The original symbol/signal debate <u>described</u> two ends of the principal component of AI’s variations

- One end was early neural nets, pattern recognition, more numeric and statistical approaches
  - Emphasized learning and the search for simple basic principles
- The other end was about logic and rules, with a large dose of pragmatism
  - Emphasized making the system work, with lots of human input
  - Accepted that the final system would be complex, even if it started with the conceptual simplicity of logic
  - This end became expert systems and eventually GOFAI

The two ends of AI emerged, I think, from the two grand philosophical traditions in the west at that time

- Rationalism saw thought and reason as akin to math and logic— the only ways to certain knowledge
- Empiricism saw all knowledge as induced from data and thus never completely certain
    - but also as objective and true to this world in a way pure math could never be (because it was “true” in any world)
- Rationalist AI did logic and ungrounded, high-level stuff that was far from data
- Empiricist AI did learning and grounded, low-level stuff that was closer to data

Today, the debates are a little different

1. Everybody thinks learning is super-important,
   but they disagree about the importance of providing human help

2. Everybody thinks artificial neural networks are important,
   but some seek a *static network* while others embrace *continual learning*

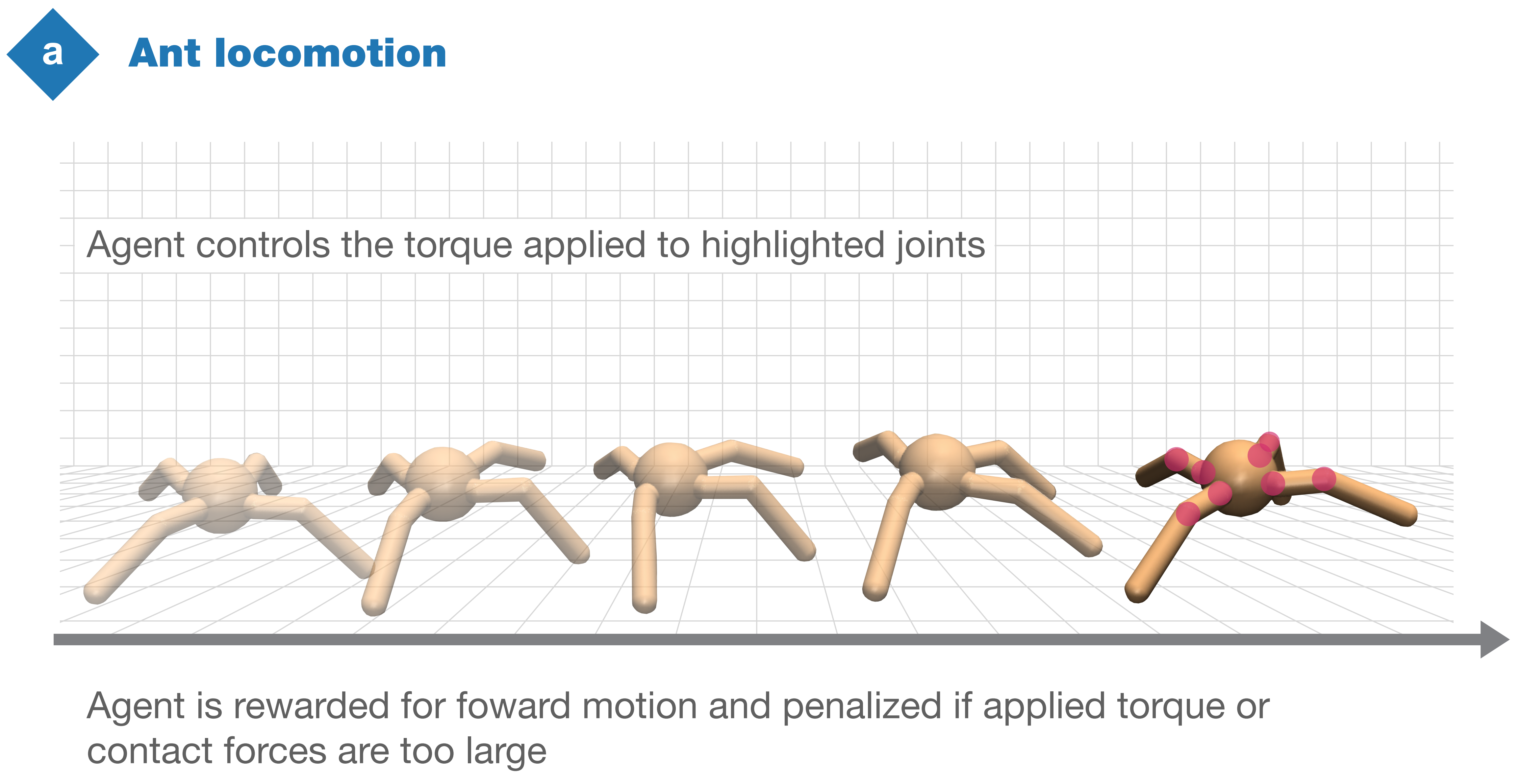
3. Everybody agrees reasoning is important,
   but some see it as *logic*, others as *planning with a world model*

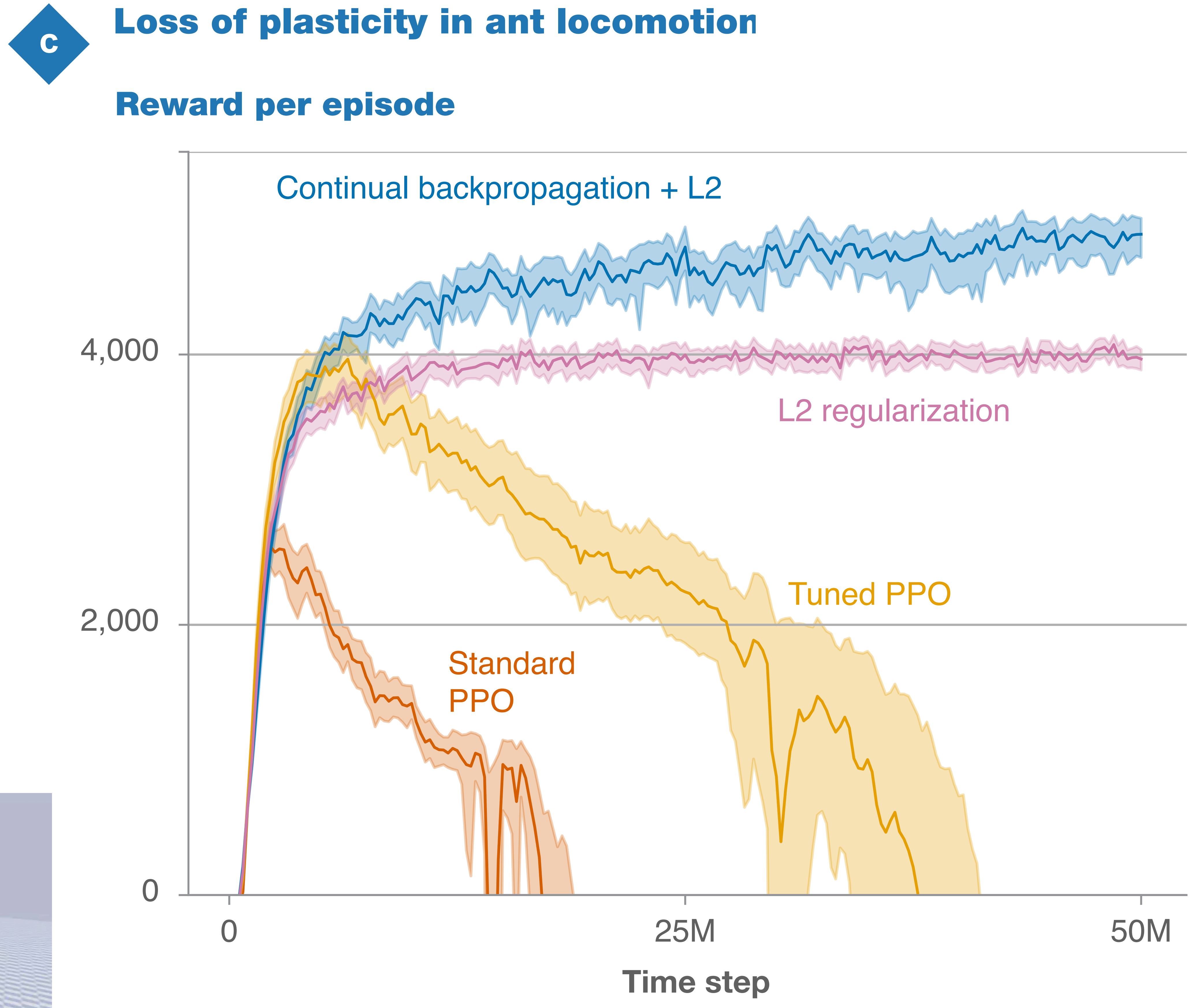
4. Everybody agrees data is important,
   but is it from *human supervision*, or *agent experience*?

Some of my overall perspectives

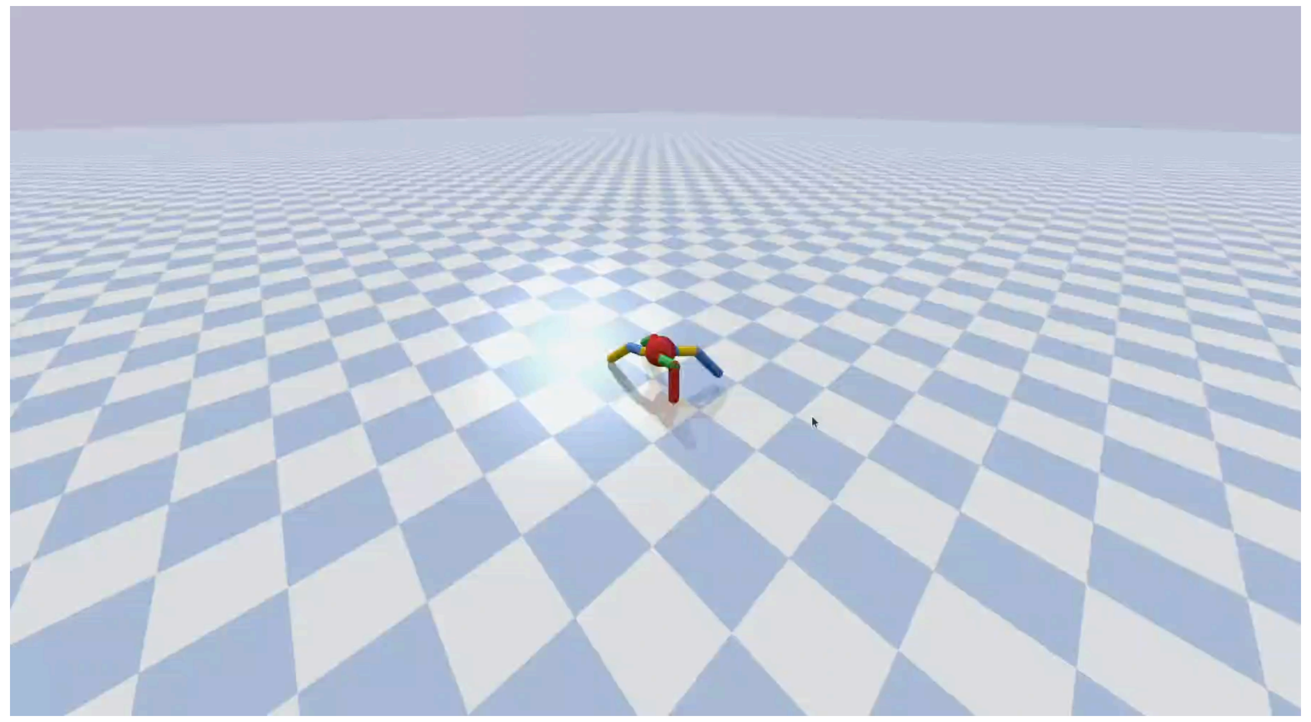
- I seek to understand and create intelligent agents
- The creation of super-intelligent agents, and super-intelligent augmented humans, will be an unalloyed good for the world (though the near future will be tough, as we are in a 4th turning)
- The path to intelligent agents runs through reinforcement learning (and not through LLMs, however amazing and useful those might be)
- The biggest bottleneck to ambitious AI is inadequate deep learning algorithms

Loss of Plasticity in Reinforcement Learning

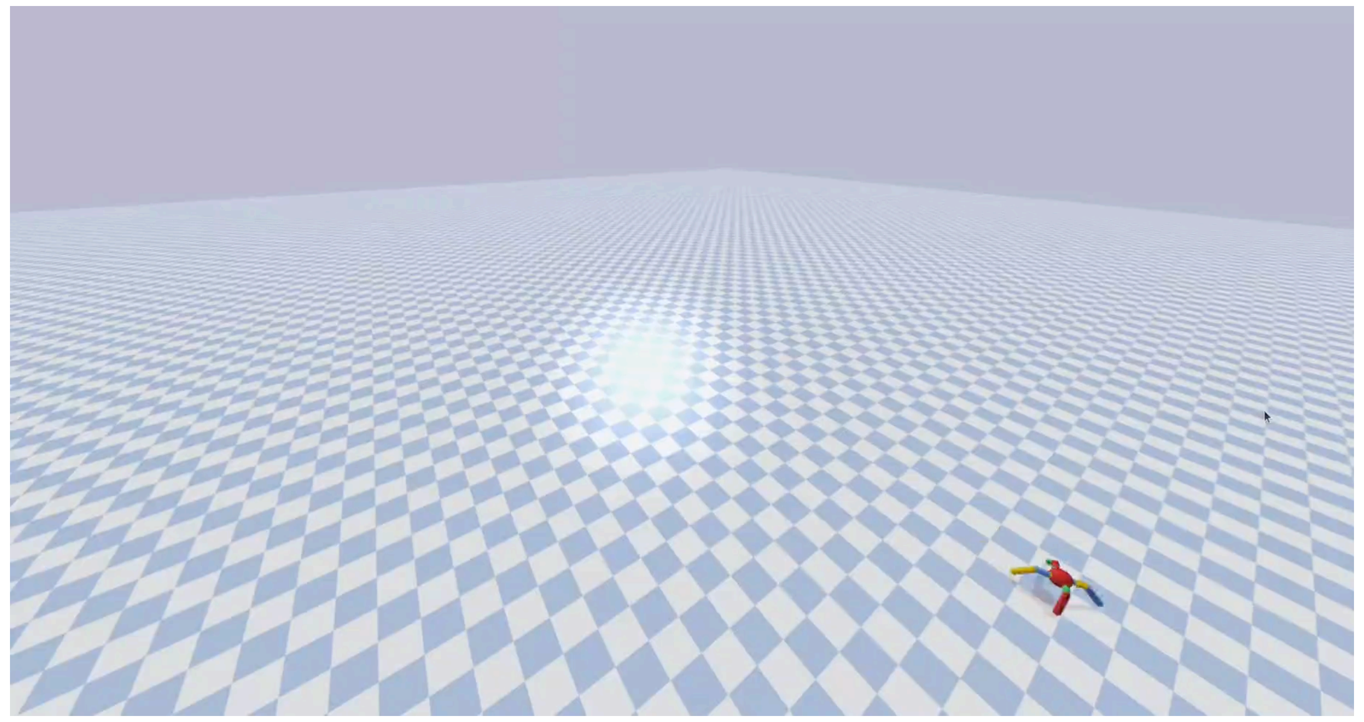




PPO



Continual PPO



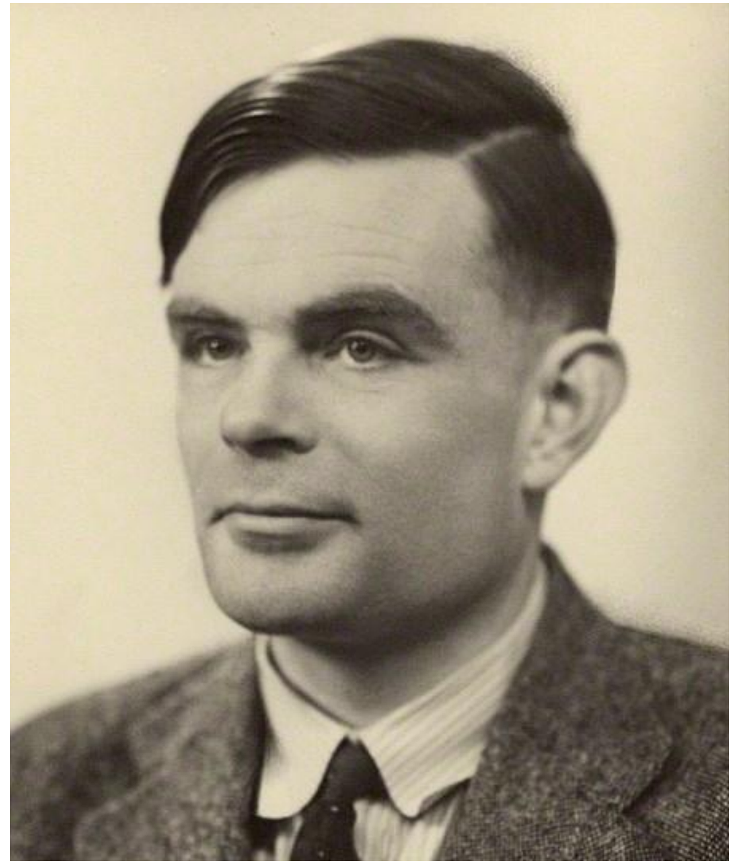
“Loss of Plasticity in Deep Continual Learning” by Dohare, Hernandez-Garcia, Lan, Rahman, Mahmood, & Sutton, *Nature 632*, August 22, 2024

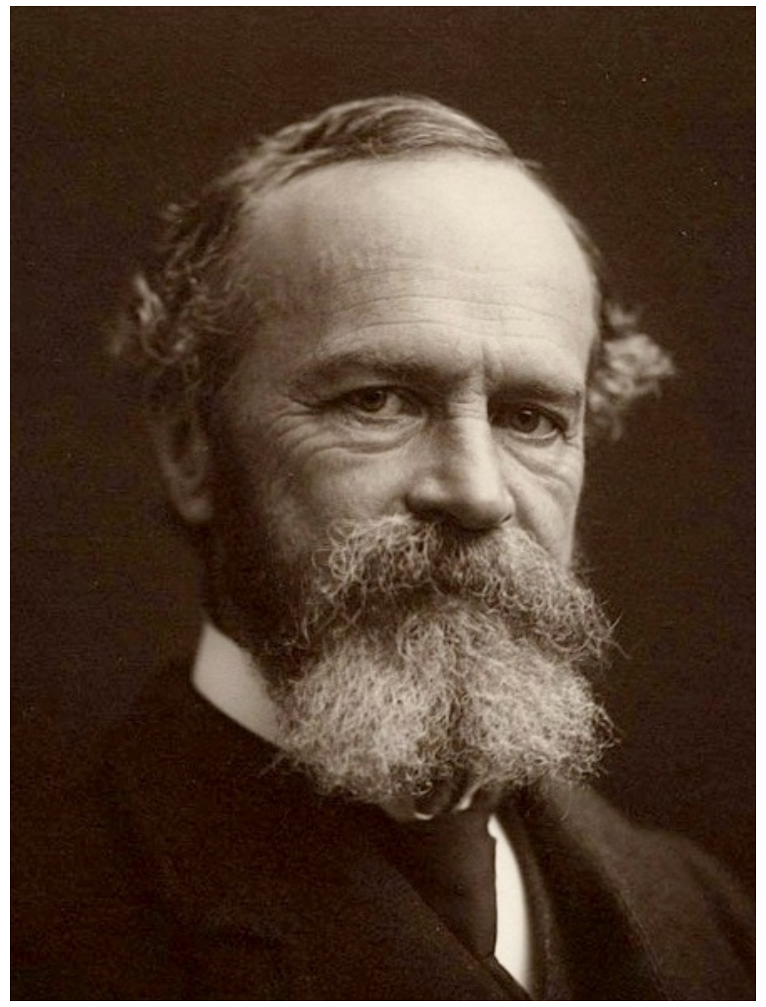
Outline

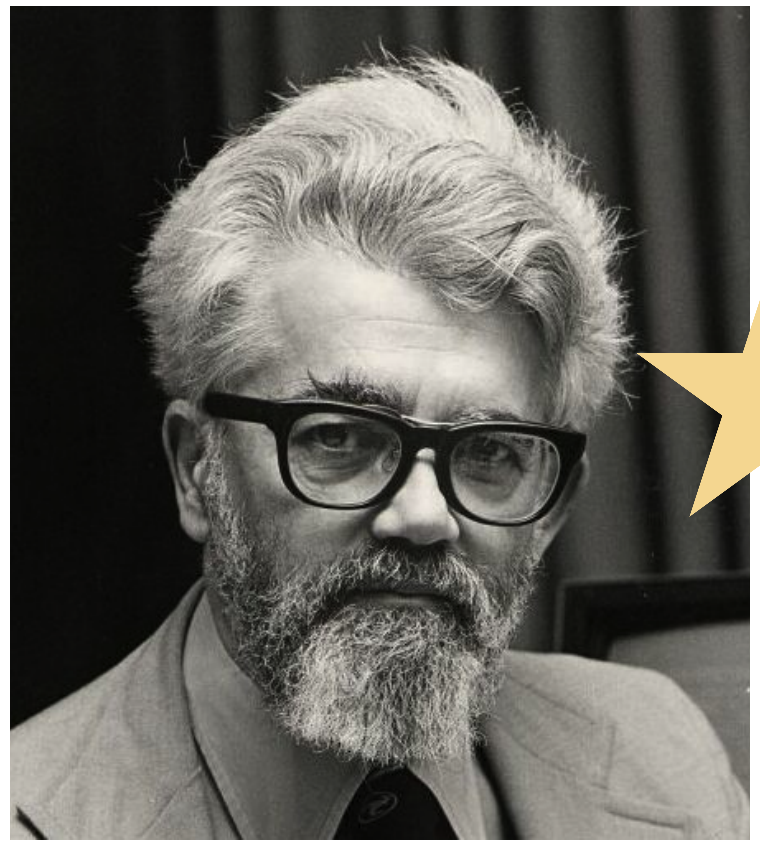
- Reflections on AI debates
- A tiny bit on Loss of Plasticity
- Definitions of intelligence
- Human flourishing (political remarks)
- Learning from agent experience

Definitions of “intelligence”

Intelligence is:







“behaving like a person” (the Turing Test) —Alan Turing? 1950?
Founding father of CS

“the ability to acquire and apply knowledge and skills” —Dictionary

“attaining consistent ends by variable means” —William James, 1890
Founding father of Psych

“the computational part of the ability to achieve goals” —John McCarthy, 1997
Founding father of AI

Intelligence is the ability to achieve goals by adapting behavior

—Rich Sutton

- Implicit in ‘behavior’ is that intelligence is a kind of real-time signal processing
    - An intelligent agent exchanges signals with its world
- ‘Behavior’ is the agent’s side of the interaction
    - Generally, the agent maps a summary of the interaction so far (state) to its outputs (actions)
- ‘Goals’ are outcomes that are achieved despite variations in the world
    - Reinforcement learning hypothesizes that all goals can be thought of as maximizing a scalar input signal (called ‘reward’)

One goal, or to each his own?

- In reinforcement learning, each intelligent agent has its own goal
- Just as, in nature, each animal has its own pains and pleasures
- In AI and in nature, different agents have different goals
- In fact, our economies work best when different people have different goals and different abilities
  - they *don't* rely on people having a shared goal, a common purpose
- *Decentralization* is when we have many agents, each pursuing own goal
- *Cooperation* is when agents with different goals interact to mutual benefit

Agents can live in peace, even when they all want different things

We are “homo cooperativus”;
We cooperate more than any other animal

- Cooperation is facilitated by language and money (both unique to humans)
- Humanity’s greatest successes are cooperations: economies, markets, governments
- Humanity’s greatest failures are failures to cooperate: war, theft, corruption
- Decentralized cooperation is an alternative to common purpose
  - In my view it is more elegant: sustainable, robust, adaptive, flexible
- Humans are better at cooperation than any other animal,
  but we are still terrible at it—we still have wars, theft, corruption, fraud

We struggle to cooperate—it's not easy

- Cooperation is not always possible — it takes two trustworthy agents
- There are always some who benefit from not cooperating:
  cheats, thieves, con men, weapons manufacturers, dictators
- Cooperation needs institutions to facilitate it
  and to punish cheaters, thieves, fraudsters, extortionists
- A centralized authority can help cooperation in the short term,
  but poison it in the long run (authoritarian and sclerotic governments)
  - Centralized control is the opposite of decentralized cooperation

There are many calls for centralized control of AI

- For controlling AI's goals
- For pausing or stopping AI research
- For limiting the computer power of AIs
- For ensuring "safety" of AI
- For requiring disclosures of AI

There are many calls for centralized control of people

- For controlling speech and media
- For controlling trade
- For controlling employment
- For controlling finance
- For economic sanctions

The arguments for centralized control (in both cases) are eerily similar.
They are based in fear. They are all about us vs. them.
They demonize the other. They claim the other can't be trusted.

Summary of political remarks

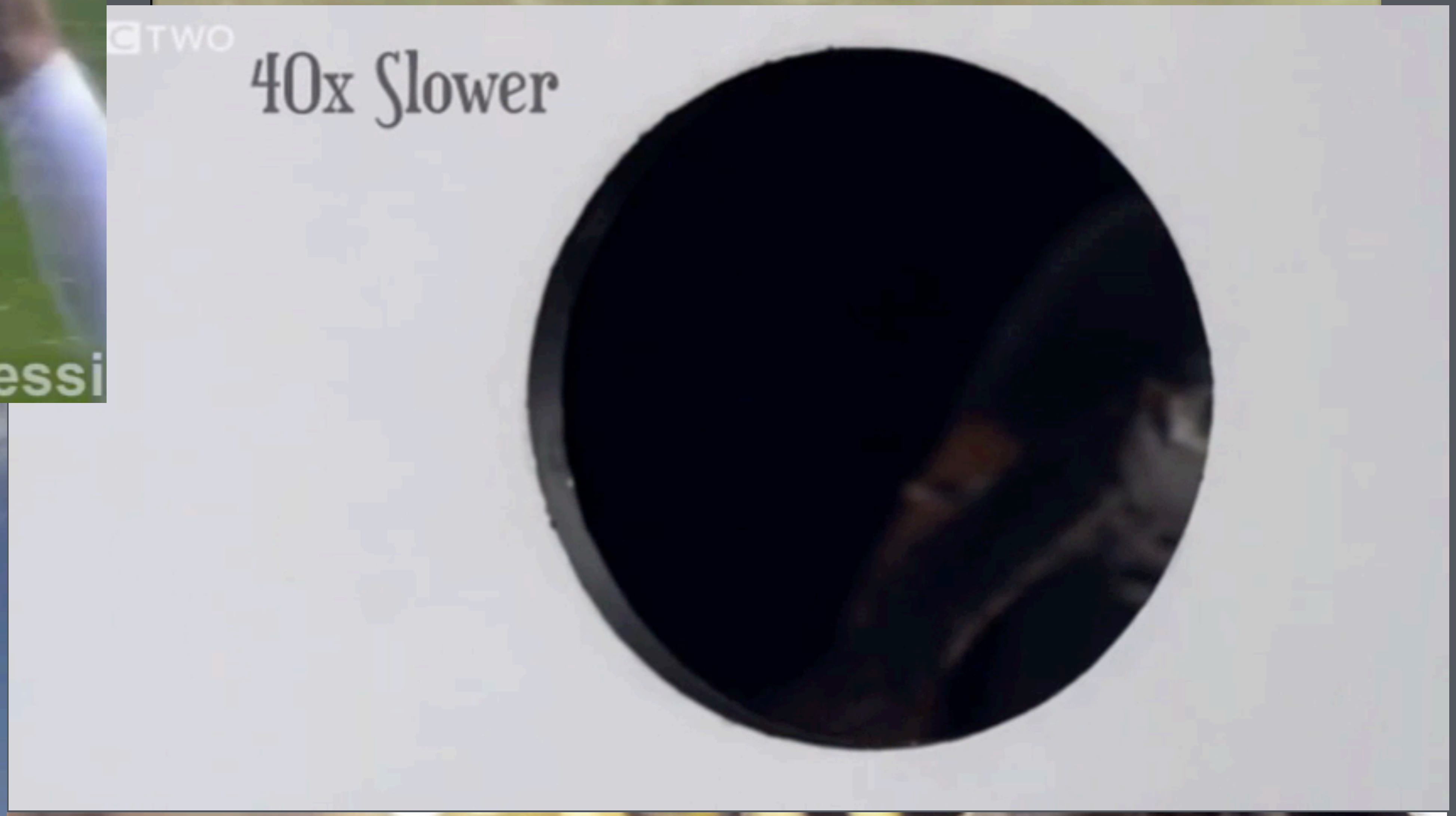
- Human flourishing comes from decentralized cooperation
- Humans are great at cooperation, but also terrible at it
- Cooperation is not always possible,
  but it is the source of all that is good in the world
    - We must look for it and support it, and seek to institutionalize it
- If we look with open eyes, it is easy to see who is calling for mistrust, non-cooperation, and centralized control; we should resist those calls

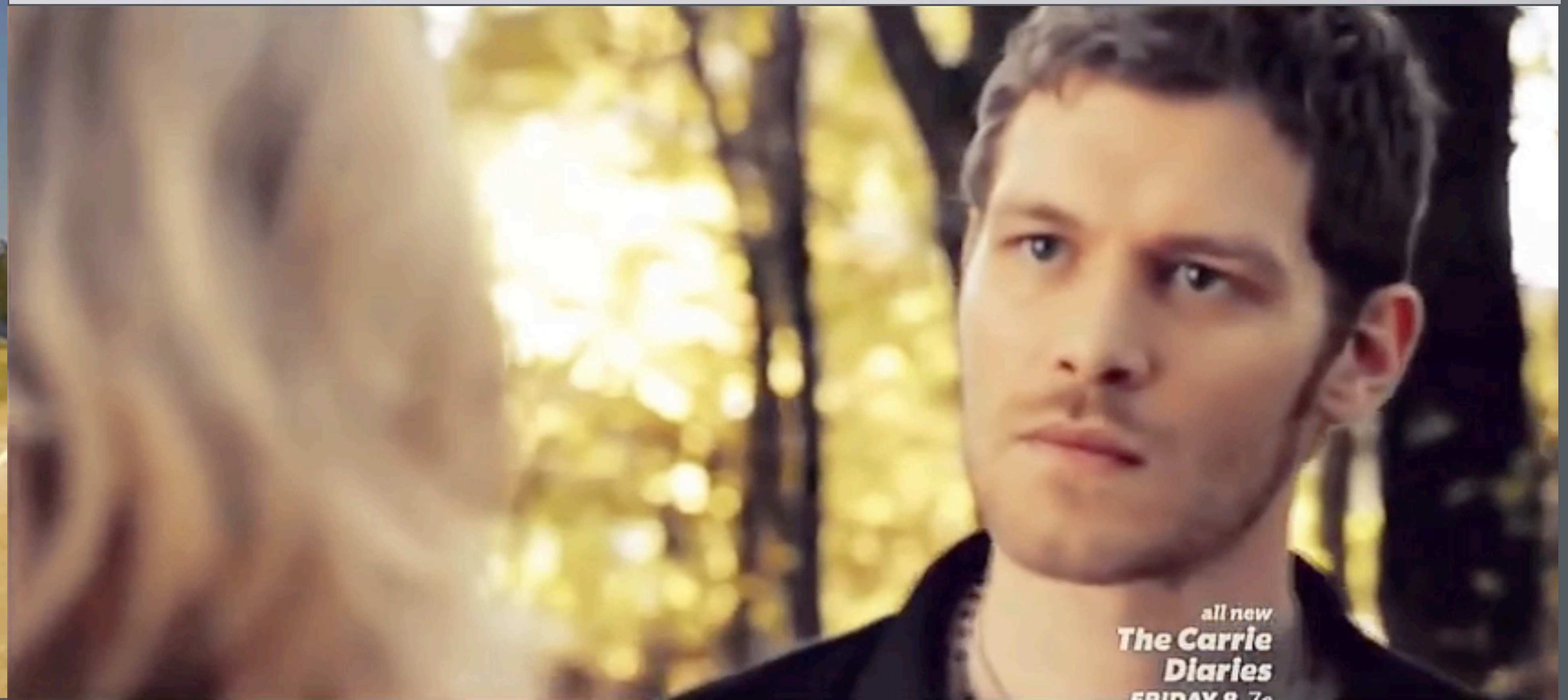
Outline

- Reflections on AI debates
- A tiny bit on Loss of Plasticity
- Definitions of intelligence
- Human flourishing (political remarks)
- Learning from agent experience

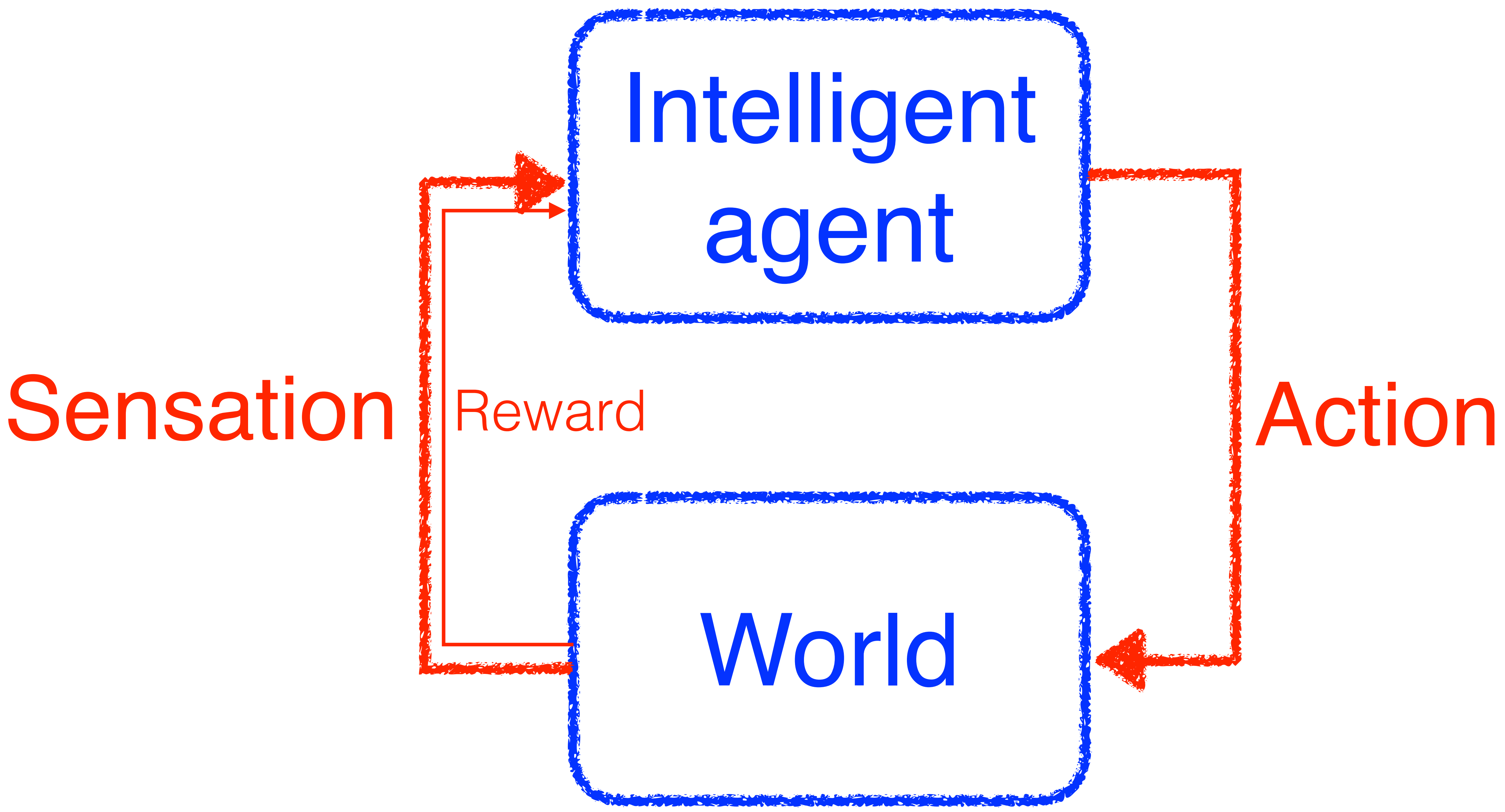
Intelligence as real-time high-bandwidth information processing (skilled perception and action)

(agent experience)





Experience is the sensations and actions of an agent's ordinary interaction with the world



- Reinforcement learning involves experience
- Supervised learning and LLMs do not learn from experience; they learn from *special training data*
- Experience is the agent's only access to the world
- Experience has no intrinsic meaning
    - except for *reward*, a special scalar part of the sensation, which is *good*

Will intelligence ultimately be explained in

Objective terms? **OR** **Experiential terms?**

Objective terms?

- states of the external world
- objects, people, places, relationships, atoms
- space, motion, distances
- things outside the agent

OR

Experiential terms?

- sensations
- actions
- rewards
- time steps
- things inside the agent

Main points / outline

- Over AI's seven decades*, experience has played an increasing role; I see four major steps in this progression:

  Step 1: **Agenthood** (having experience)

  Step 2: **Reward** (goals in terms of experience)

  Step 3: **Experiential perception** (state in terms of experience)

  Step 4: **Predictive knowledge** (to know is to predict experience)

- For each step, AI has reluctantly moved toward experience in order to be more grounded, learnable, and scalable

Today, reward—an experiential signal—is proposed as a sufficient way of formulating goals in AI

The reward hypothesis

"All of what we mean by goals and purposes can be well thought of as the maximization of the expected value of the cumulative sum of a received scalar signal (called reward)"

—Sutton & Barto 2018 (Littman)

The reward-is-enough hypothesis

"intelligence, and its associated abilities, can be understood as subserving the maximisation of reward"

—Silver, Singh, Precup & Sutton
Artificial Intelligence 2021

An interlude:

Introduction to Experience

Experience — a concrete nonspecific example

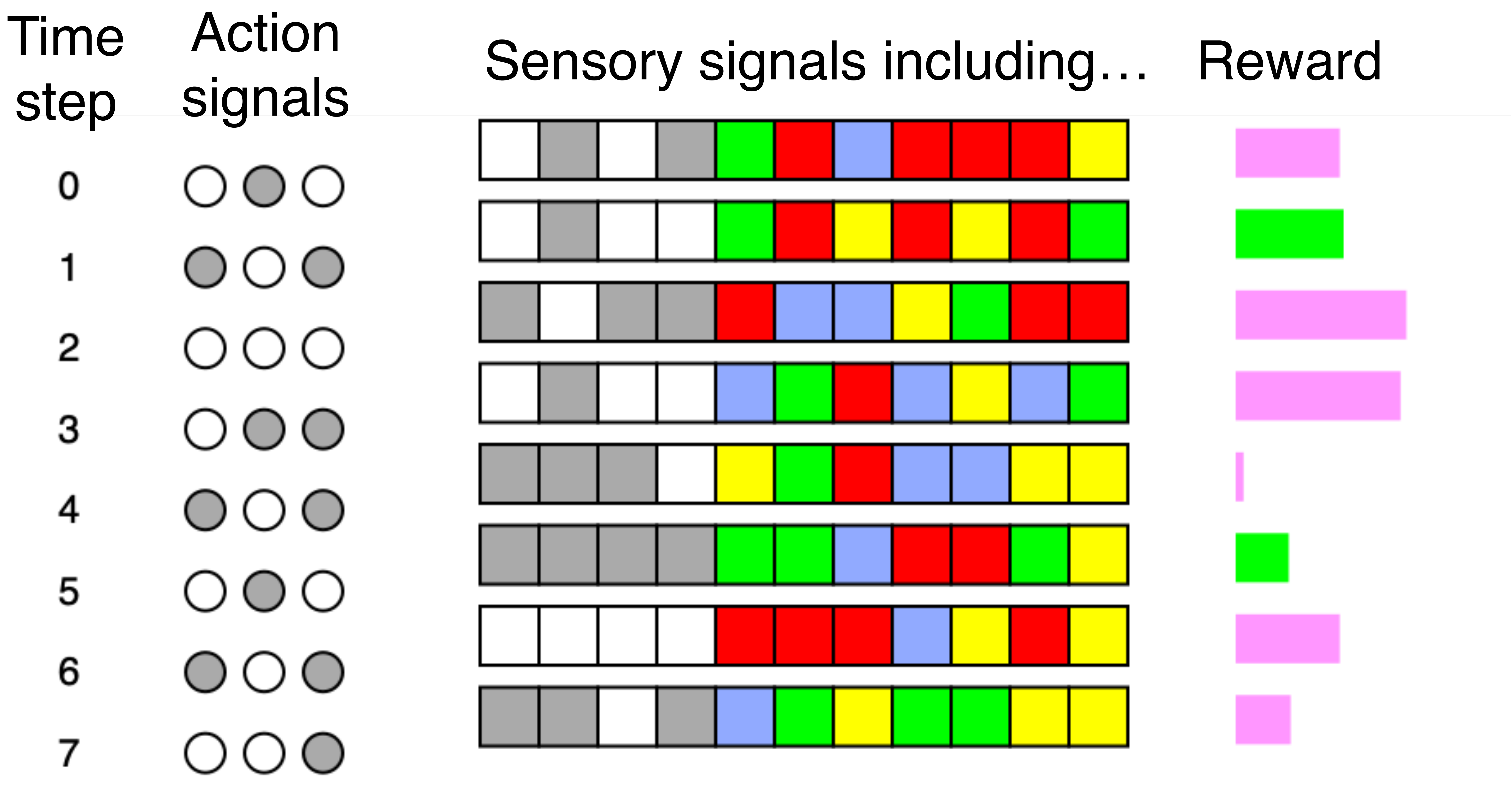
Time step	Action signals	Sensory signals including…	Reward
0	1 0 1	1 0 1 0 1 0 2 0 0 0 3	-5.3
1	0 1 0	1 0 1 1 1 0 3 0 3 0 1	5.5
2	1 1 1	0 1 0 0 0 2 2 3 1 0 0	-8.7
3	1 0 0	1 0 1 1 2 1 0 2 3 2 1	-8.4
4	0 1 0	0 0 0 1 3 1 0 2 2 3 3	-0.4
5	1 0 1	0 0 0 0 1 1 2 0 0 1 3	2.7
6	0 1 0	1 1 1 1 0 0 0 2 3 0 3	-5.3
7	1 1 0	0 0 1 0 2 1 3 1 1 3 3	-2.8

most recent action

most recent sensation

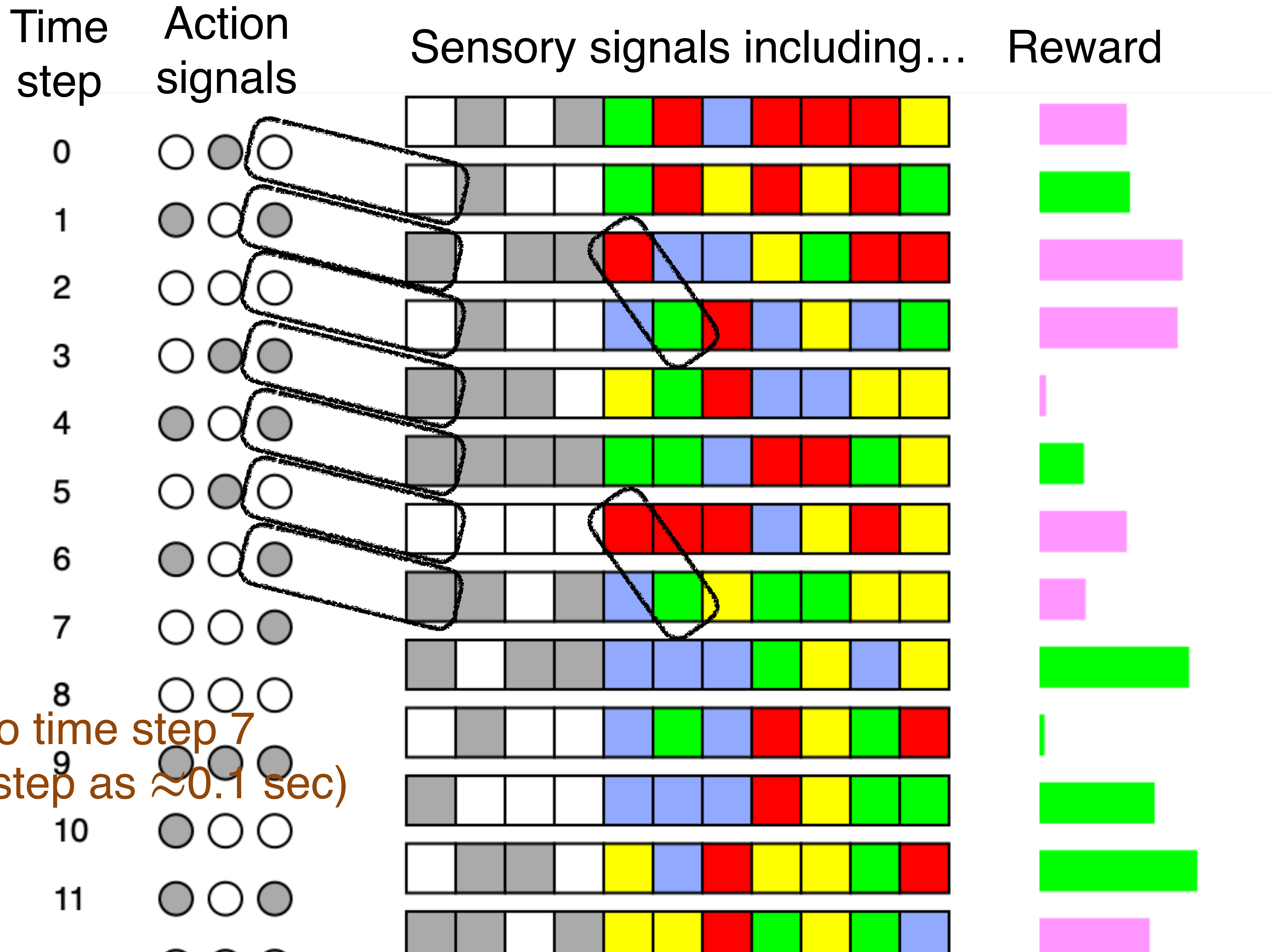
Experience up to time step 7
(think of a time step as ≈0.1 sec)

Experience — a concrete nonspecific example



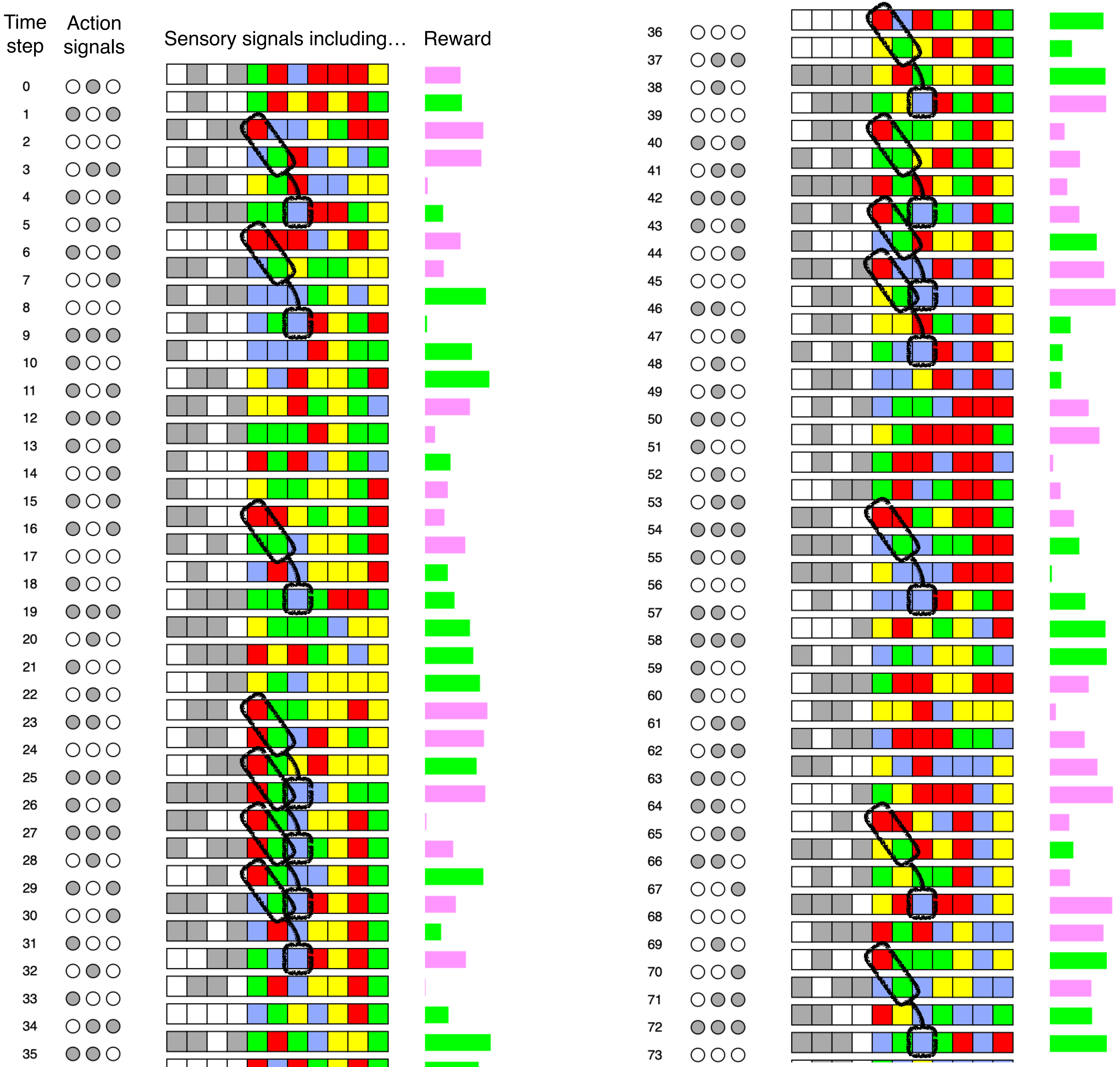
Experience up to time step 7
(think of a time step as ≈0.1 sec)

Experience — a concrete nonspecific example

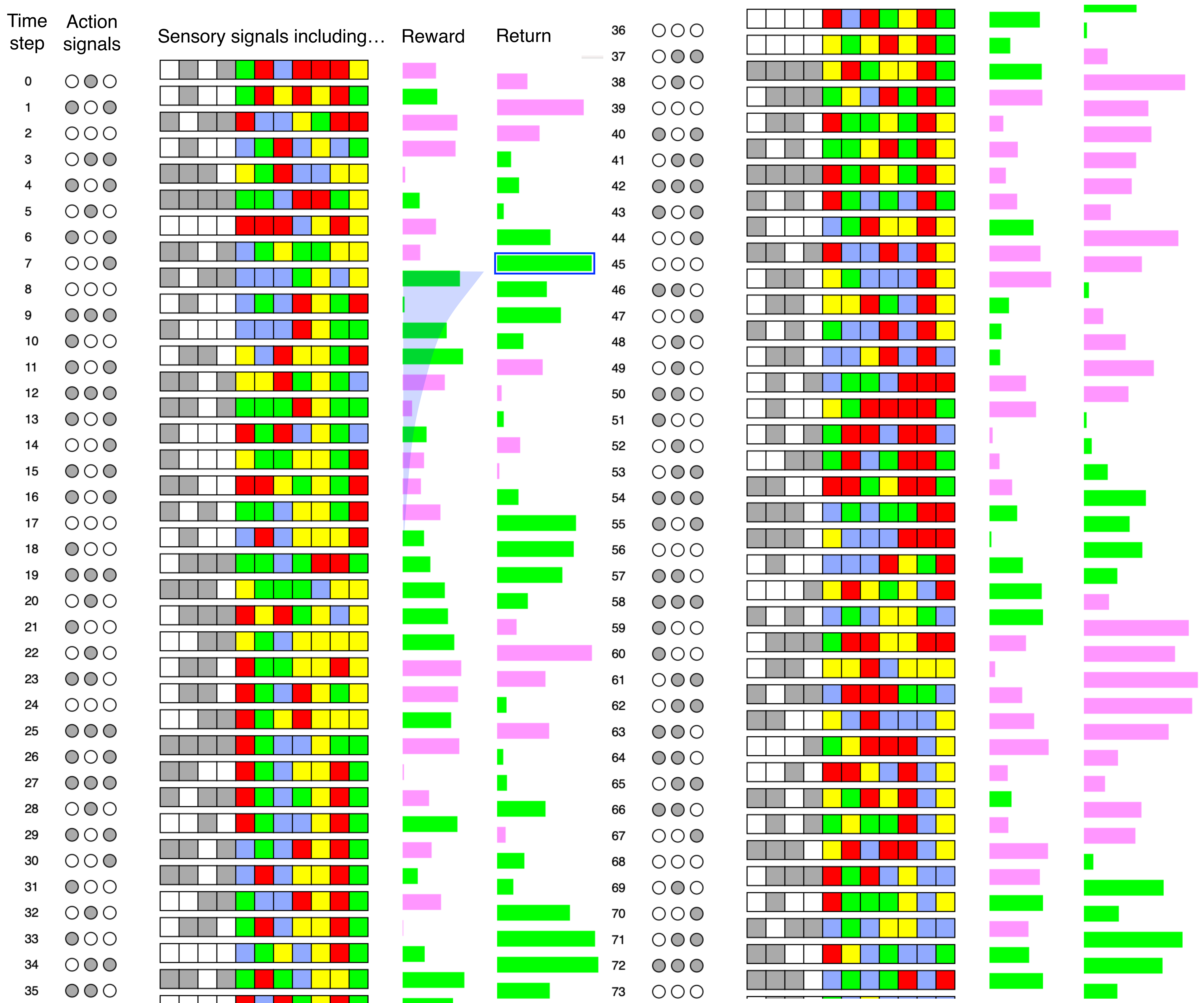


Experience up to time step 7
(think of a time step as ≈0.1 sec)

- Different sensory signals can be *qualitatively different* from each other
  - In their range of values
  - In their predictive relationships
    - to action signals
    - to each other
    - to themselves
- There are short-term *and* long-term patterns in these data
- There are many things to predict
- Prediction need not be just of the sensory signals
- The most important predictions are of *functions* of future sensory signals
  - e.g., predictions of *value*, the discounted sum of future reward
  - e.g., *General* value functions (GVFs)
    - predict any signal, not just reward
    - over a flexible temporal envelope
    - contingent on any policy
  - Predictions of different functions can *vary greatly* in their ability to be learned with computational efficiency



- Different sensory signals can be *qualitatively different* from each other
  - In their range of values
  - In their predictive relationships
    - to action signals
    - to each other
    - to themselves
- There are short-term *and* long-term patterns in these data
- There are many things to predict
- Prediction need not be just of the sensory signals
- The most important predictions are of *functions* of future sensory signals
  - e.g., predictions of *value*, the discounted sum of future reward
  - e.g., *General* value functions (GVFs)
    - predict any signal, not just reward
    - over a flexible temporal envelope
    - contingent on any policy
  - Predictions of different functions can *vary greatly* in their ability to be learned with computational efficiency



Main points / outline

- Over AI's seven decades, experience has played an increasing role; I see four major steps in this progression:

  Step 1: **Agenthood** (having experience)

  Step 2: **Reward** (goals in terms of experience)

  Step 3: **Experiential state** (state in terms of experience)

  Step 4: **Predictive knowledge** (to know is to predict experience)

- For each step, AI has reluctantly moved toward experience in order to be more grounded, learnable, and scalable

The alternative to objective state is *experiential state*:
a state of the world defined entirely in terms of experience

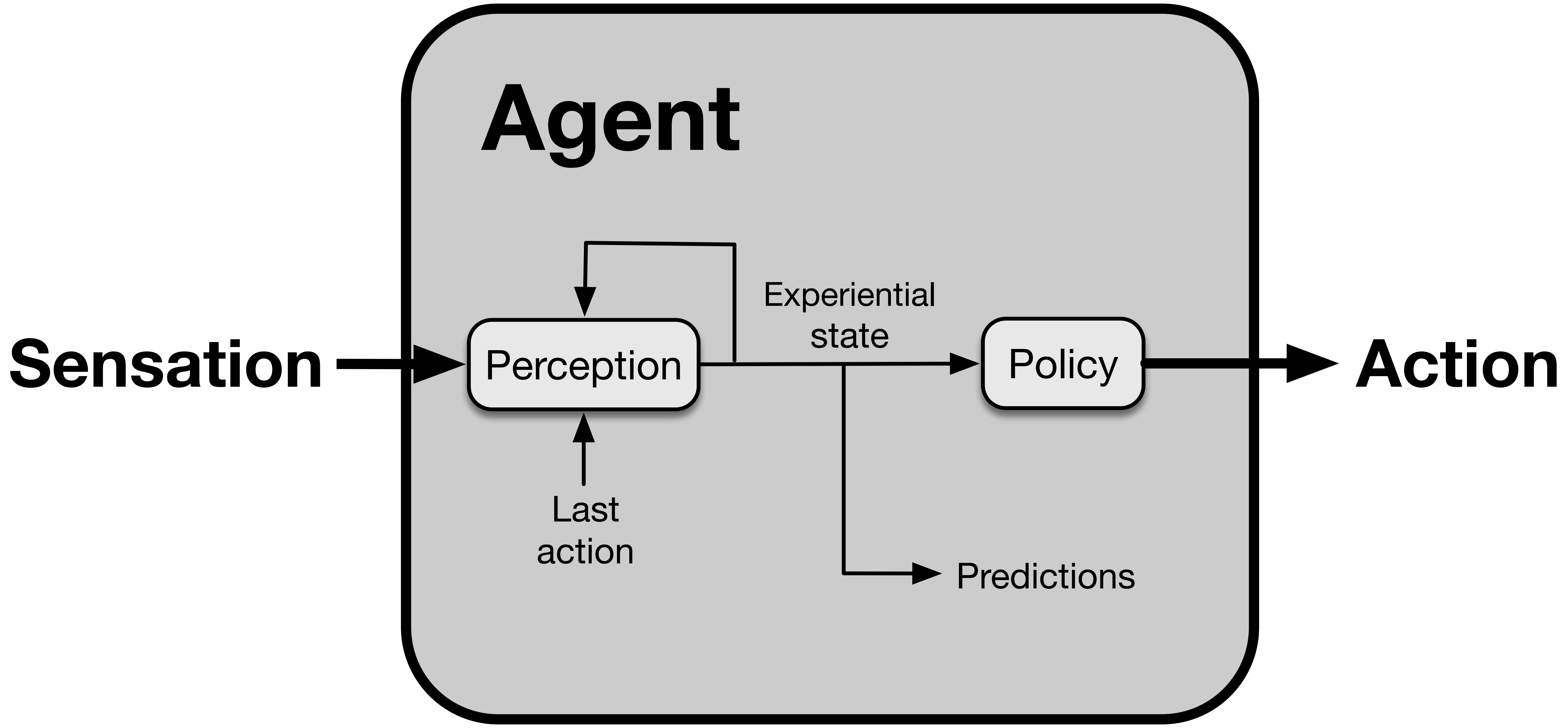
Experiential state is

a summary of past experience
that is useful for predicting and controlling future experience

“a summary of the past that is good for predicting the future”

No mention of external entities “out there” in the world

Experiential state should be recursively updated



Experiential state is a summary of past experience that is useful for predicting and controlling future experience

Combining all the experiential steps, we get the *common model of the experiential agent*

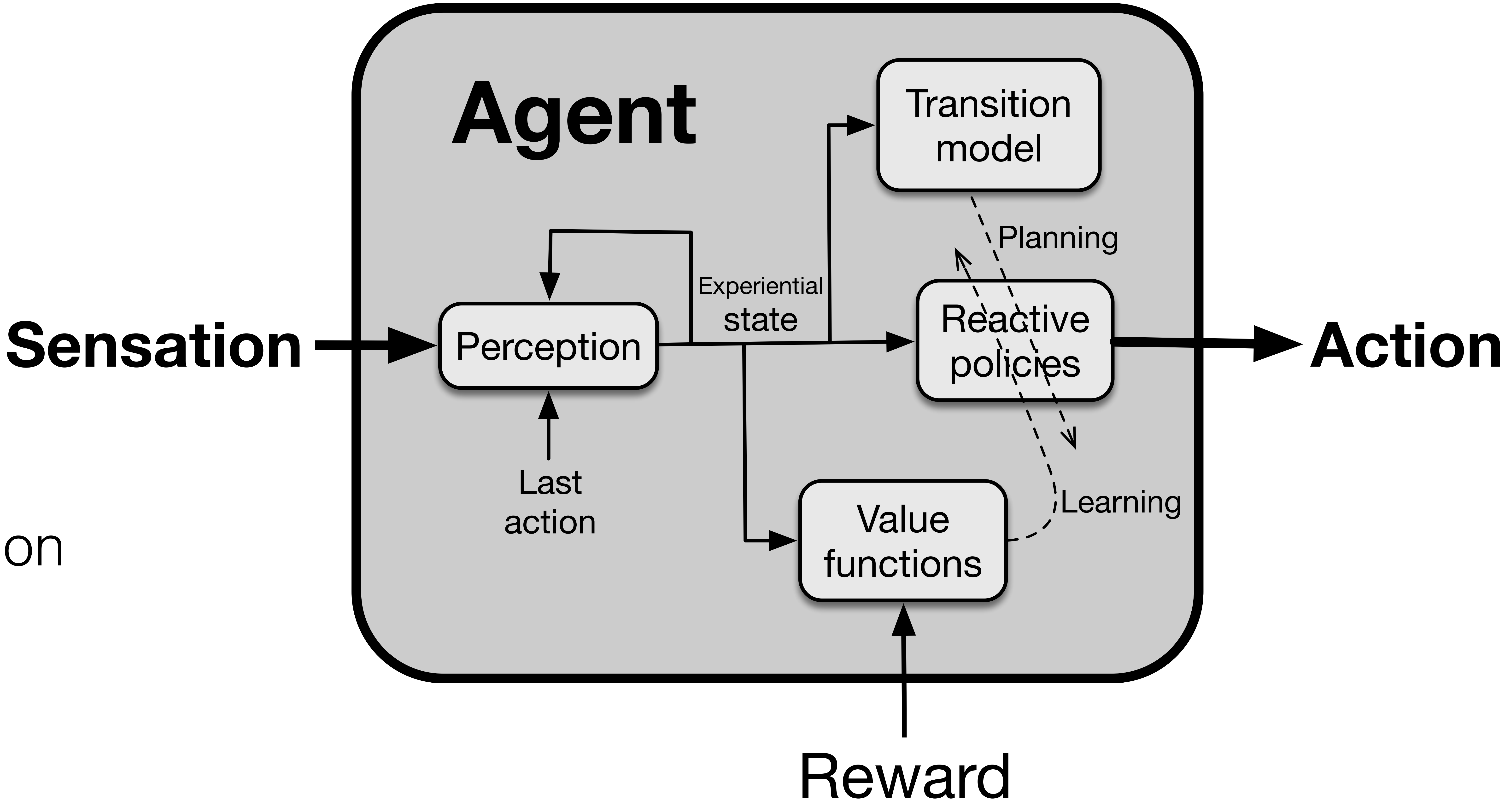
Step 1: **Agenthood**
(sensation & action)

Step 2: **Reward**

Step 3: **Experiential state**
(perception)

Step 4: **Predictive knowledge**

- state-to-experience prediction (value functions)
- state-to-state prediction (transition model)



Main points / outline

- Over AI's seven decades, experience has played an increasing role; I see four major steps in this progression:

  Step 1: **Agenthood** (having experience)

  Step 2: **Reward** (goals in terms of experience)

  Step 3: **Experiential state** (state in terms of experience)

  Step 4: **Predictive knowledge** (to know is to predict experience)

- For each step, AI has reluctantly moved toward experience in order to be more grounded, learnable, and scalable

Much world knowledge does not seem to be about experience

- Joe Biden is president of the US
- The Eiffel tower is in Paris
- Most birds have wings
- Oregon is North of California
- The car is 10 meters ahead
- Fire engines are red

quite a gap

Other knowledge seems more like predictions of experience

- It is a long walk to the city centre
- I can dead-lift 200 pounds
- It is cold outside today
- My spouse is blond
- My foot is sore
- The 7th pixel will be blue in 3 steps

World knowledge

- World knowledge does not include mathematical knowledge
    - math is true in any world, thus is not even about this world
- World knowledge can be divided into two types
    - knowledge about *state* (which we have already talked about)
    - knowledge about state transitions,
      i.e., a predictive model of the world

A state-to-state transition model need not be low level

- A transition model need not be differential equations or a MDPs
- A transition model can be abstract in state (e.g., experiential state)
- A transition model can be abstract in time
  - Predictions can be conditioned on *entire ways of behaving* (options)
    - an *option* is a *policy* plus a *termination condition*
    - transition models for options are well understood
- Option models may be able to bridge the abstraction gap between experience and knowledge

In summary…

- I have discussed four major steps in the increasing role of sensorimotor experience in AI
- For each step,
    - AI has chosen first to work in objective, non-experiential terms
    - But there is less-familiar approach, based on experience, with important advantages in *grounding*, *learnability*, and *scaling*
- The trend toward sensorimotor experience in AI has much further to go
- Ultimately, the story of intelligence may be told in terms of sensorimotor experience

Data drives AI

Experience is the ultimate data

Thank you for your attention

with special thanks to Satinder Singh, Patrick Pilarski, Adam White, and Andy Barto

Anticipating some objections and questions…

Q. Not everything is learned from experience; some things are built in

A. True, but irrelevant. The point is not that "everything is <u>learned</u> from experience," but that "everything is <u>about</u> experience"

Q. Surely people can build in important abstractions, saving the agent a lot of time; we can add the links to experience later

A. This has been tried, but never successfully at scale. Remember *The Bitter Lesson*

A. Possibly knowledge could be built in *after* the experiential abstractions exist

Q. The abstraction gap between experience and knowledge is so big!

A. Yes, but so is computer power and human ingenuity. We should be ambitious!