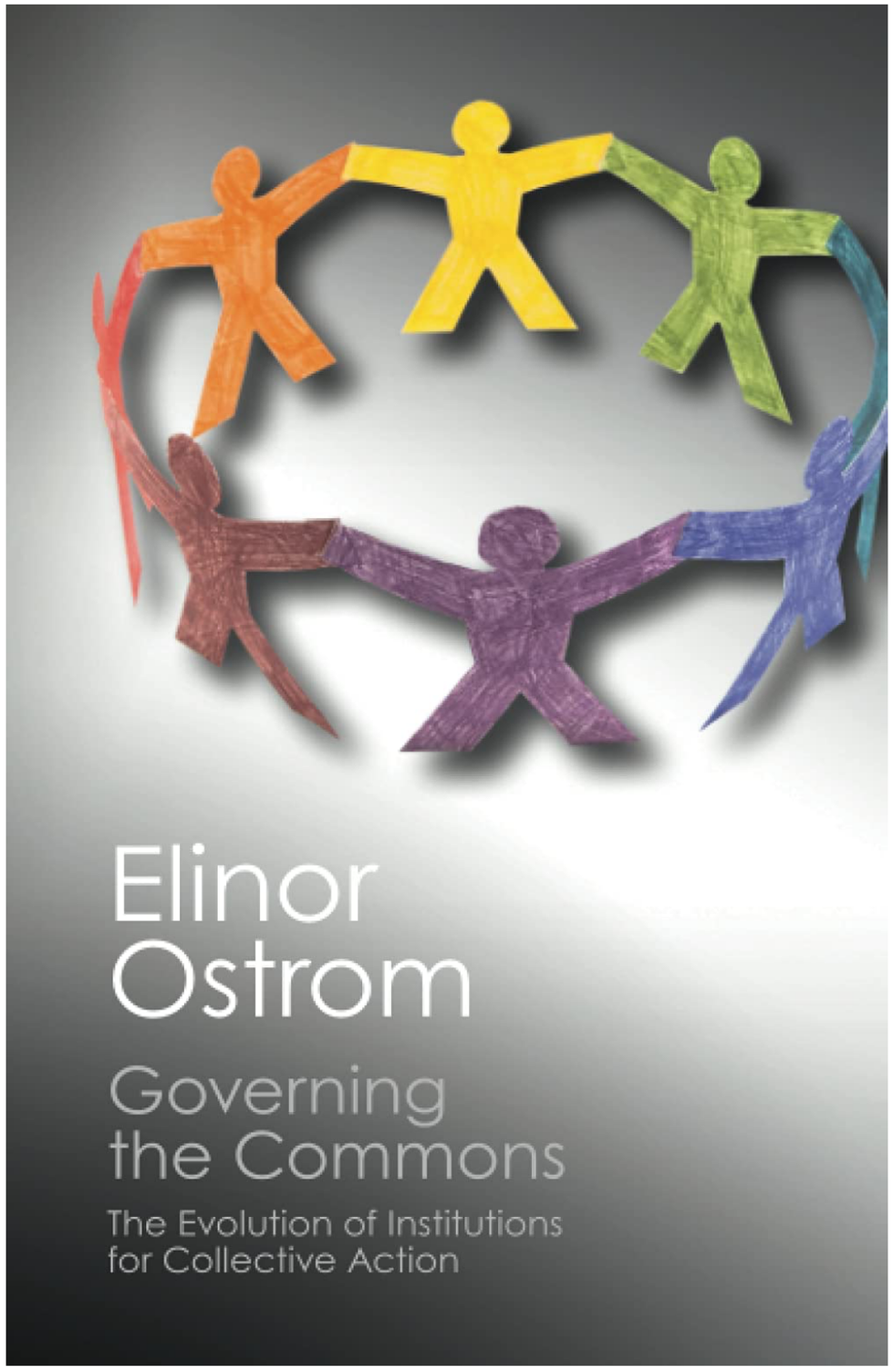
- Norms constrain the behaviour of group members, by capturing and encoding sanctions for transgressions.

- They are seen central to ensure global coordination.

- Humans are very good in organising themselves around social norms in a fully decentralised way.

  - Social norms key for civilisation.

- A key problem in AI is to design agents that “understand” norms and follow them.


Elinor
Ostrom
Governing
the Commons
The Evolution of Institutions
for Collective Action

Cooperation and Reputation Dynamics with Reinforcement Learning

Nicolas Anastassacos
University College London
The Alan Turing Institute
nanastas@cs.ucl.ac.uk

Julian García
Monash University
julian.garcia@monash.edu

Stephen Hailes
University College London
s.hailes@ucl.ac.uk

Mirco Musolesi
University College London
The Alan Turing Institute
University of Bologna
m.musolesi@ucl.ac.uk

ABSTRACT

Creating incentives for cooperation is a challenge in natural and artificial systems. One potential answer is reputation, whereby agents trade the immediate cost of cooperation for the future benefits of having a good reputation. Game theoretical models have shown that specific social norms can make cooperation stable, but how agents can independently learn to establish effective reputation mechanisms on their own is less understood. We use a simple model of reinforcement learning to show that reputation mechanisms generate two coordination problems: agents need to learn how to coordinate on the meaning of existing reputations and collectively agree on a social norm to assign reputations to others based on their behavior. These coordination problems exhibit multiple equilibria, some of which effectively establish cooperation. When we train

1 INTRODUCTION

Cooperation is important in natural and artificial systems [2]. It allows for agents with individual goals to reach beneficial group outcomes, even when group and individual incentives are not perfectly aligned [34]. If cooperation is costly but the benefits of cooperation can be enjoyed by all agents, the temptation to pay no cost is a dominant strategy and cooperation is hard to establish and maintain, unless a specific mechanism is in place to foster cooperation [21].

One popular mechanism is direct reciprocity, which allows for agents to meet repeatedly [10, 37], thereby creating incentives to punish past defections, making cooperation viable via reciprocal strategies like Tit-for-Tat [3]. When agents are anonymous or cannot interact repeatedly, they can use a reputation mechanism to condition their cooperative actions, e.g., cooperating only with

The Role of AI in Designing Decentralised Institutions

- The AI research community can help in designing effective institutions that lead to desirable outcomes.

- Applications:
  - Planning institutions;
  - Setting incentives;
  - International relation analysis;
  - Human-AI interactions;
  - …

Main Reference

- The main reference is: Dafoe et al. Open Problems in Cooperative AI. arXiv preprint arXiv:2012.08630.

- Acknowledgement: Some definitions and concepts in this lecture were taken from that paper.