Autonomous and Adaptive Systems

Intelligent Agents

Mirco Musolesi

mircomusolesi@acm.org

VOL. LIX. NO. 236.] [October, 1950

MIND
A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

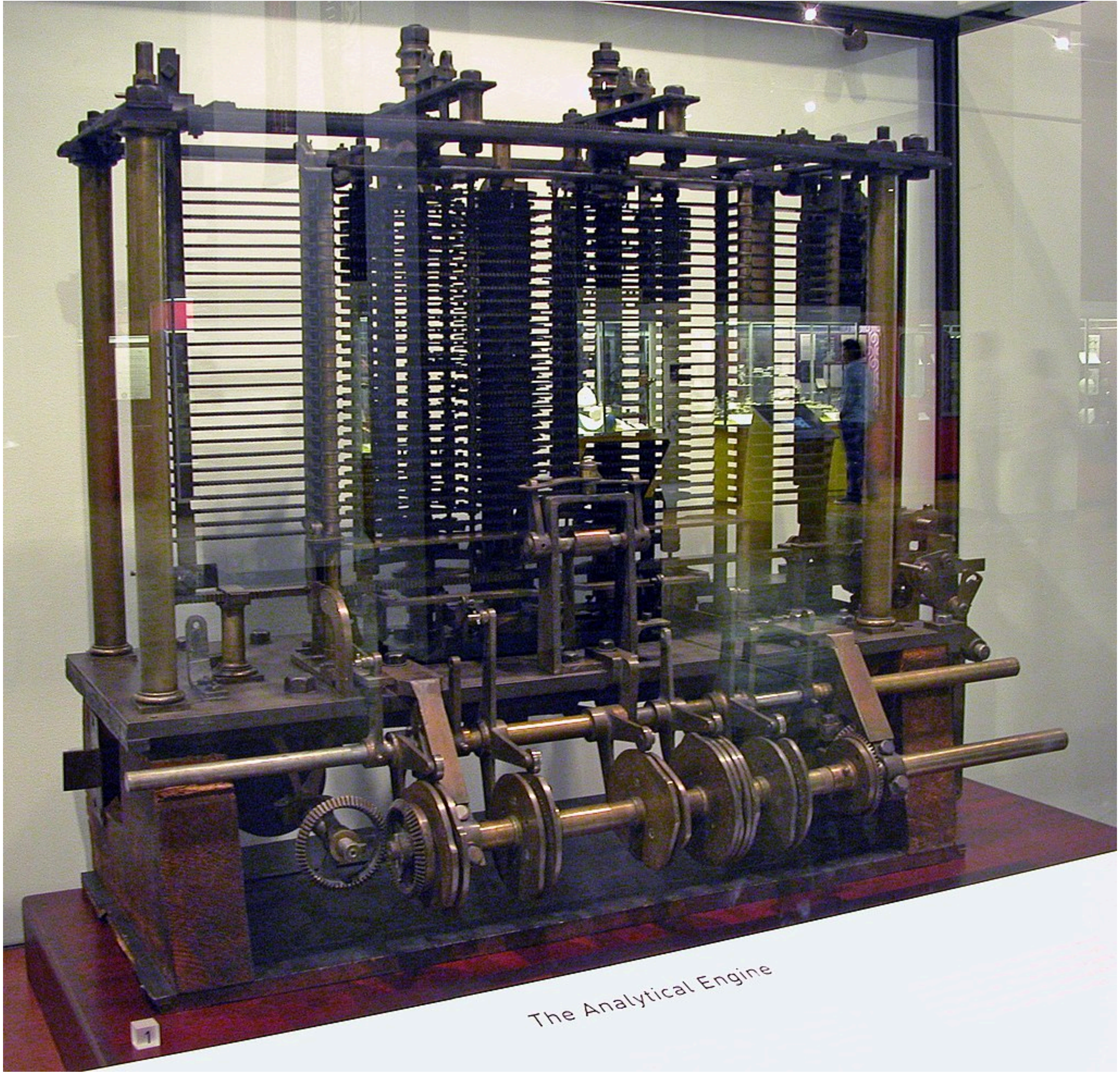
course only a finite part can have been used at any one time. Likewise only a finite amount can have been constructed, but we can imagine more and more being added as required. Such computers have special theoretical interest and will be called infinitive capacity computers.

The idea of a digital computer is an old one. Charles Babbage, Lucasian Professor of Mathematics at Cambridge from 1828 to 1839, planned such a machine, called the Analytical Engine, but it was never completed. Although Babbage had all the essential ideas, his machine was not at that time such a very attractive prospect. The speed which would have been available would be definitely faster than a human computer but something like 100 times slower than the Manchester machine, itself one of the slower of the modem machines. The storage was to be purely mechanical, using wheels and cards.

The fact that Babbage's Analytical Engine was to be entirely mechanical will help us to rid ourselves of a superstition. Importance is often attached to the fact that modem digital computers are electrical, and that the nervous system also is electrical. Since Babbage's machine was not electrical, and since all digital computers are in a sense equivalent, we see that this use of electricity cannot be of theoretical importance. Of course electricity usually comes in where fast signalling is concerned, so that it is not surprising that we find it in both these connections. In the nervous system chemical phenomena are at least as important as electrical. In certain computers the storage system is mainly acoustic. The feature of using electricity is thus seen to be only a very superficial similarity. If we wish to find such similarities we should look rather for mathematical analogies of function.

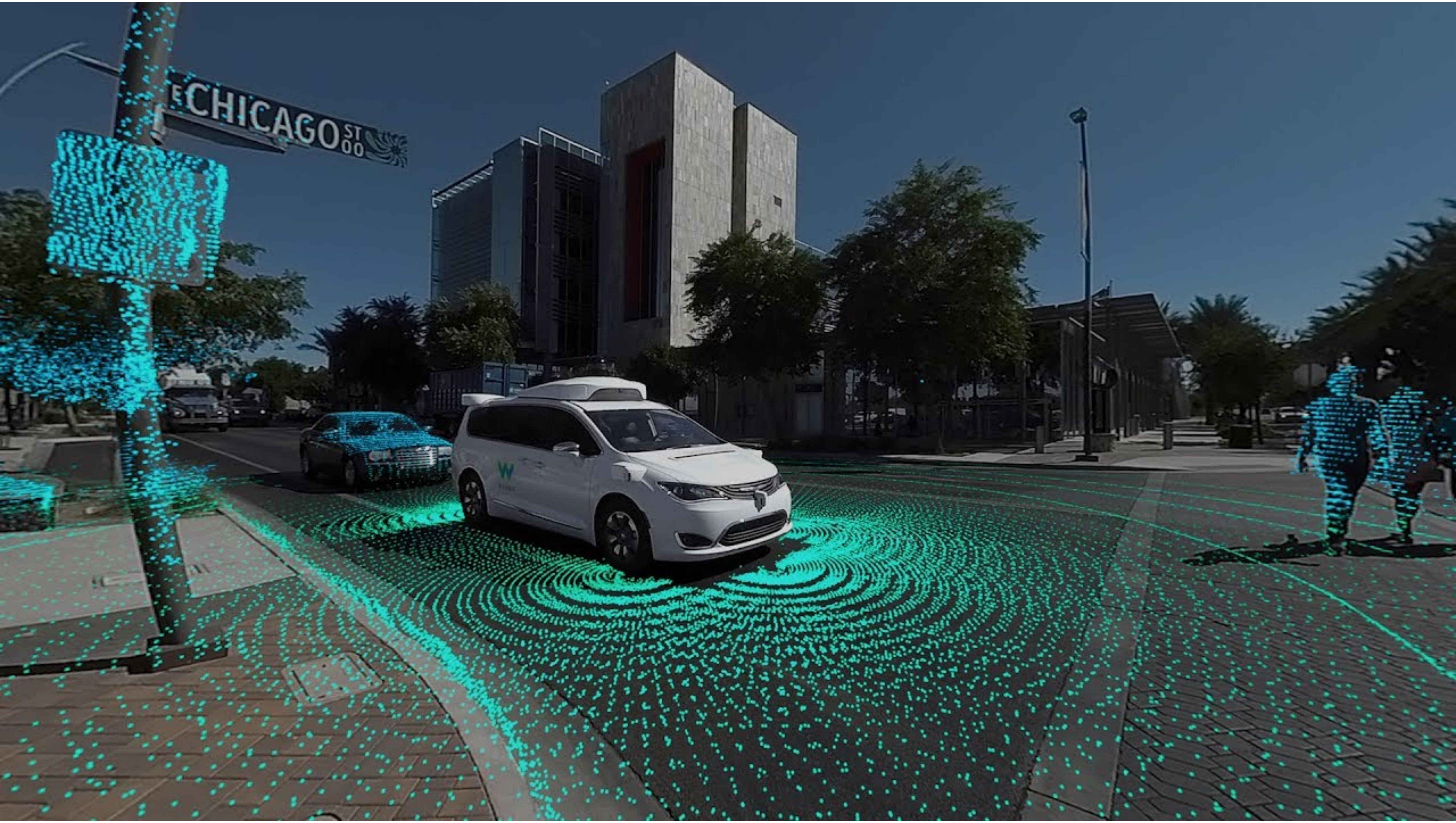
5. *Universality of Digital Computers.*

The digital computers considered in the last section may be classified amongst the 'discrete state machines'. These are the machines which move by sudden jumps or clicks from one quite definite state to another. These states are sufficiently different for the possibility of confusion between them to be ignored. Strictly speaking there are no such machines. Everything really moves continuously. But there are many kinds of machine which can profitably be *thought of* as being discrete state machines. For instance in considering the switches for a lighting system it is a convenient fiction that each switch must be definitely on or definitely off. There must be intermediate positions, but for most purposes we can forget about them. As an example of a discrete state machine we might

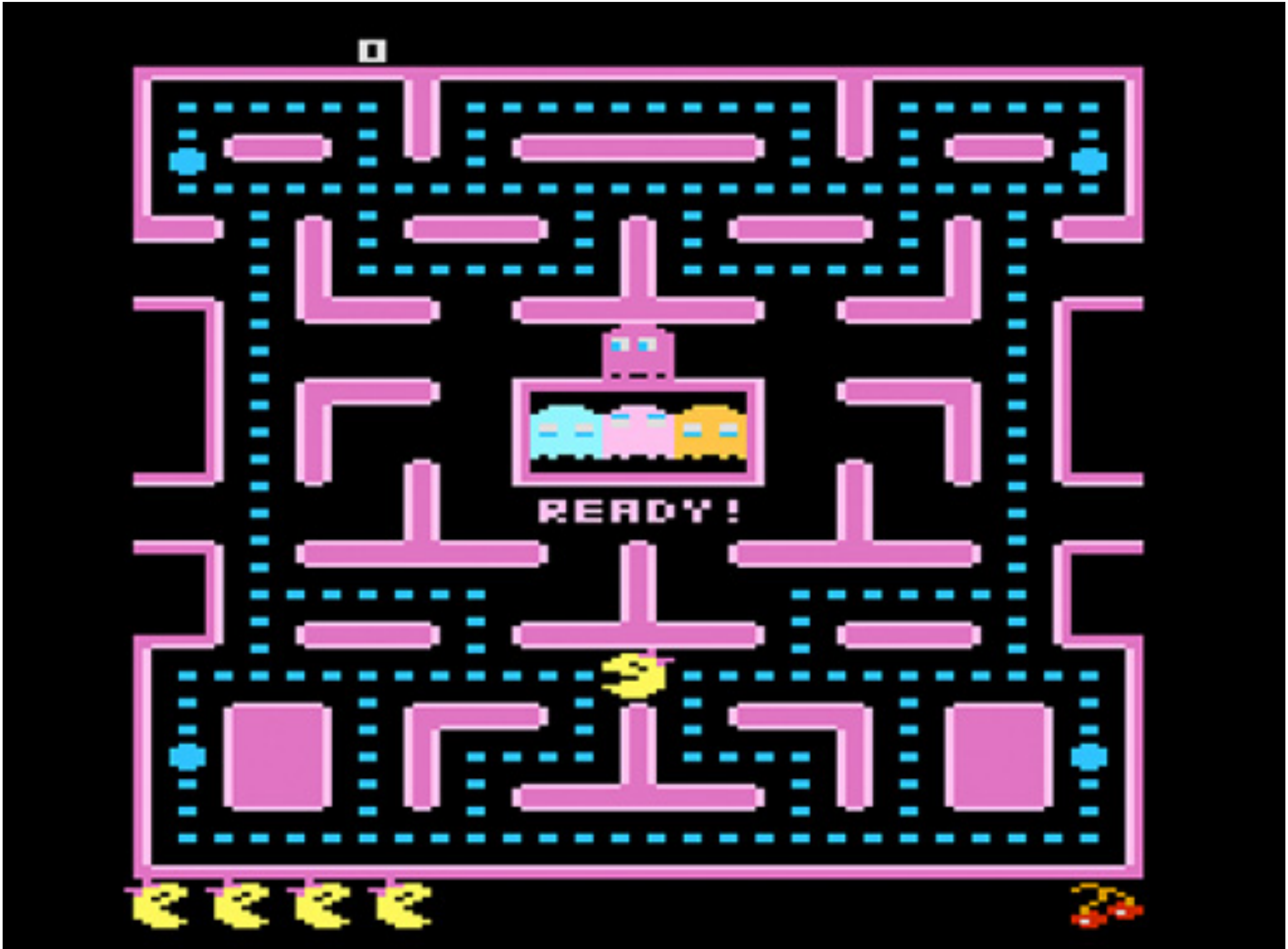


Credit: Wikipedia


WAYMO



https://www.youtube.com/watch?v=B8R148hFxPw


0
READY!

Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih   Koray Kavukcuoglu   David Silver   Alex Graves   Ioannis Antonoglou

Daan Wierstra   Martin Riedmiller

DeepMind Technologies

{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

Abstract

We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future

ARTICLE

doi:10.1038/nature16961

Mastering the game of Go with deep neural networks and tree search

David Silver[1]*, Aja Huang[1]*, Chris J. Maddison[1], Arthur Guez[1], Laurent Sifre[1], George van den Driessche[1], Julian Schrittwieser[1], Ioannis Antonoglou[1], Veda Panneershelvam[1], Marc Lanctot[1], Sander Dieleman[1], Dominik Grewe[1], John Nham[2], Nal Kalchbrenner[1], Ilya Sutskever[2], Timothy Lillicrap[1], Madeleine Leach[1], Koray Kavukcuoglu[1], Thore Graepel[1] & Demis Hassabis[1]

The game of Go has long been viewed as the most challenging of classic games for artificial intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves. Here we introduce a new approach to computer Go that uses 'value networks' to evaluate board positions and 'policy networks' to select moves. These deep neural networks are trained by a novel combination of supervised learning from human expert games, and reinforcement learning from games of self-play. Without any lookahead search, the neural networks play Go at the level of state-

ARTICLE

doi:10.1038/nature24270

Mastering the game of Go without human knowledge

David Silver[1]*, Julian Schrittwieser[1]*, Karen Simonyan[1]*, Ioannis Antonoglou[1], Aja Huang[1], Arthur Guez[1], Thomas Hubert[1], Lucas Baker[1], Matthew Lai[1], Adrian Bolton[1], Yutian Chen[1], Timothy Lillicrap[1], Fan Hui[1], Laurent Sifre[1], George van den Driessche[1], Thore Graepel[1] & Demis Hassabis[1]

A long-standing goal of artificial intelligence is an algorithm that learns, *tabula rasa*, superhuman proficiency in challenging domains. Recently, AlphaGo became the first program to defeat a world champion in the game of Go. The tree search in AlphaGo evaluated positions and selected moves using deep neural networks. These neural networks were trained by supervised learning from human expert moves, and by reinforcement learning from self-play. Here we introduce an algorithm based solely on reinforcement learning, without human data, guidance or domain knowledge beyond game rules. AlphaGo becomes its own teacher: a neural network is trained to predict AlphaGo's own move selections and also the winner of AlphaGo's games. This neural network improves the strength of the tree search, resulting in higher quality

https://www.youtube.com/watch?v=6EQAsrfUIyo

COMPUTER SCIENCE

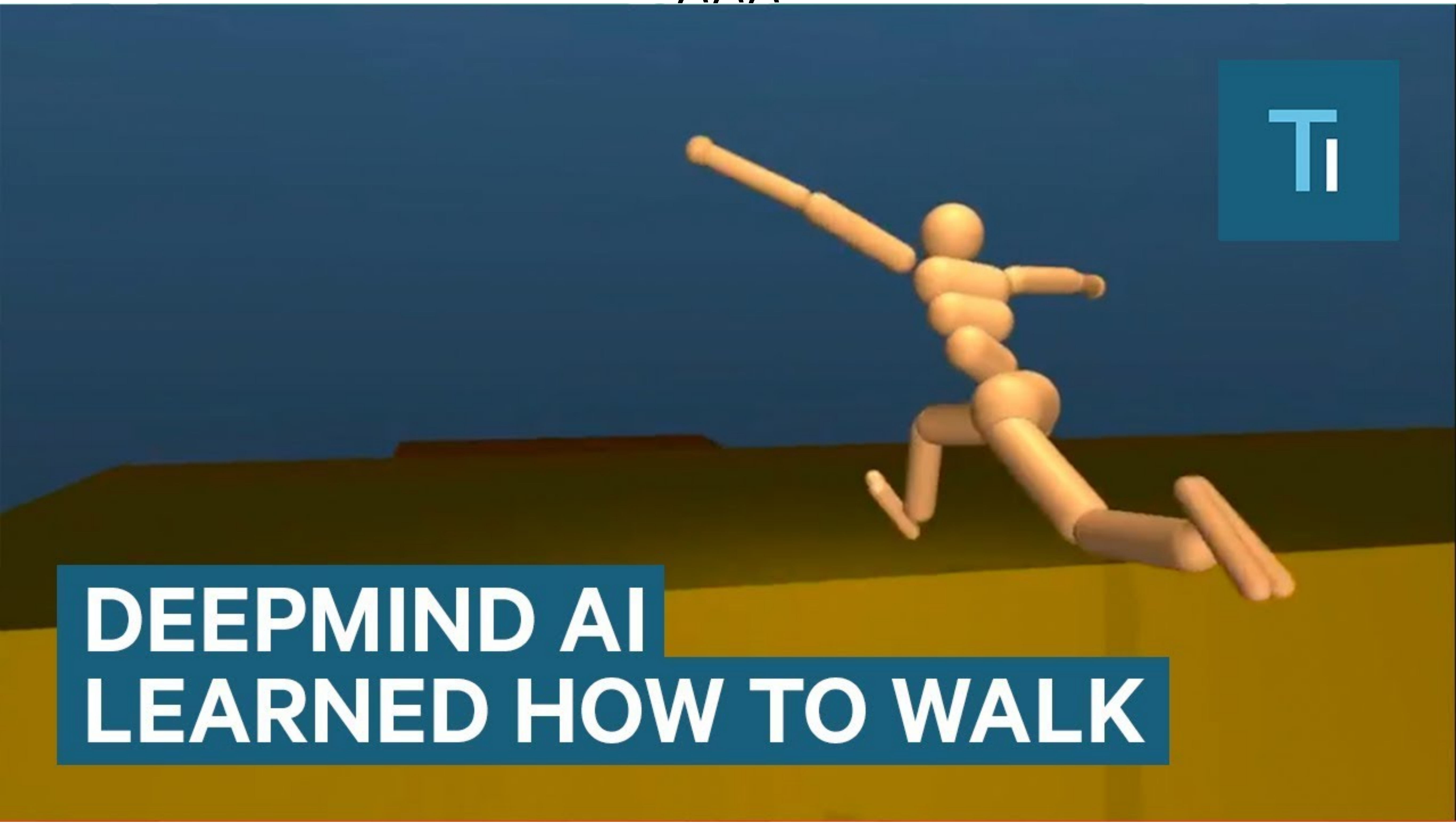
DeepStack: Expert-level artificial intelligence in heads-up no-limit poker

Matej Moravčík,[1,2*] Martin Schmid,[1,2*] Neil Burch,[1] Viliam Lisý,[1,3] Dustin Morrill,[1] Nolan Bard,[1] Trevor Davis,[1] Kevin Waugh,[1] Michael Johanson,[1] Michael Bowling[1]†

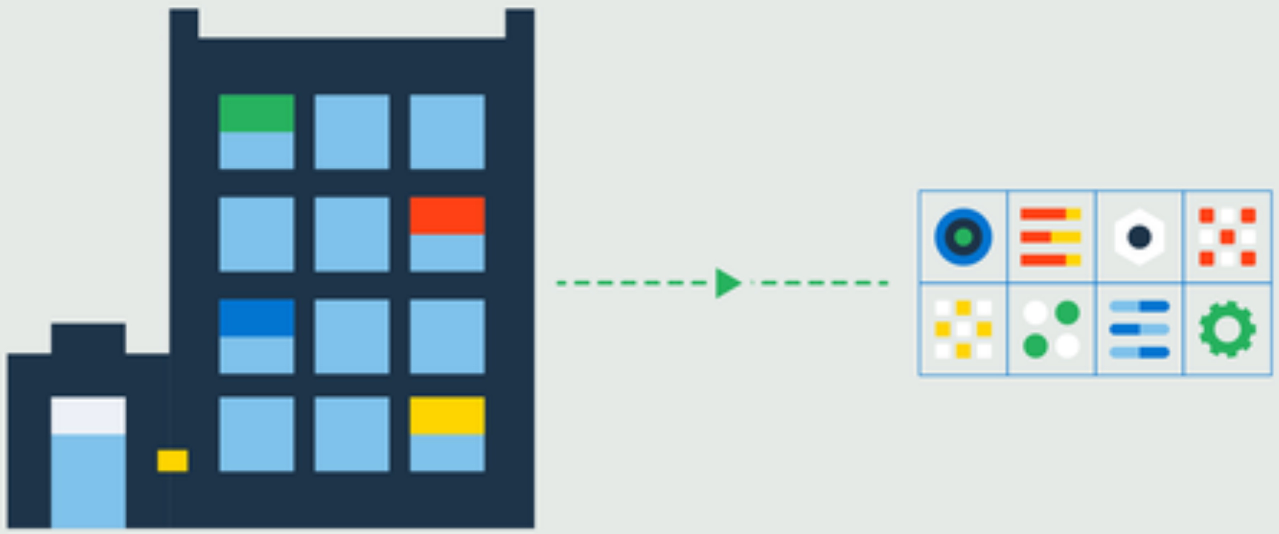
Artificial intelligence has seen several breakthroughs in recent years, with games often serving as milestones. A common feature of these games is that players have perfect information. Poker, the quintessential game of imperfect information, is a long-standing challenge problem in artificial intelligence. We introduce DeepStack, an algorithm for imperfect-information settings. It combines recursive reasoning to handle information asymmetry, decomposition to focus computation on the relevant decision, and a form of intuition that is automatically learned from self-play using deep learning. In a study involving 44,000 hands of poker, DeepStack defeated, with statistical significance, professional poker players in heads-up no-limit Texas hold'em. The approach is theoretically sound and is shown to produce strategies that are more difficult to exploit than prior approaches.

Likely as a result of this loss of information, such programs are behind expert human play. In 2015, the computer program Claudico lost to a team of professional poker players by a margin of 91 milli–big blinds per game (mbb/g) (*19*), which is a "huge margin of victory" (*20*). Furthermore, it was recently shown that abstraction-based programs from the Annual Computer Poker Competition have massive flaws (*21*). Four such programs (including top programs from the 2016 competition) were evaluated using a local best response technique that produces an approximate lower bound on how much a strategy can lose. All four abstraction-based programs are beatable by more than 3000 mbb/g, whereas simply folding each game results in 750 mbb/g.

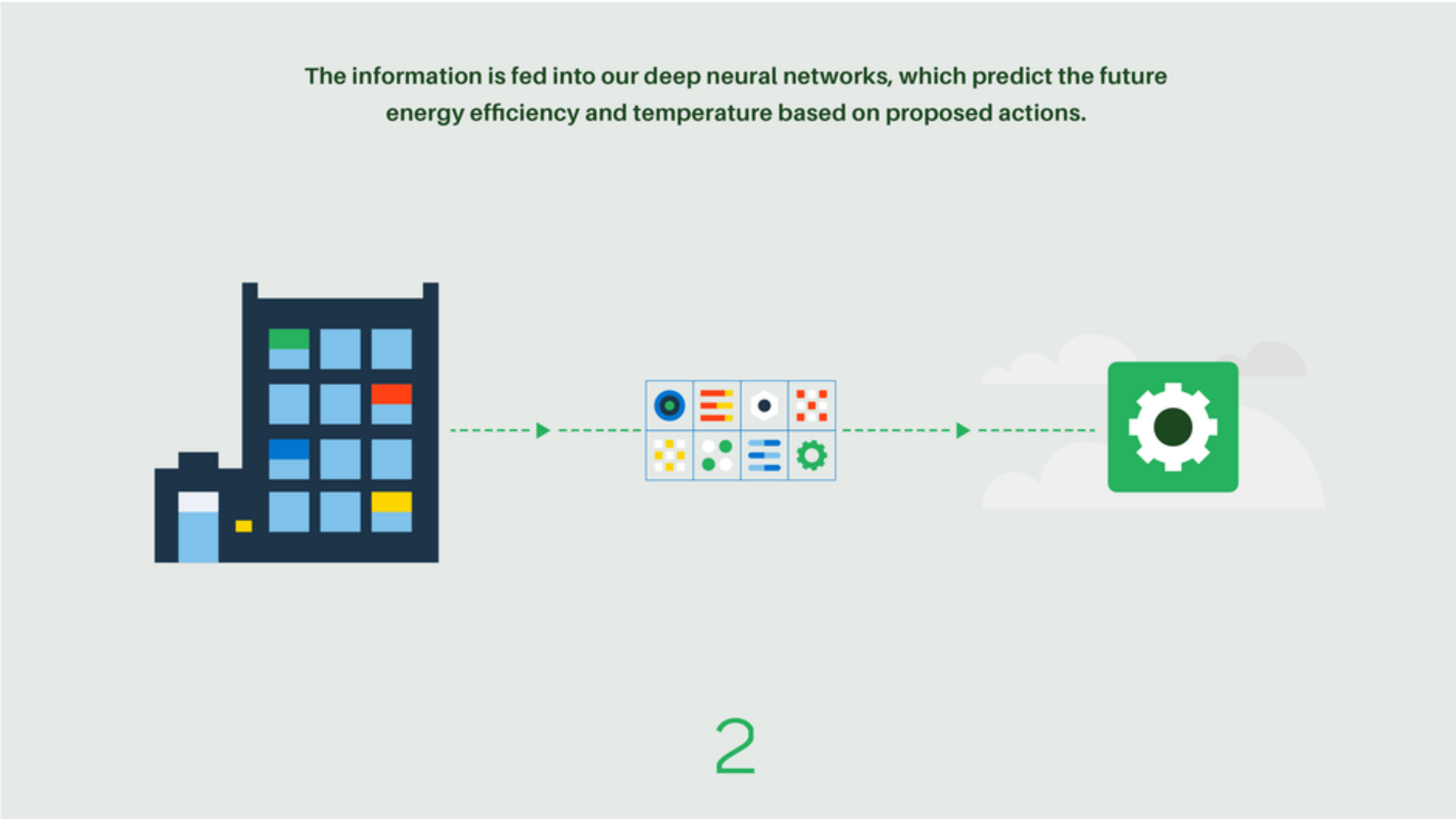
DeepStack takes a fundamentally different approach. It continues to use the recursive reasoning of CFR to handle information asymmetry. However, it does not compute and store a complete strategy prior to play and so has no need for explicit abstraction. Instead, it considers each particular situation as it arises during play, but not in isolation. It avoids reasoning about the


DEEPMIND AI
LEARNED HOW TO WALK

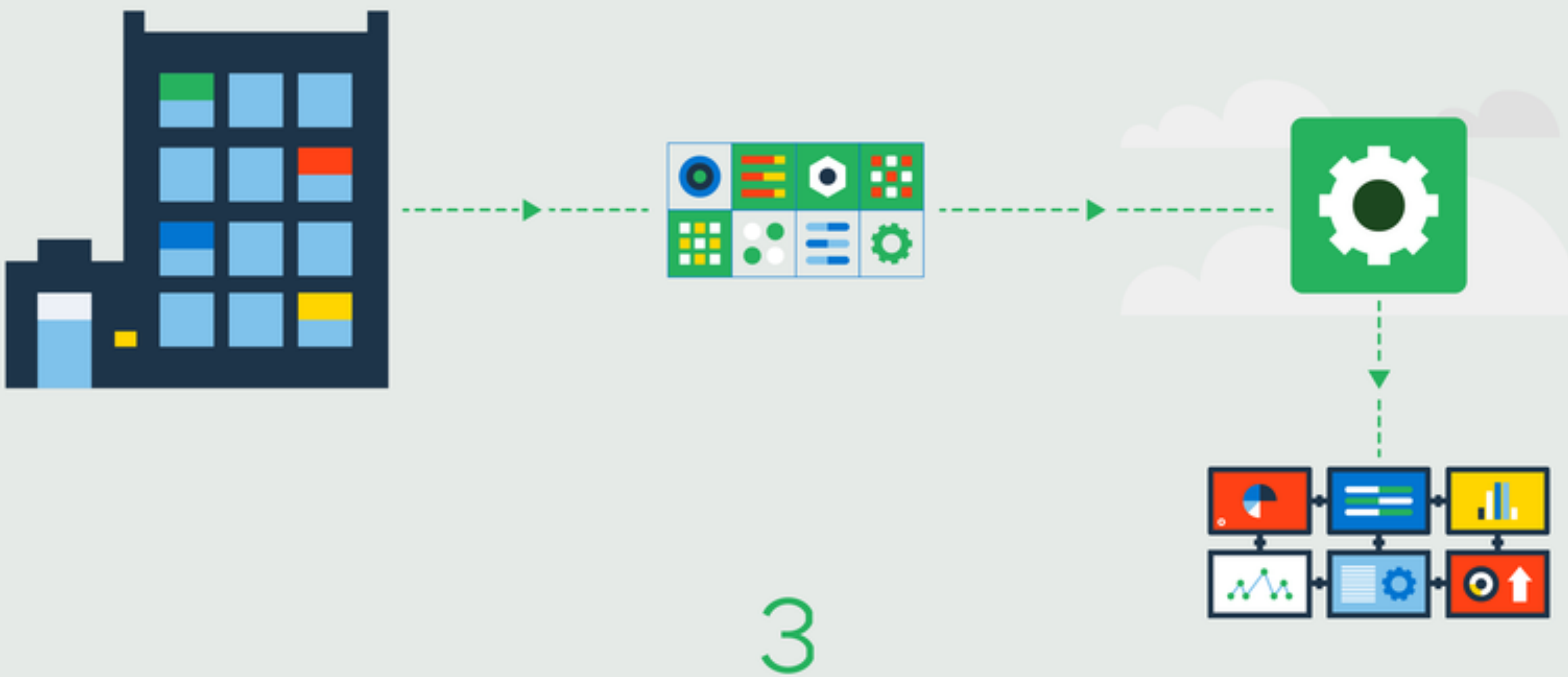
Every five minutes, our cloud-based AI pulls a snapshot of the data centre cooling system as represented by thousands of physical sensors.



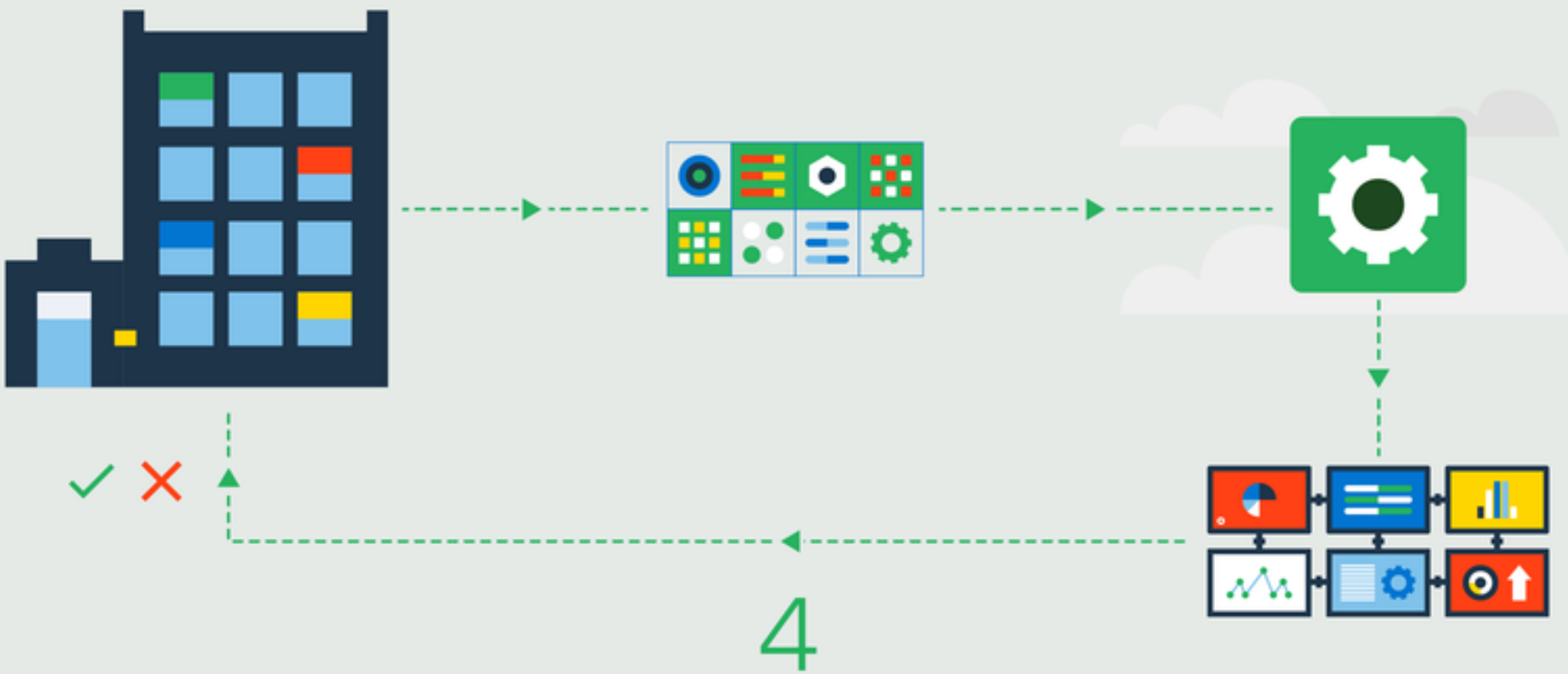
1


The information is fed into our deep neural networks, which predict the future energy efficiency and temperature based on proposed actions.
2

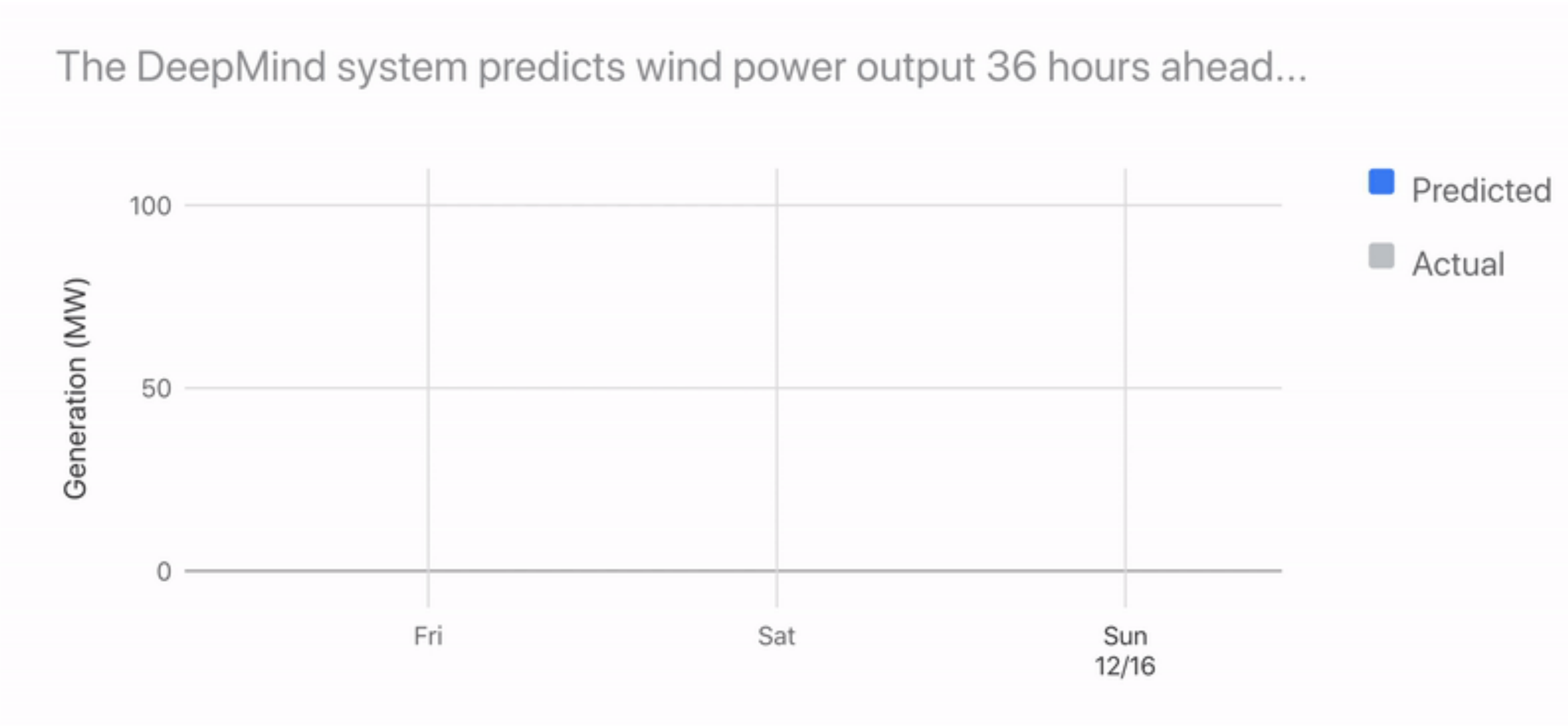
The AI selects actions that satisfy safety constraints and minimise future energy consumption.



Optimal actions are sent back to the data centre, where the local system verifies them against its own safety constraints before implementation.



Source: https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control



Source: https://deepmind.com/blog/article/machine-learning-can-boost-value-wind-energy

RL in Android

- RL used in Android for:
  - Adaptive battery:
    - It is used to learn and anticipate future battery use
  - Adaptive brightness of the video:
    - Algorithm learns preferences in terms of brightness from the user

The Brave New World

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

Microsoft Research

Abstract

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4 [Ope23], was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

Reinforcement Learning for Generative AI

ChatGPT

Gemini

ANTHROP\C

Talk to Claude

Deep Reinforcement Learning from Human Preferences

Paul F Christiano
OpenAI
paul@openai.com

Jan Leike
DeepMind
leike@google.com

Tom B Brown
Google Brain*
tombbrown@google.com

Miljan Martic
DeepMind
miljanm@google.com

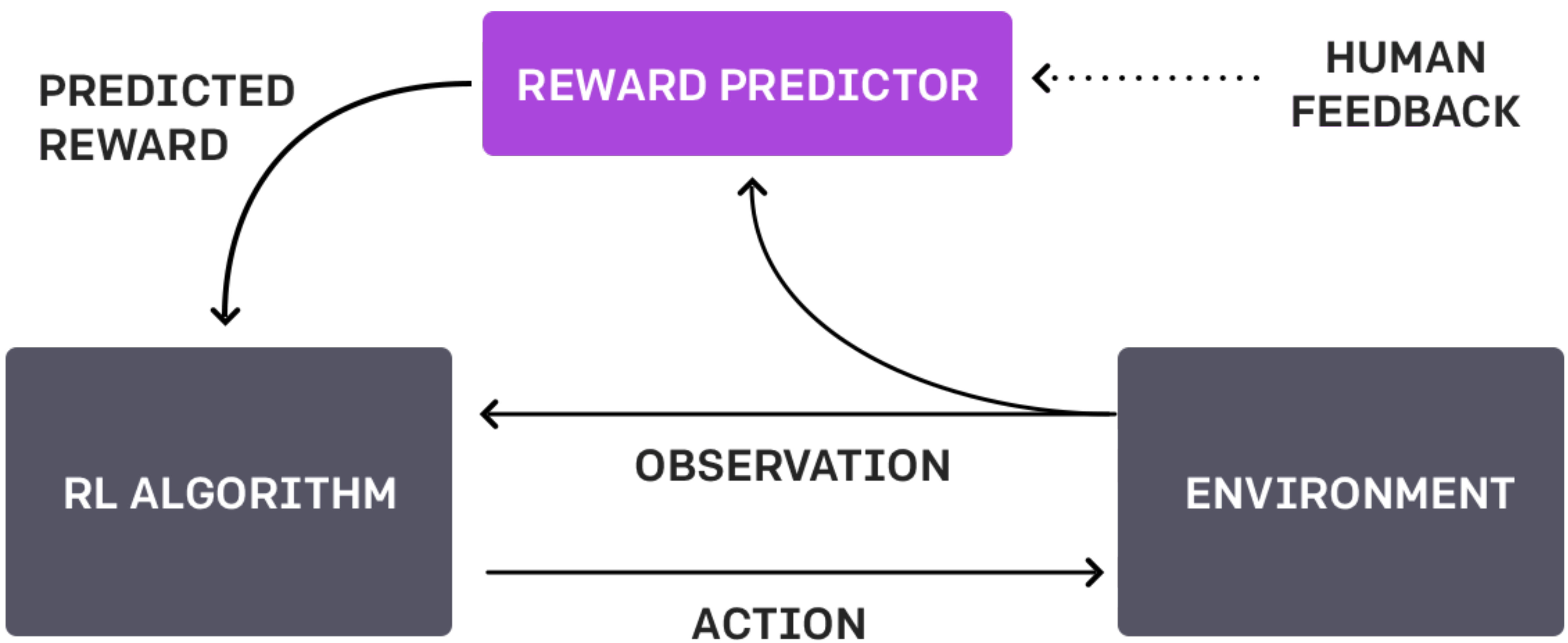
Shane Legg
DeepMind
legg@google.com

Dario Amodei
OpenAI
damodei@openai.com

Abstract

For sophisticated reinforcement learning (RL) systems to interact usefully with real-world environments, we need to communicate complex goals to these systems. In this work, we explore goals defined in terms of (non-expert) human preferences between pairs of trajectory segments. We show that this approach can effectively solve complex RL tasks without access to the reward function, including Atari games and simulated robot locomotion, while providing feedback on less than 1% of our agent's interactions with the environment. This reduces the cost of human oversight far enough that it can be practically applied to state-of-the-art RL systems. To demonstrate the flexibility of our approach, we show that we can successfully train complex novel behaviors with about an hour of human time. These behaviors and environments are considerably more complex than any which have been previously learned from human feedback.

Reinforcement Learning from Human Feedback (RLHF)



Source: OpenAI

Reinforcement Learning for Generative AI: State of the Art, Opportunities and Open Research Challenges

Giorgio Franceschelli GIORGIO.FRANCESCHELLI@UNIBO.IT
Department of Computer Science and Engineering, University of Bologna, Bologna, Italy

Mirco Musolesi M.MUSOLESI@UCL.AC.UK
Department of Computer Science, University College London, London, United Kingdom
Department of Computer Science and Engineering, University of Bologna, Bologna, Italy

Abstract

Generative Artificial Intelligence (AI) is one of the most exciting developments in Computer Science of the last decade. At the same time, Reinforcement Learning (RL) has emerged as a very successful paradigm for a variety of machine learning tasks. In this survey, we discuss the state of the art, opportunities and open research questions in applying RL to generative AI. In particular, we will discuss three types of applications, namely, RL as an alternative way for generation without specified objectives; as a way for generating outputs while concurrently maximizing an objective function; and, finally, as a way of embedding desired characteristics, which cannot be easily captured by means of an objective function, into the generative process. We conclude the survey with an in-depth discussion of the opportunities and challenges in this fascinating emerging area.

“Generative” Agents

Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park
Stanford University
Stanford, USA
joonspk@stanford.edu

Joseph C. O’Brien
Stanford University
Stanford, USA
jobrien3@stanford.edu

Carrie J. Cai
Google Research
Mountain View, CA, USA
cjcai@google.com

Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Percy Liang
Stanford University
Stanford, USA
pliang@cs.stanford.edu

Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu

Figure 1: Generative agents are believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents plan their days, share news, form relationships, and coordinate group activities.

ABSTRACT

Believable proxies of human behavior can empower interactive applications ranging from immersive environments to rehearsal spaces for interpersonal communication to prototyping tools. In

authors write; they form opinions, notice each other, and initiate conversations; they remember and reflect on days past as they plan the next day. To enable generative agents, we describe an architecture that extends a large language model to store a complete record


2025 Top 10 Strategic Technology Trends
AI imperatives and risks
Agentic AI
AI Governance Platforms
Disinformation Security
New frontiers of computing
Post-Quantum Cryptography
Ambient Invisible Intelligence
Energy-Efficient Computing
Hybrid Computing
Human-machine synergy
Spatial Computing
Polyfunctional Robots
Neurological Enhancement
Source: Gartner
© 2024 Gartner, Inc. and/or its affiliates.
All rights reserved. 3185862
Gartner

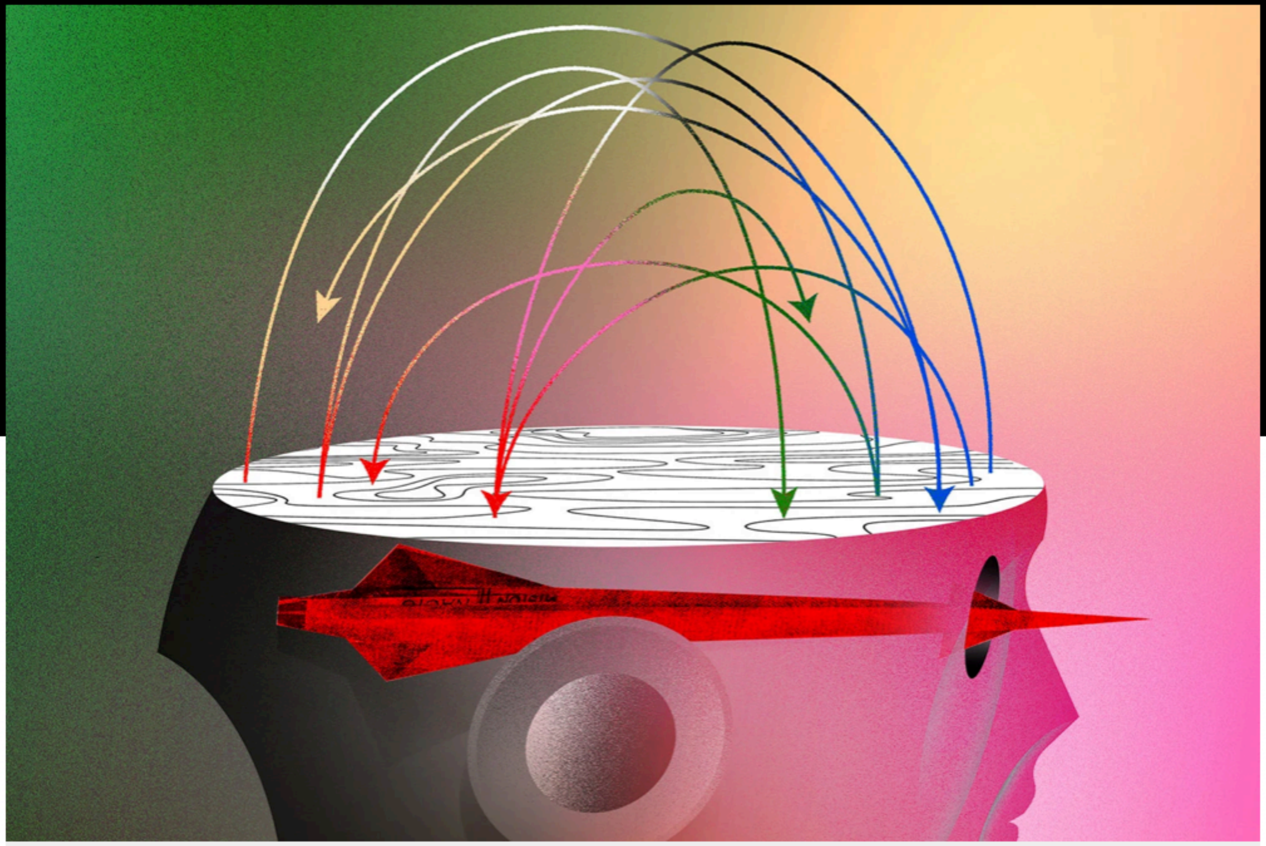
ANALYSIS

AI Has Entered the Situation Room

Data lets us see with unprecedented clarity—but reaping its benefits requires changing how foreign policy is made.

JUNE 19, 2023, 11:00 PM

By Stanley McChrystal, a retired four-star U.S. Army general and an advisor to Rhombus Power, and Anshu Roy, the founder and CEO of Rhombus Power.



BRIAN STAUFFER ILLUSTRATION FOR FOREIGN POLICY

Definition of Intelligent Agents

- An *intelligent agent* is an entity that perceives its environment and takes actions that maximise the probability of achieving its goals.

- Agents might be *physically situated* (we call them *robots*) or not (we refer to them as *software agents*, *softbots*, *bots*, etc.).

Definition of Adaptive Agents

- An adaptive agent is an entity that is capable of responding (by adapting) to its environment.
- The environment might also include other agents.

Definition of Autonomous Agents

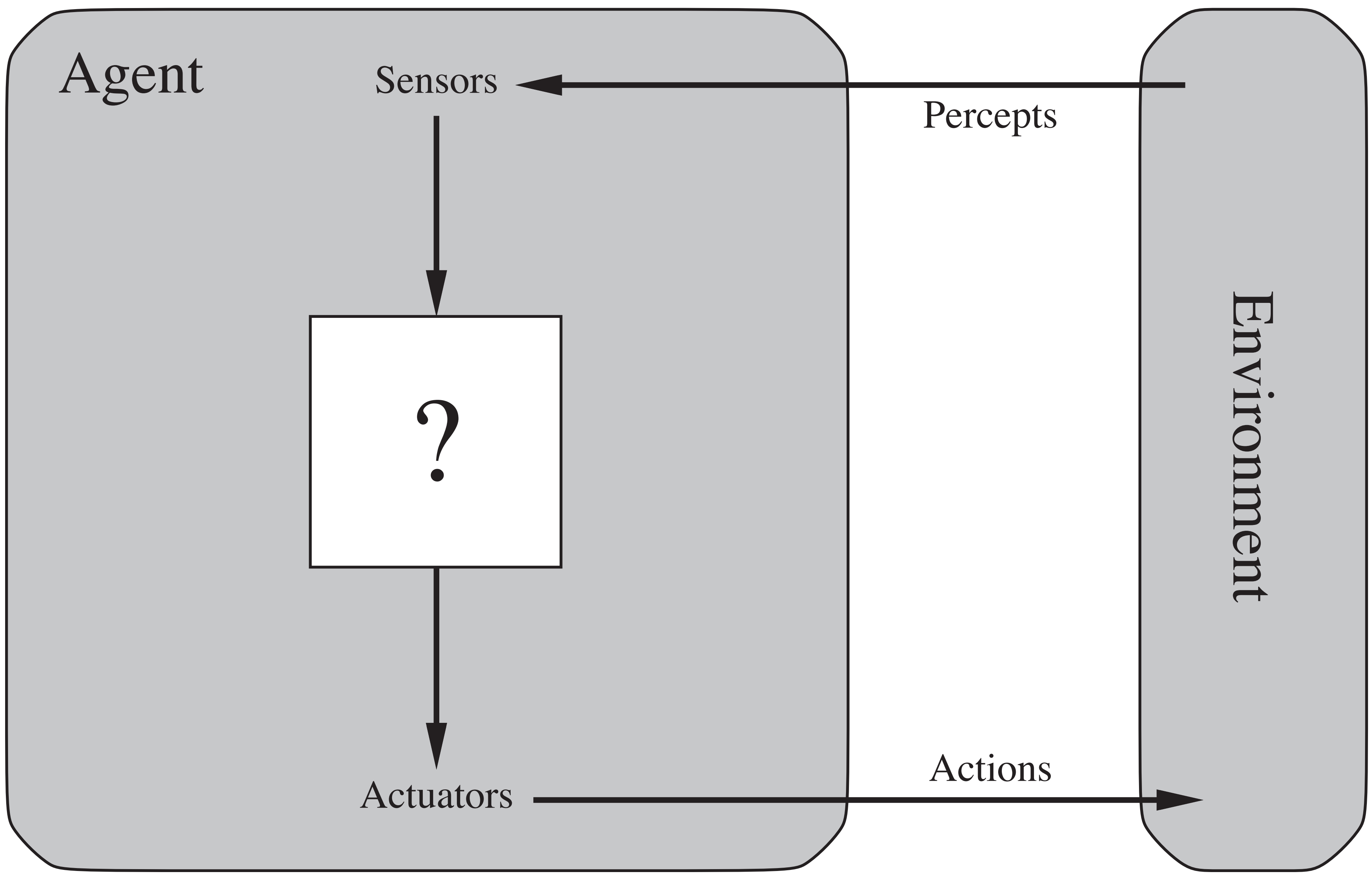
- An *autonomous agent* is an entity that relies only on its perception and acts in the world independently from its designer.
- An autonomous agent should be able to compensate for partial knowledge.
- From a practical point of view, it makes sense to provide the agent with *some knowledge of the world* and the *ability to learn*.
- As in the case of evolution, in which animals/humans have a sufficient number of built-in reflexes in order to survive long enough to learn by themselves, it is reasonable to provide an artificial intelligent agent with some initial knowledge as well as an ability to learn.

Learning and Autonomous Agents

- After sufficient experience of its environment, an intelligent agent can become effectively independent of its prior knowledge.
- For this reason, learning allows an intelligent agent to survive in a vast variety of environments.

Designing Agents

- Dimensions that should be taken into consideration in designing agents are the following:
  - Performance
  - Environment
  - Actuators
  - Sensors


Agent
Sensors
Percepts
?
Environment
Actuators
Actions

	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Health patient, minimise cost, lawsuit	Patient, hospital, staff	Display questions, tests, treatments, etc.	Keyboard entry, patients' answers
Satellite image analysis system	Correct image categorisation	Downlink from satellite	Display categorisation of scenes	Colour pixel arrays
Part-picking robot	Percentage parts in correct bins	Conveyor belt with parts, bins	Jointed arm and hand	Camera, joint angle sensors
Refinery controller	Maximise purity, yield, safety	Refinery, operators	Valves, pumps, heaters, displays	Temperature, pressure, sensors
Interactive language tutor	Maximise student's test score	Set of students	Displays exercises, suggestions	Keyboard entry
Automatic display of advertisements	Click rates/ purchase conversion	Websites, online retailers, users	Display advertisements	Automatic extraction of content, clicks

Characteristics of the Environments

- We have different dimensions:
  - Fully observable vs partially observable;
  - Deterministic vs stochastic;
  - Episodic vs sequential;
  - Static vs dynamic;
  - Discrete vs continuous;
  - Single agent vs multi-agent.

A Categorization of Intelligent Agents

- There are essentially four basic kinds of agent programs:
  - Simple reflex agents;
  - Model-based reflex agents;
  - Goal-based agents; and
  - Utility-based agents.
- The behaviour of these agents can be *hard-wired* or it can be acquired, improved and optimise through *learning*.

Simple Reflex Agents

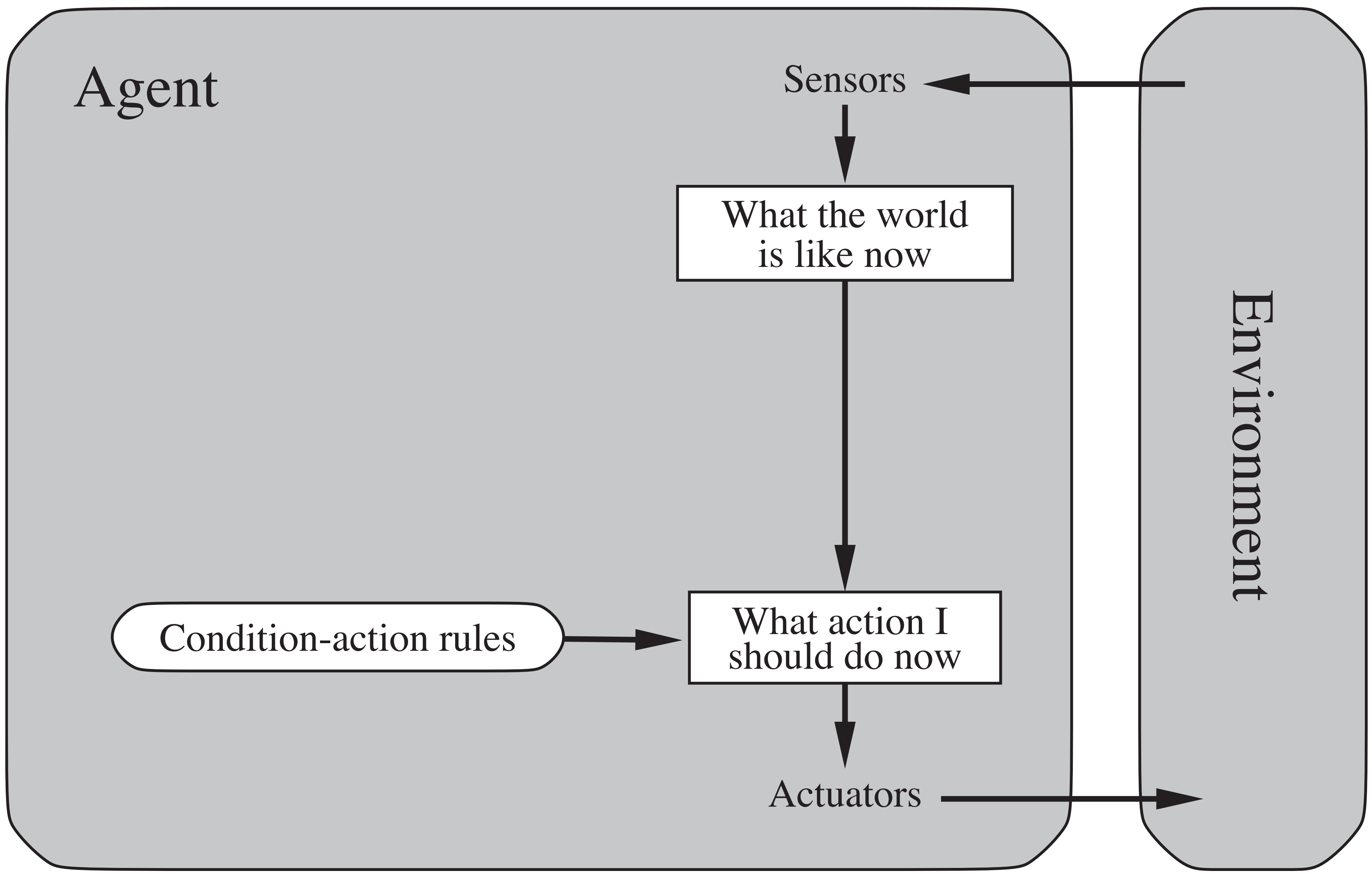
- *Simple reflex agents* select actions on the basis of the current percept, ignoring the rest of the percept history.

- These are the most basic form of agents, they are based on condition-action rules (also called simulation-action rules, productions or if-then rules).

- Humans also have a set of automatic (learned) responses.

- They are suitable for situations where the decisions can be made on the current observation.

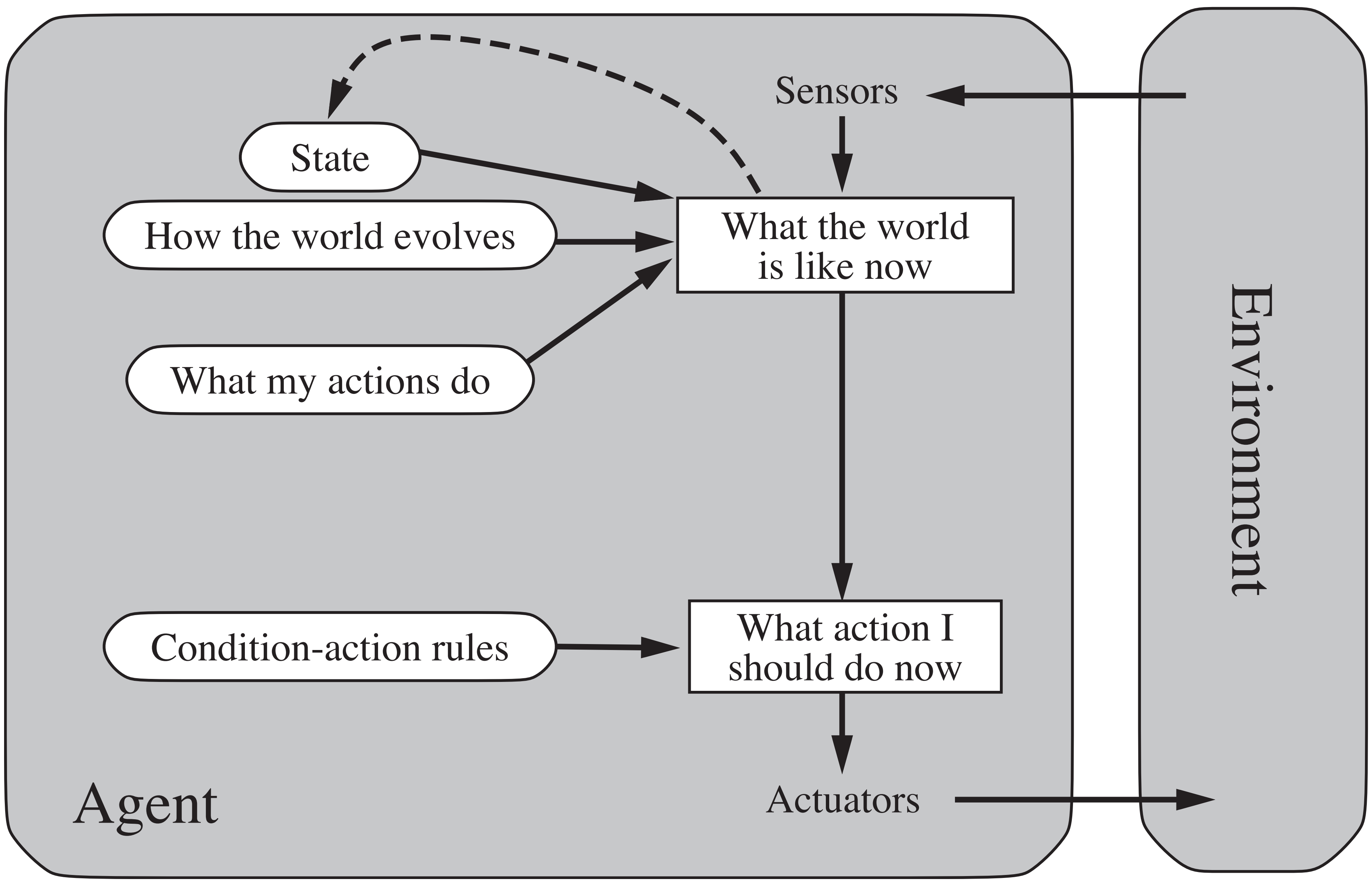
Simple-Reflex Agent



Model-based Agents

- A model-based agent is an agent that keeps an internal state and depends on two types of knowledge:
  - how the world evolves independently from the agent
  - how the actions of the agent affects the world.
- An internal state is essentially used to keep track of what is not possible to see right now. It depends on the percept history and, for this reason, it reflects at least some of the unobserved aspects of the current state.

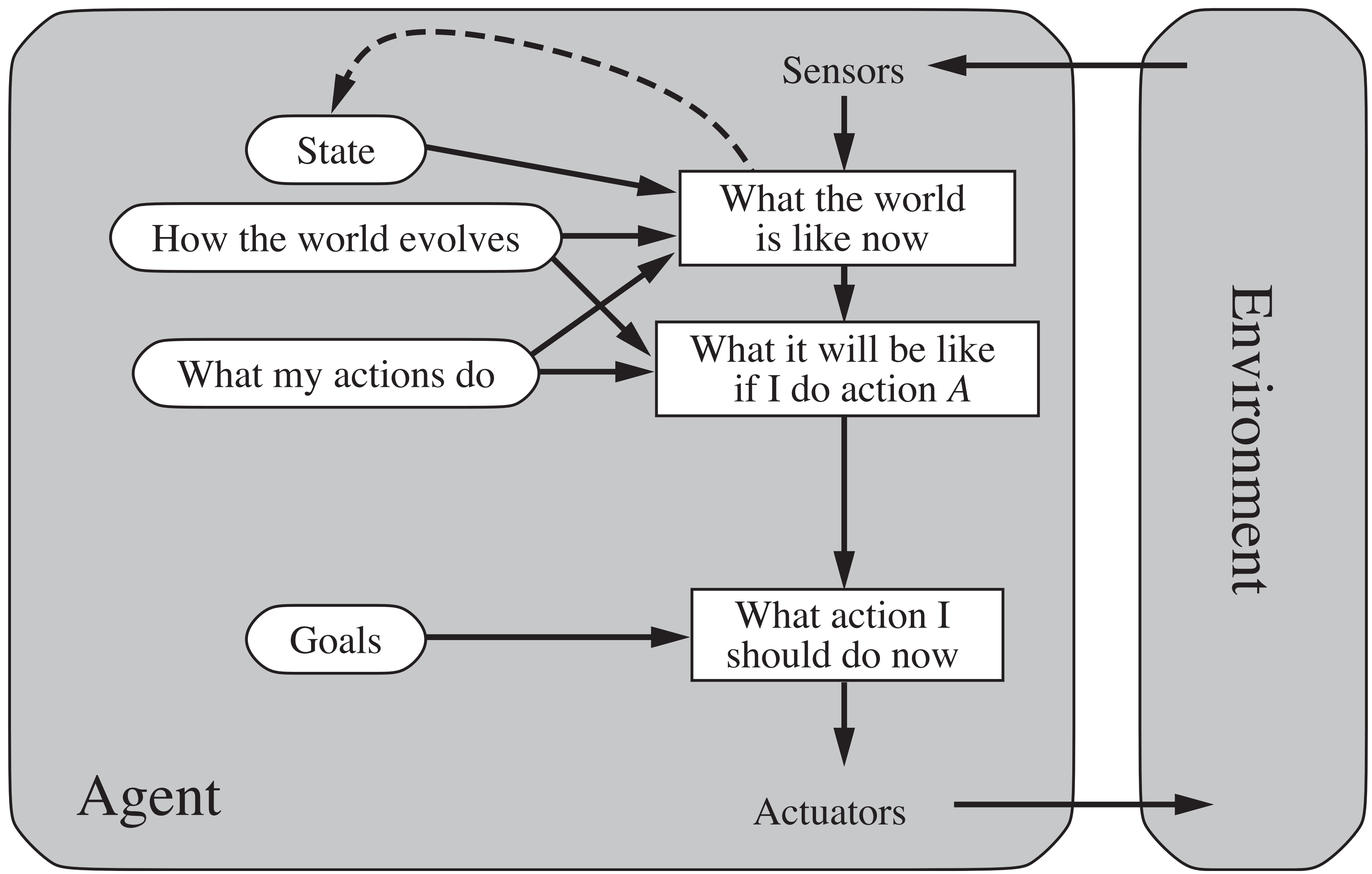
Model-based Agent



Goal-based Agents

- *Goal-based agents* act to achieve their goals.
- Sometimes goal-based satisfaction is straightforward, when goal satisfaction results immediately from a single action.
- In other cases, an agent has to consider a long sequence of actions to achieve their goal through *search* and *planning*.

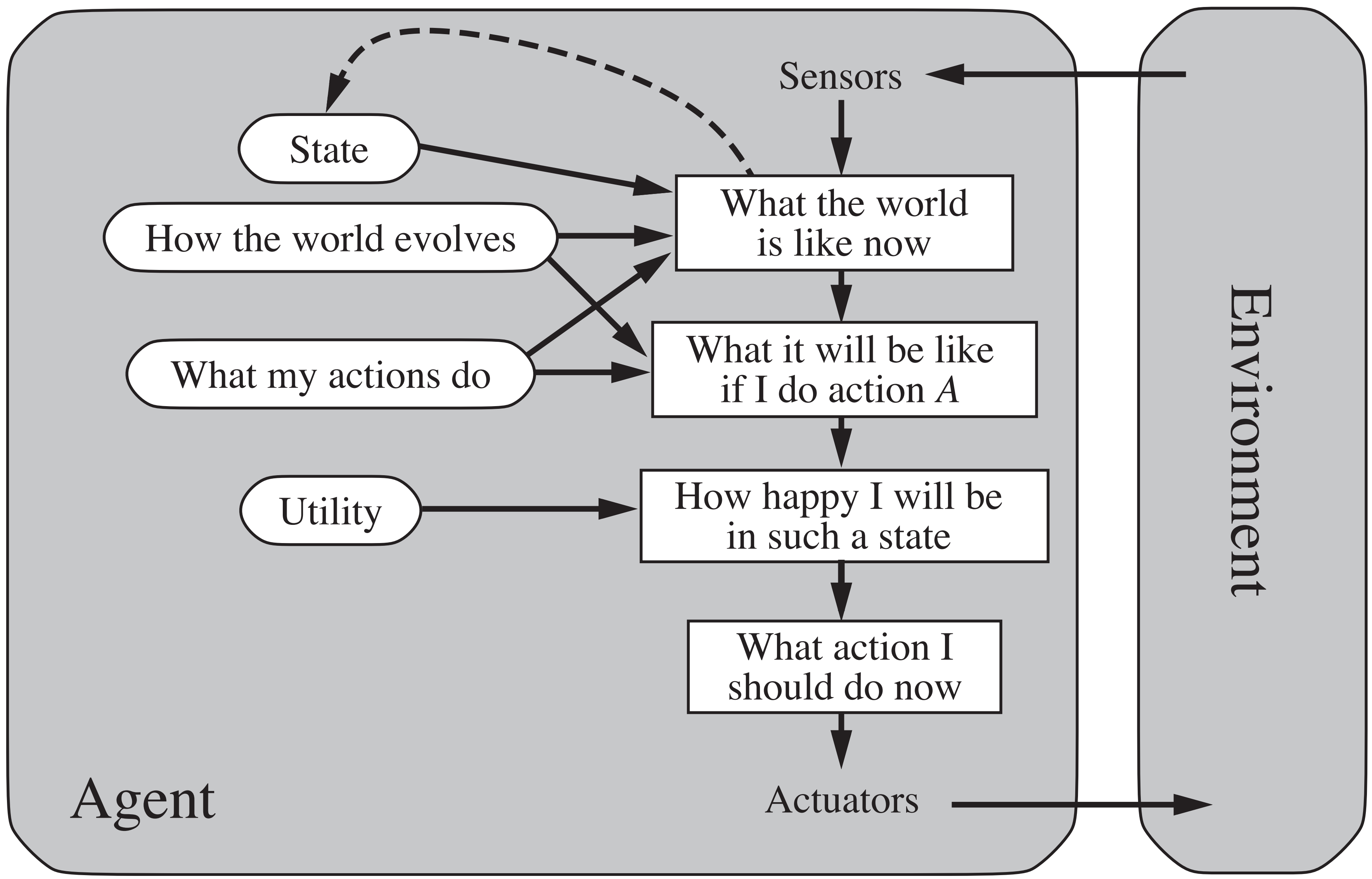
Model-based Goal-based Agent



Utility-based Agents

- Goals are not really sufficient to generate “high-quality behaviour” in most environments.
- There are states preferred to others, which we model using *utilities.*
- *Utility functions* maps a state (or a sequence of states) onto a real number, which describes the preferences of an agent.
- Example: taxi routing to destination (quicker route is preferable).

Model-based Utility-based Agent



Learning

- The behaviour of the agents can be pre-programmed and fixed or it can be *learned*.
- In order to allow for learning agents need a *learning component.*
- The learning component can be based on a model of the world and the gain towards a certain goal (possibly expressed in terms of change of the value of utility functions) can be expressed through rewards.
- In the module we will explore in depth a type of learning called Reinforcement Learning based on rewards following a certain action.

Autonomous and Adaptive Systems

- In this module, we will adopt a very broad view of intelligent, adaptive and autonomous agents.
- Following Herbert Simon's famous definition, we will consider more broadly "machines that think, that learn, and that create".
- Of course "thinking" is an ill-defined concept, which we will try to discuss during the module in detail.
- We will take a system-view, i.e., we will consider the design of intelligent and autonomous agents considering the algorithmic and implementation issues and the interactions between the different entities at various levels of abstraction.

Building Intelligent Machines

- An underlying theme of the module will be the definition of intelligence.
- Recall our definition of intelligent agents:
  - An *intelligent agent* is an entity that perceives its environment and takes actions that maximise the probability of achieving its goals.
- Our goal at the end will be to build intelligent machines, but we will also compare machine intelligence with human and animal intelligence.
- This is a fascinating aspect of the study of these systems: they help us in understanding what intelligence means in abstract and what makes us humans.

Artificial General Intelligence

- In the recent years we had several successes in building more and more sophisticated systems.

- Human-level Artificial Intelligence or Artificial General Intelligence is the intelligence of an entity (agent, machine, etc.), which has the capacity of learning any task as a human being.

- We will discuss the current state-of-the-art in this area and we will also outline the challenges and opportunities ahead.

Suggested Readings

- A key reference of this lecture can be found in Chapter 2:

Stuart Russell and Peter Norvig. Artificial Intelligence. A Modern Approach. Fourth Edition. Pearson. 2021.

Some of the definitions of this lecture were taken from that book.

Credits

- Some of the figures are from: Stuart Russell and Peter Norvig. Artificial Intelligence. A Modern Approach. Fourth Edition Pearson. 2020.