# The Increasing Role of Sensorimotor Experience in AI


Rich Sutton

University of Alberta

DeepMind Alberta

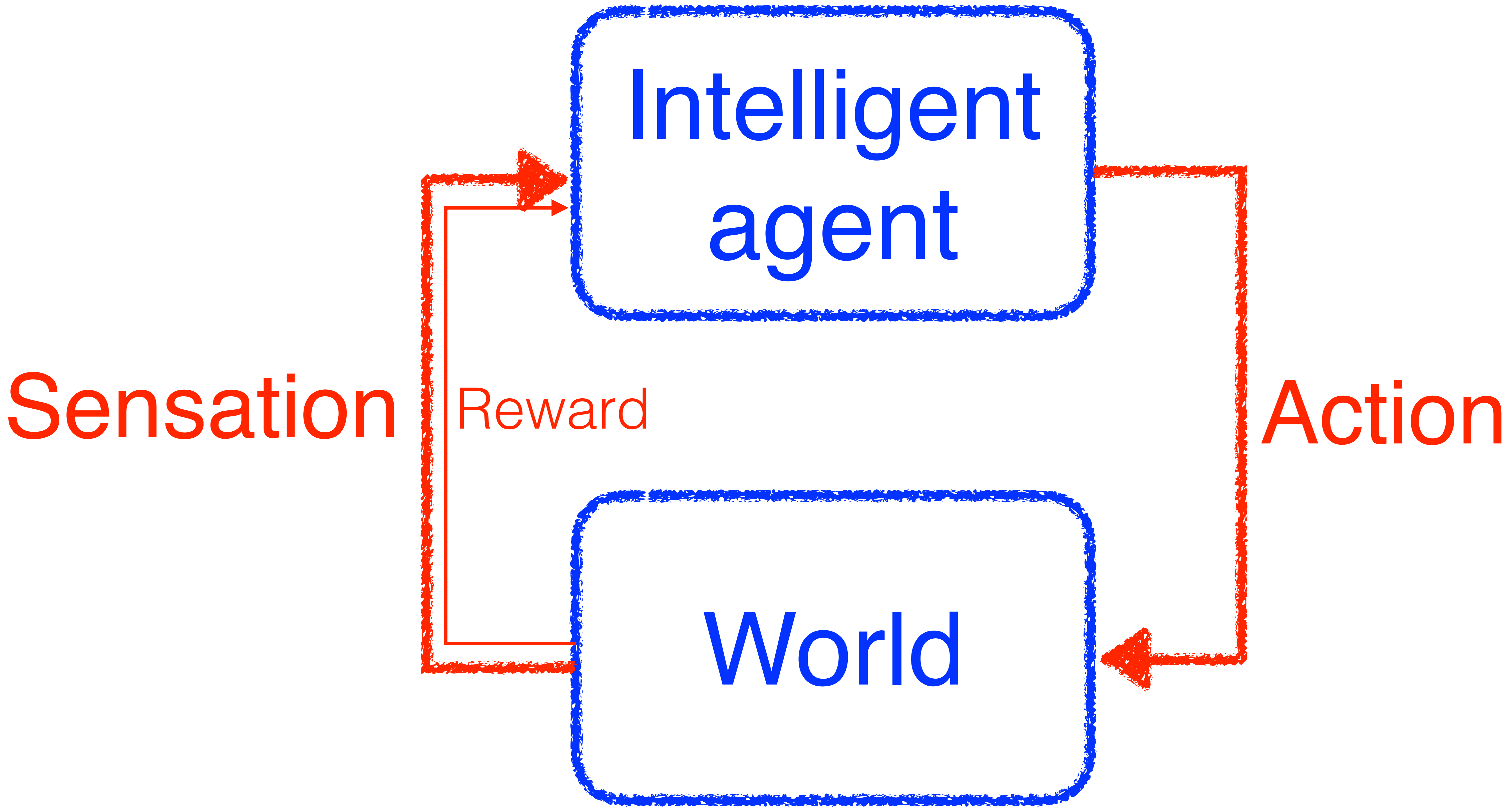
Alberta Machine Intelligence Institute

Reinforcement Learning and Artificial Intelligence Lab



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

# *Sensorimotor experience* is the sensations and actions of an agent's ordinary interaction with the world




- Reinforcement learning involves experience
- Predictive learning involves experience
- Supervised learning does not involve experience; it learns from *special training data*

- Experience is the agent's only access to the world

- Experience has meaning only by its relationship to other experience
  - except for *reward*, a special scalar part of the sensation, which is *good*

# Will intelligence ultimately be explained in

Experiential terms? **OR** Objective terms?

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

- states of the external world
- objects, people, places, relationships, atoms
- space, motion, distances
- things outside 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 state** (state in terms of experience)

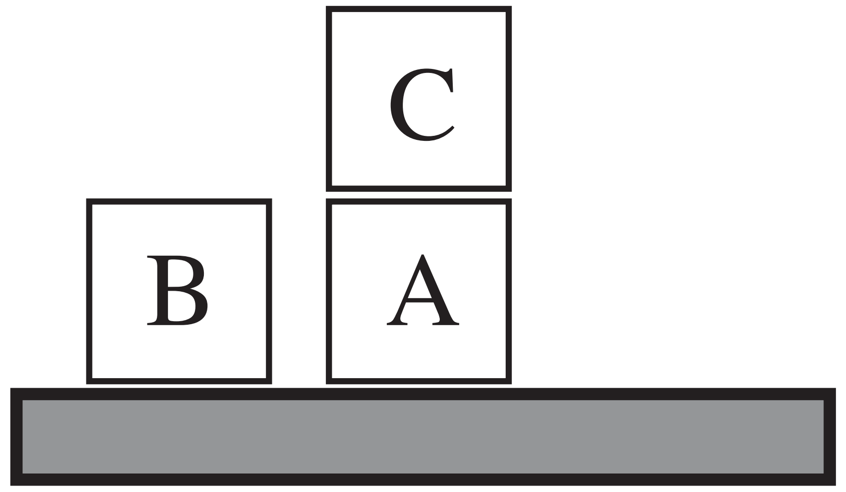
  Step 4: **Predictive knowledge** (to know is to predict experience)

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

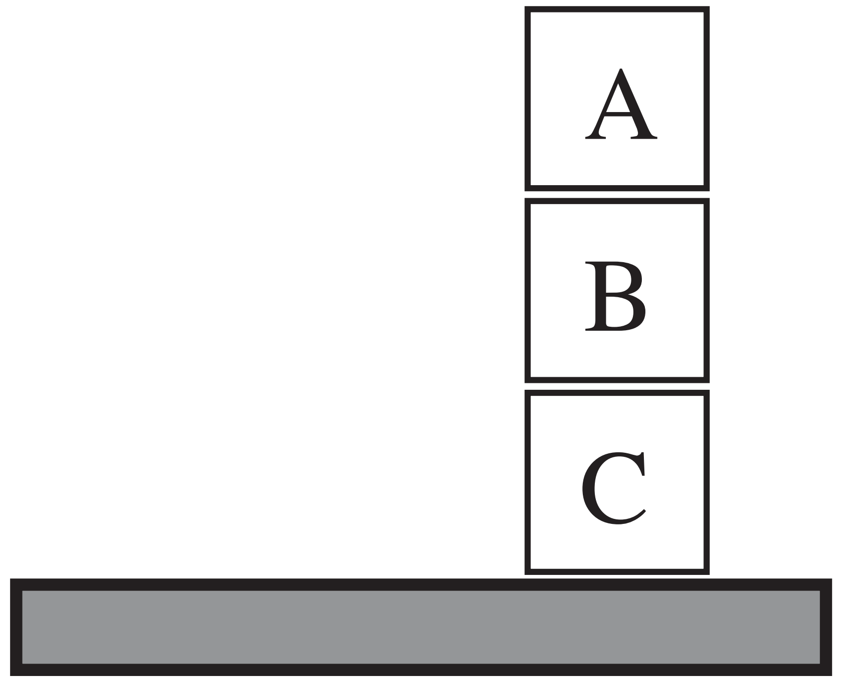
# Step 1: Agenthood

(having experience)

# Experience was rare in early AI systems (1954–1985)




Start State




Goal State

—Figures from Russell & Norvig, 2008

- Most AI systems were problem solvers and question answerers with no sensations or actions (robotics was an exception)
  - A typical problem was a start state and a goal state, with operators defined not as actions, but as state transitions
  - A solution was a sequence of operators guaranteed to go from start to goal
  - There was no sensing or acting (operators were deterministic)
    The solution was never actually executed!

# Early AI systems did not involve experience; They could:

—*Principles of Artificial Intelligence* by Nils Nilsson 1980

“diagnose diseases,

plan the synthesis of complex organic chemical compounds,

solve differential equations in symbolic form,

understand limited amounts of human speech and natural language text, and

write small computer programs to meet formal specifications”

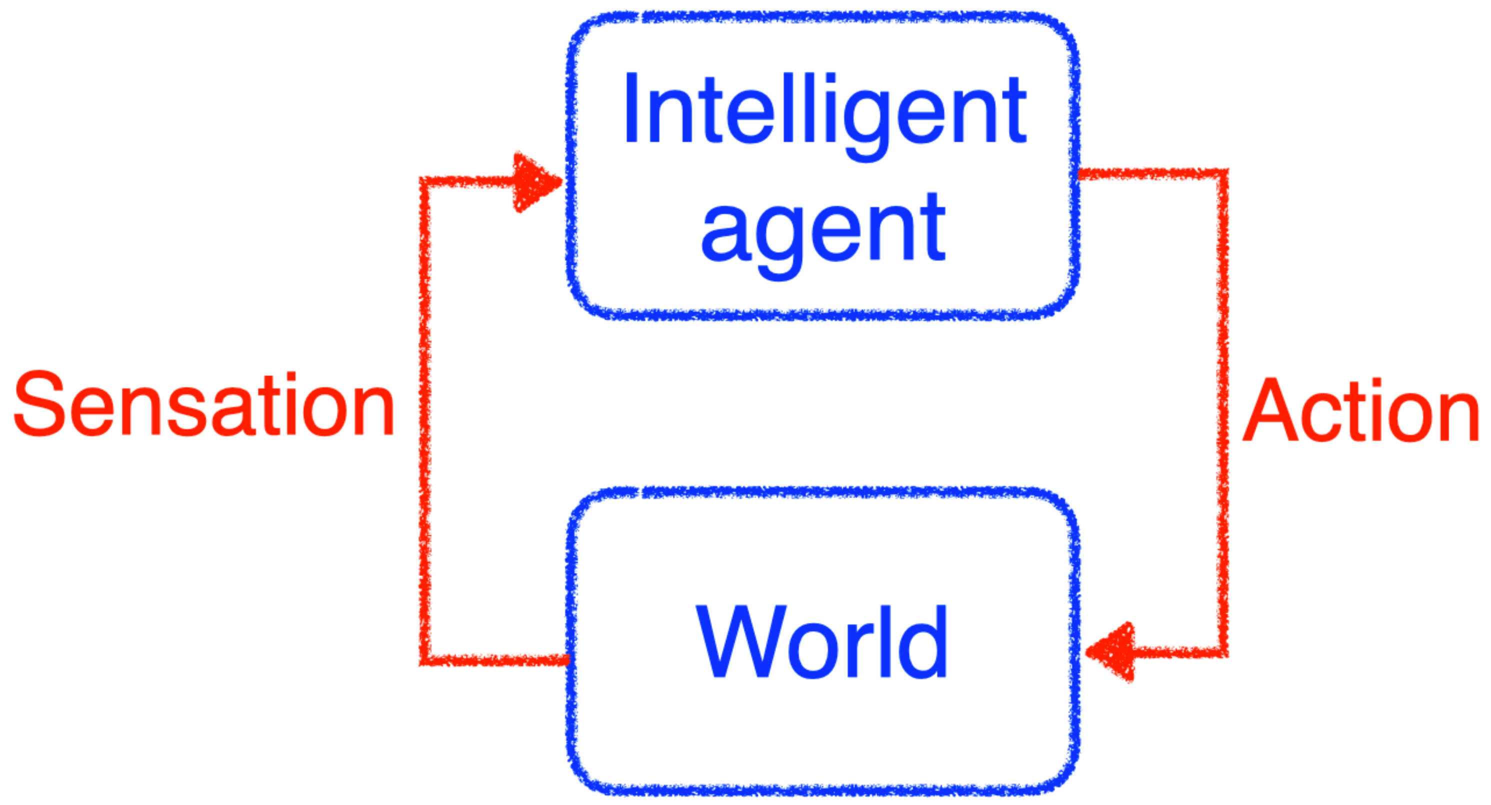
# For 30 years now AI has focused on building *agents*

From the 1995 edition of the standard AI textbook (Russell & Norvig):

“The unifying theme of the book is the concept of an intelligent agent”

“In this view, the problem of AI is to describe and build agents that receive percepts from the environment and perform actions”

Experience used to be rare in AI, but now it is the standard, modern approach



# 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

# Step 2: Reward

(goals in terms of experience)

# Today, reward (a single number over time) is proposed as a sufficient way of formulating goals in AI

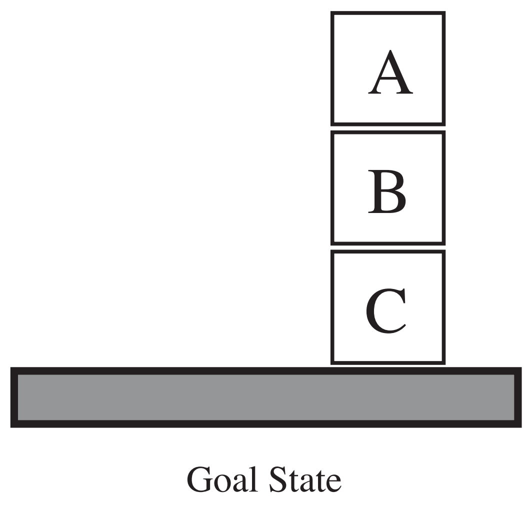
## The reward-is-enough hypothesis

> "intelligence, and its associated abilities, can be understood as subserving the maximisation of reward"
>
> —Silver, Singh, Precup & Sutton
> *Artificial Intelligence* 2021

# But still, for many, reward is <u>not</u> enough

- Enough for animals maybe, enough for engineering okay, but not enough for people, not enough for intelligence
  - A single number? From outside the mind!?
- Reward just seems too small. Too reductive. Too demeaning.
- Surely peoples' goals are grander
  - to raise a family, to save the planet, to contribute to human understanding, or to make the world a better place
- Surely our goals are more than just maximizing our pleasure and comfort!

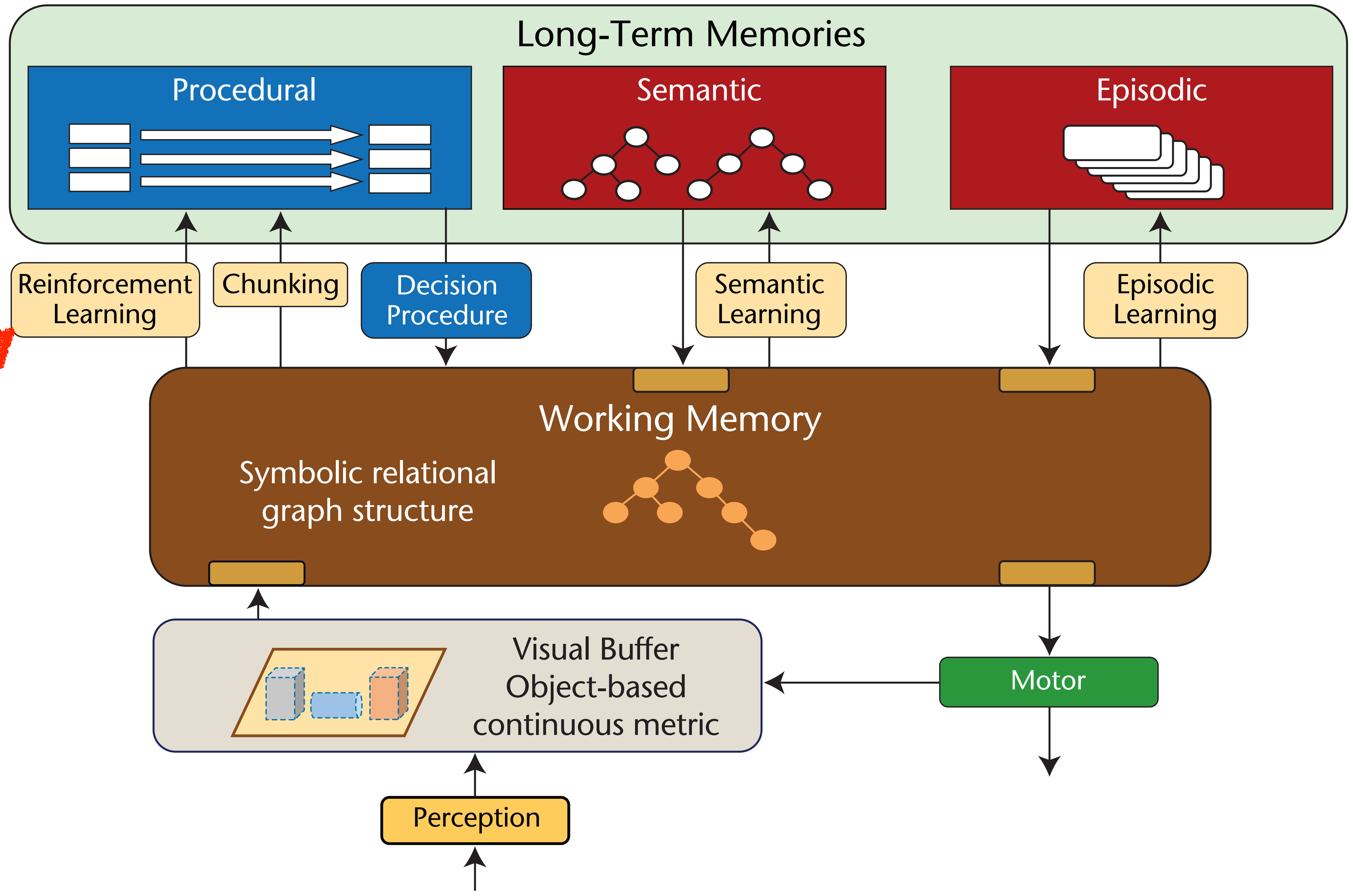
# AI is still uneasy with reward, but is coming around




- Early problem-solving AI, and even the *latest edition* of the standard AI textbook, define goals as world states to reach, not experience
    - But it also has chapters on reinforcement learning, using reward
- With the rise of machine learning in AI, the reward formulation of goals is becoming standard
    - For example, Markov decision processes are now one standard way of formulating planning in AI
    - Reward is “the cherry on top of the cake of intelligence” (Yann LeCun)

# The Soar cognitive architecture now includes reward

- Soar is classic GOFAI
  (1980s, Newell, Laird, Rosenbloom…)
  - Production rules, symbols
- Since 2008 it has included a form of reward and reinforcement learning




—Laird, Lebiere & Rosenbloom. *A Standard Model of the Mind*, AI Magazine 2017

# 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

An interlude:

# Introduction to Experience

# Experience — a concrete nonspecific example

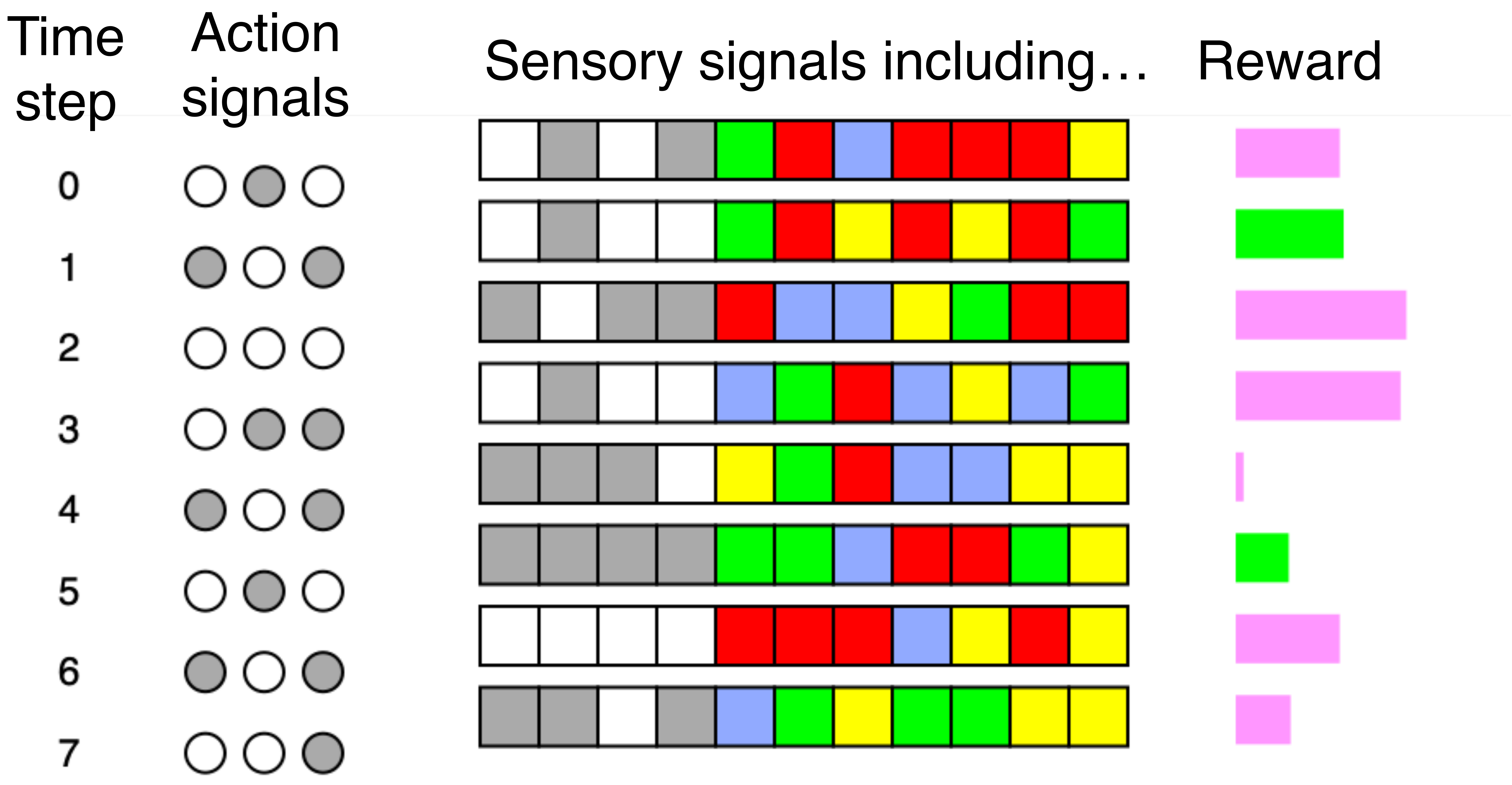
| Time step | Action signals | Sensory signals including… | Reward |
|---|---|---|---|
| | | 1 0 1 0 1 0 2 0 0 0 3 | -5.3 |
| 0 | 1 0 1 | 1 0 1 1 1 0 3 0 3 0 1 | 5.5 |
| 1 | 0 1 0 | 0 1 0 0 0 2 2 3 1 0 0 | -8.7 |
| 2 | 1 1 1 | 1 0 1 1 2 1 0 2 3 2 1 | -8.4 |
| 3 | 1 0 0 | 0 0 0 1 3 1 0 2 2 3 3 | -0.4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

most recent action

most recent sensation

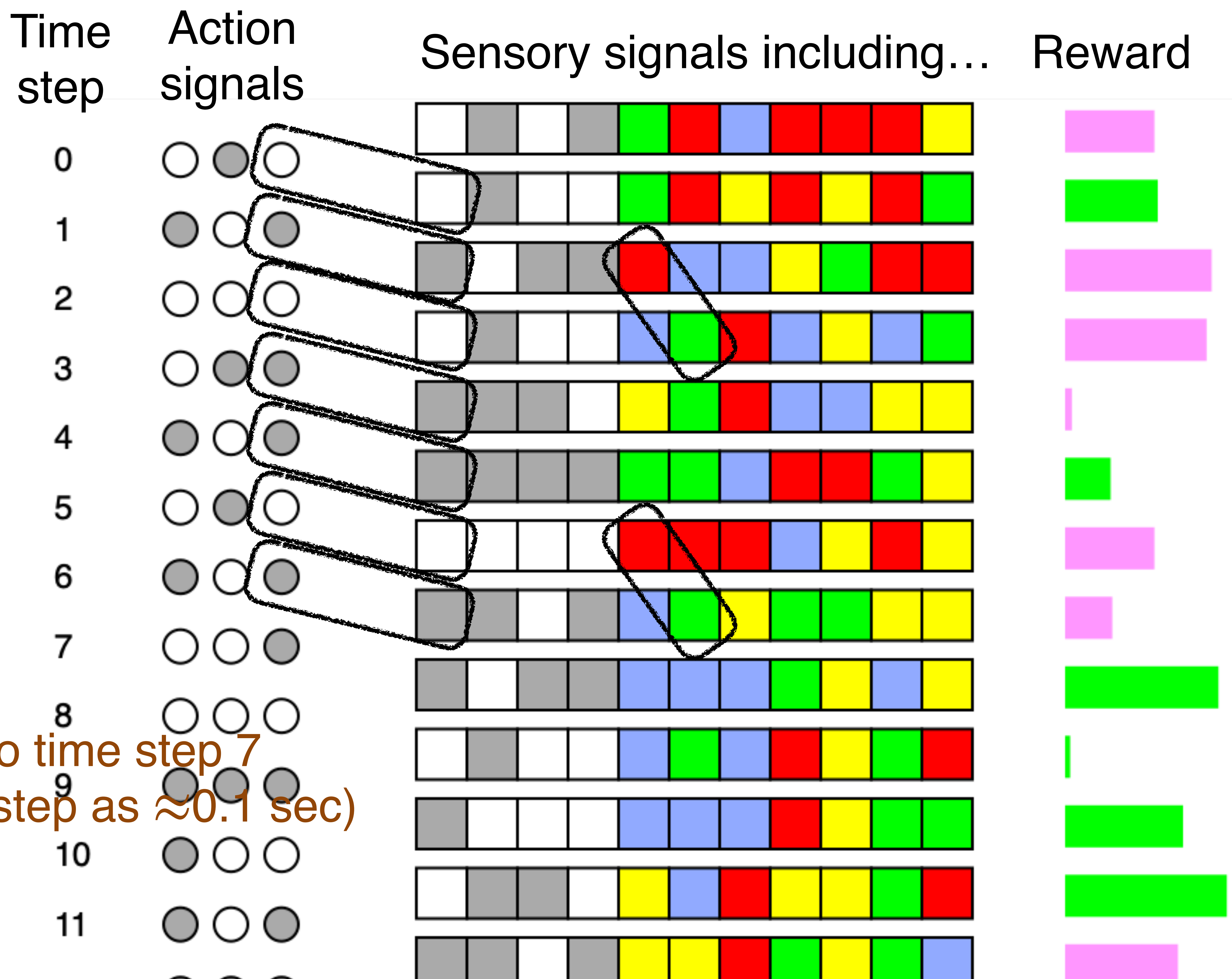
Experience up to time step 7
(think of a time step as ≈0.1 sec)

# Experience — a concrete nonspecific example



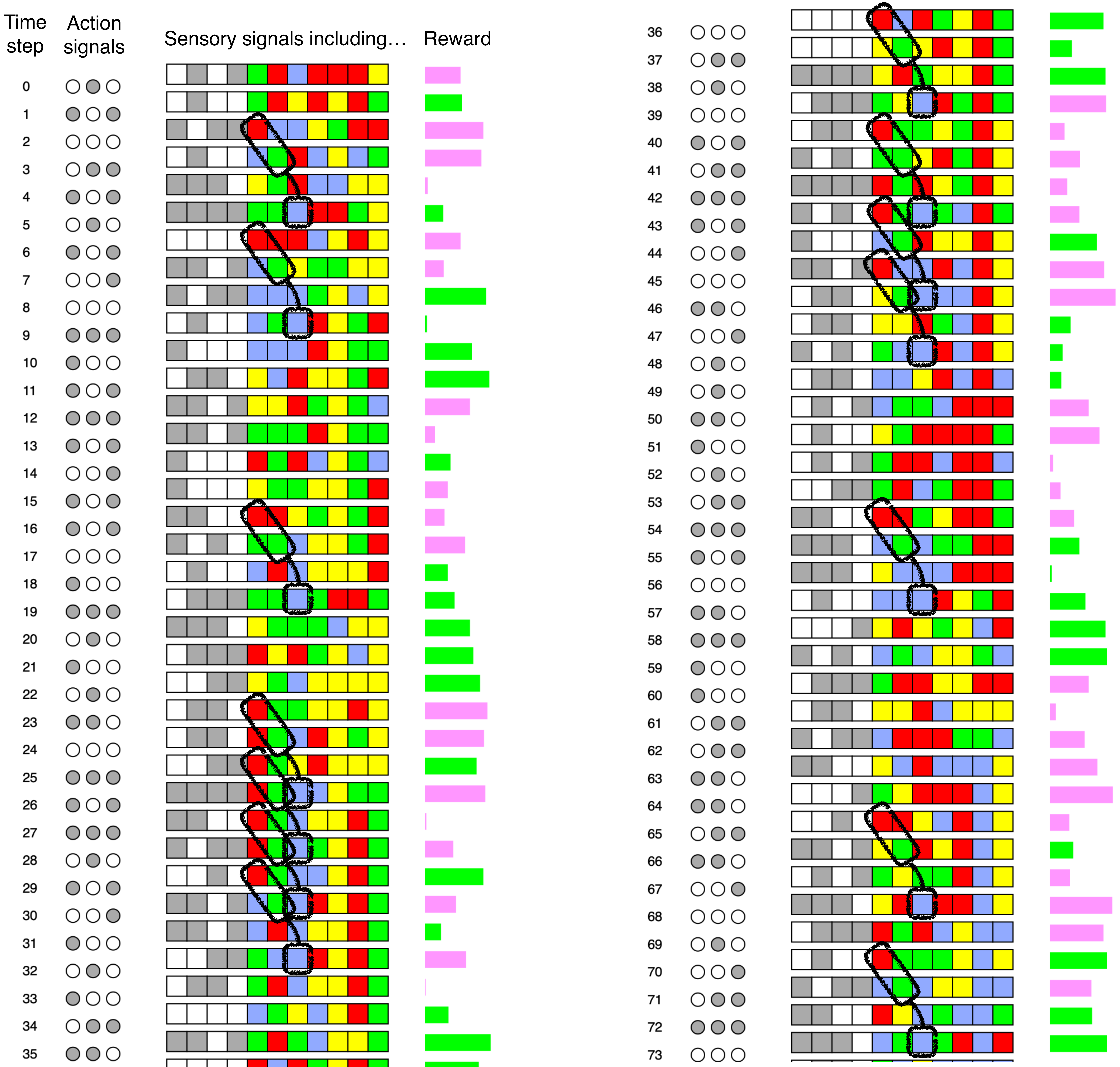

Experience up to time step 7
(think of a time step as ≈0.1 sec)

# Experience — a concrete nonspecific example

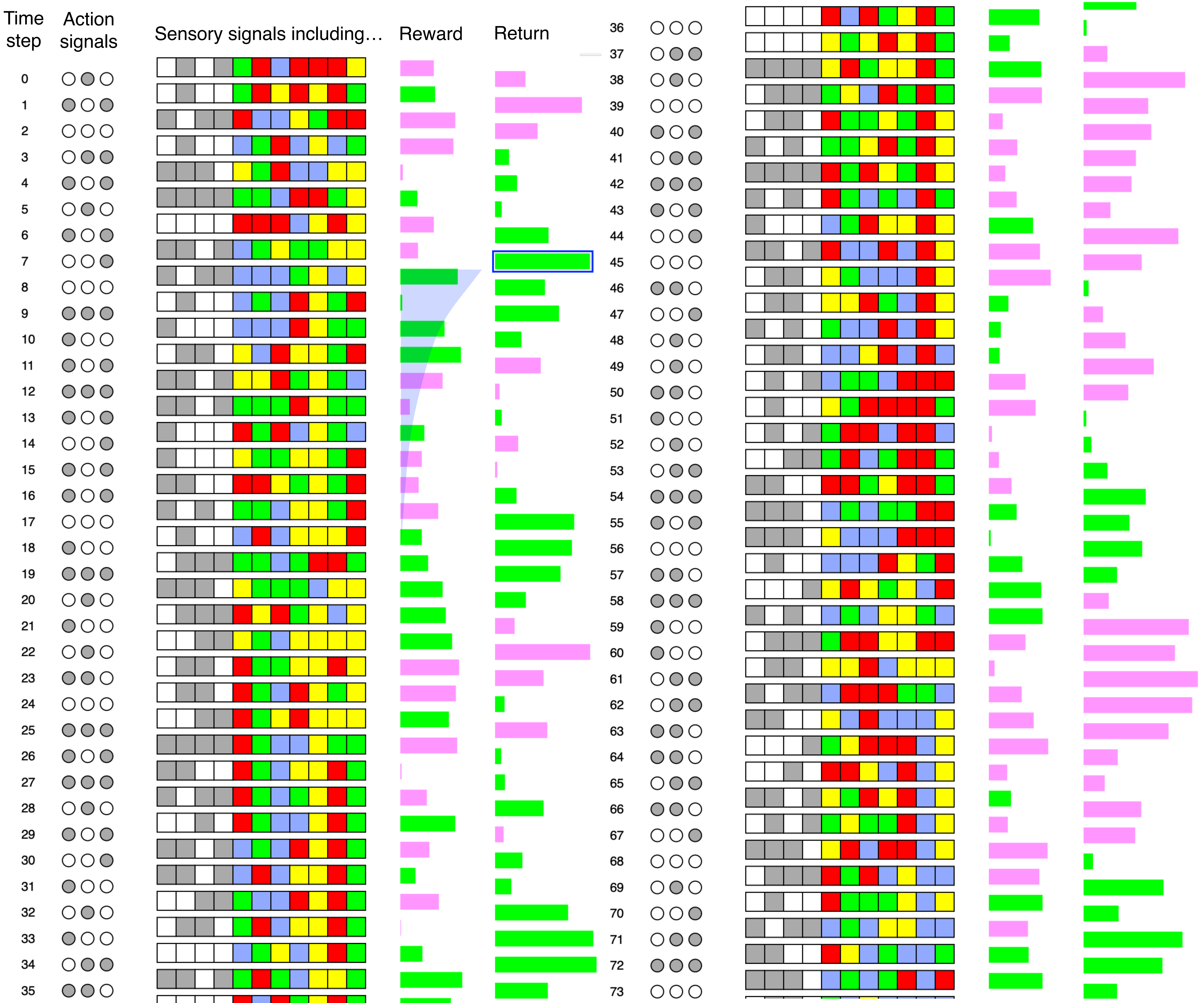



Experience up to time step 7
(think of a time step as ≈0.1 sec)

- 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



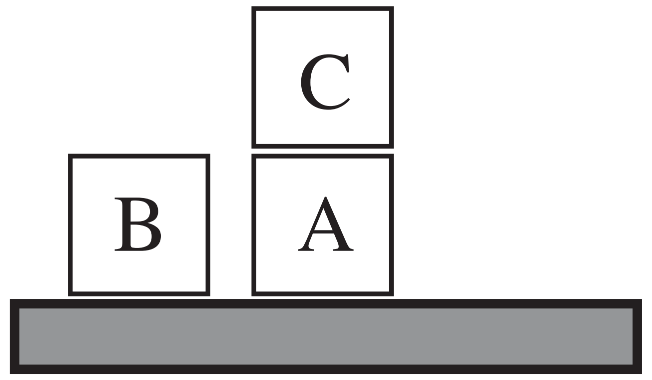
- Different sensory signals can be *qualitatively different* from each other
  - In their range of values
  - In their predictive relationships
    - to action signals
    - to each other
    - to themselves
- There are short-term *and* long-term patterns in these data
- There are many things to predict
- 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



# Step 3: Experiential state

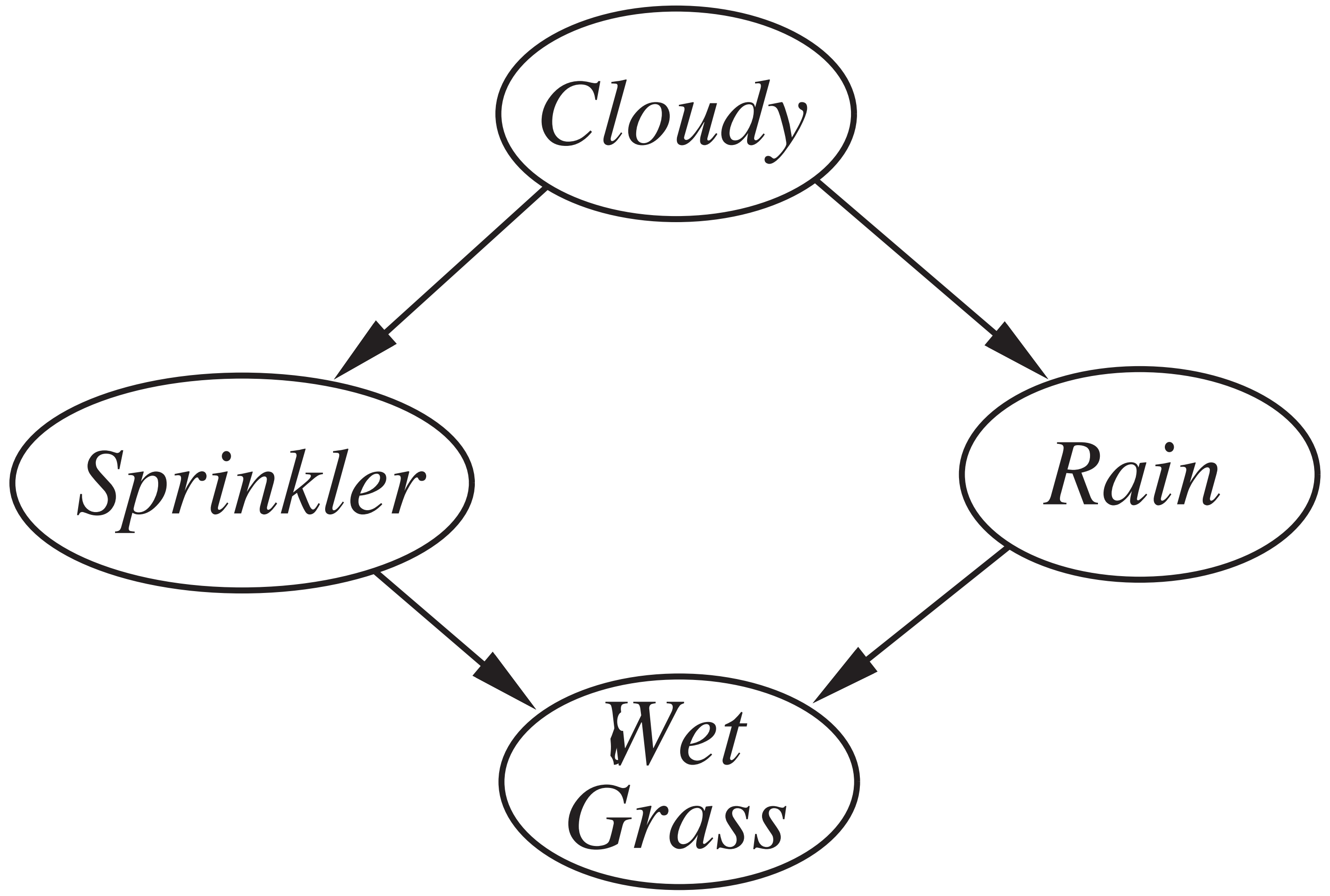
(state in terms of experience)

# Conventionally in AI, state has been characterized in terms of the external world (objective state)




- Classically, perception produced symbolic propositions whose truth values were assumed to match the world, e.g.,

  *On*(*BlockC*, *BlockA*), *Loves* (*John*, *Mary*)




- In probabilistic graphical models,
  state is a probability distribution over world state variables
- In POMDPs (Partially observable Markov decision processes) state is a probability distribution over underlying discrete world states (belief state)
- Such *objective state* representations are far from experience

—Figures from Russell & Norvig, 2008

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*

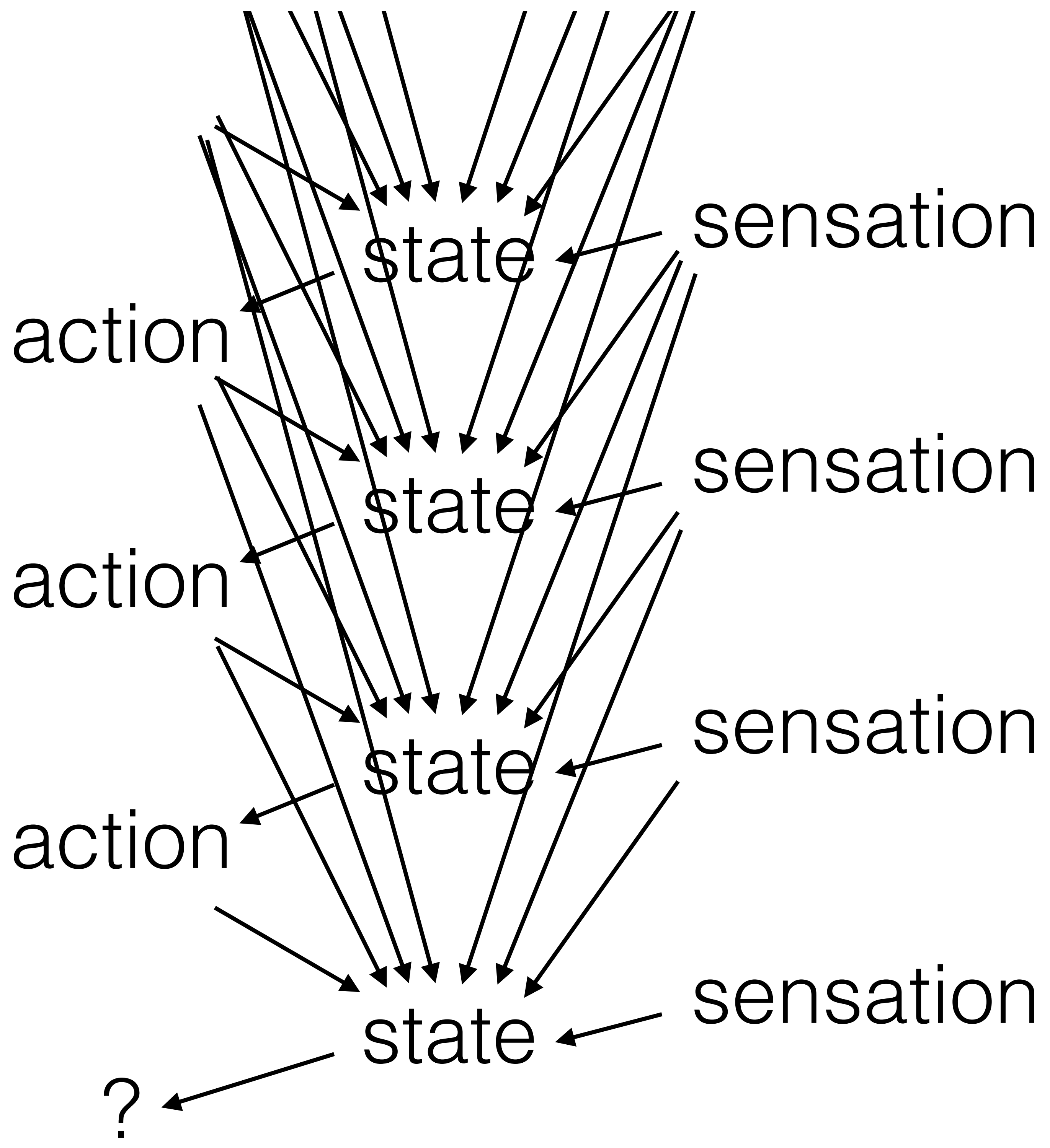
No mention of external entities “out there” in the world

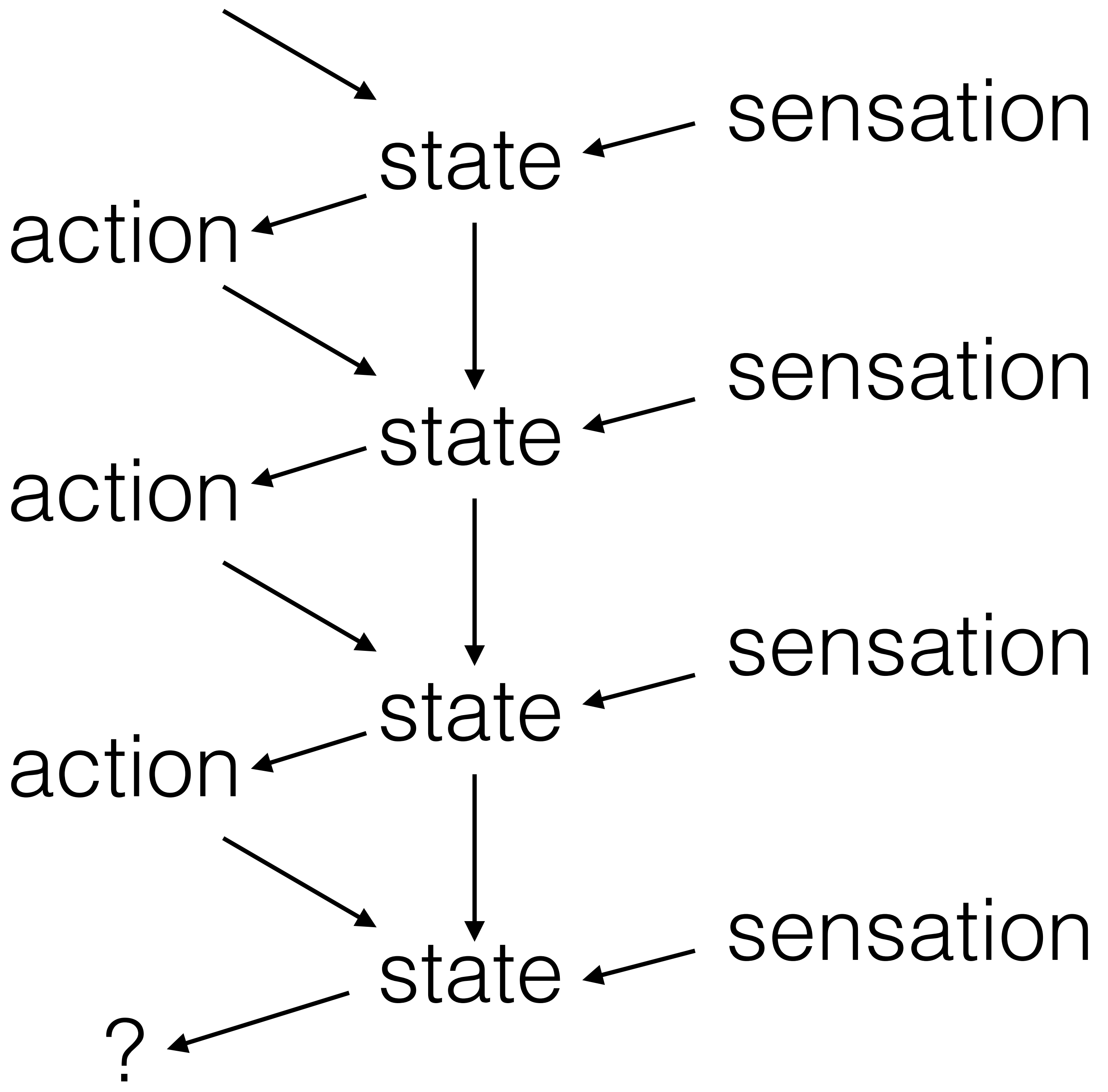
# Some modern AI embraces experiential state

- Most commonly it is simply build in, e.g.,
  - the last four video frames of Atari video input to DQN
  - including one or more recent actions
- Compression approaches to AI
- LSTMs in deep learning
- Predictive State Representations, Spectral methods

Such approaches learn (or discover) their experiential state

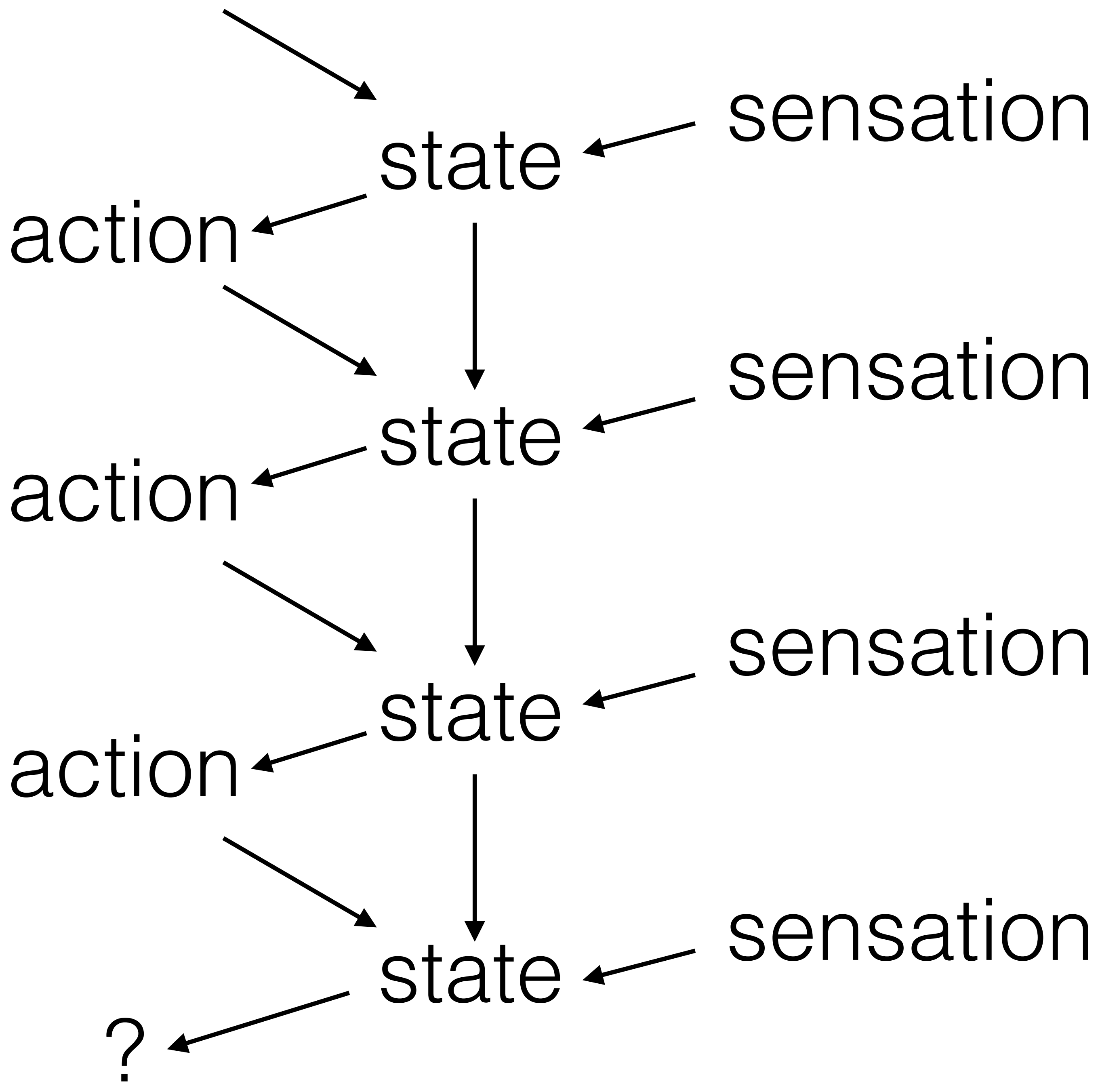
# Experiential state should be recursively updated

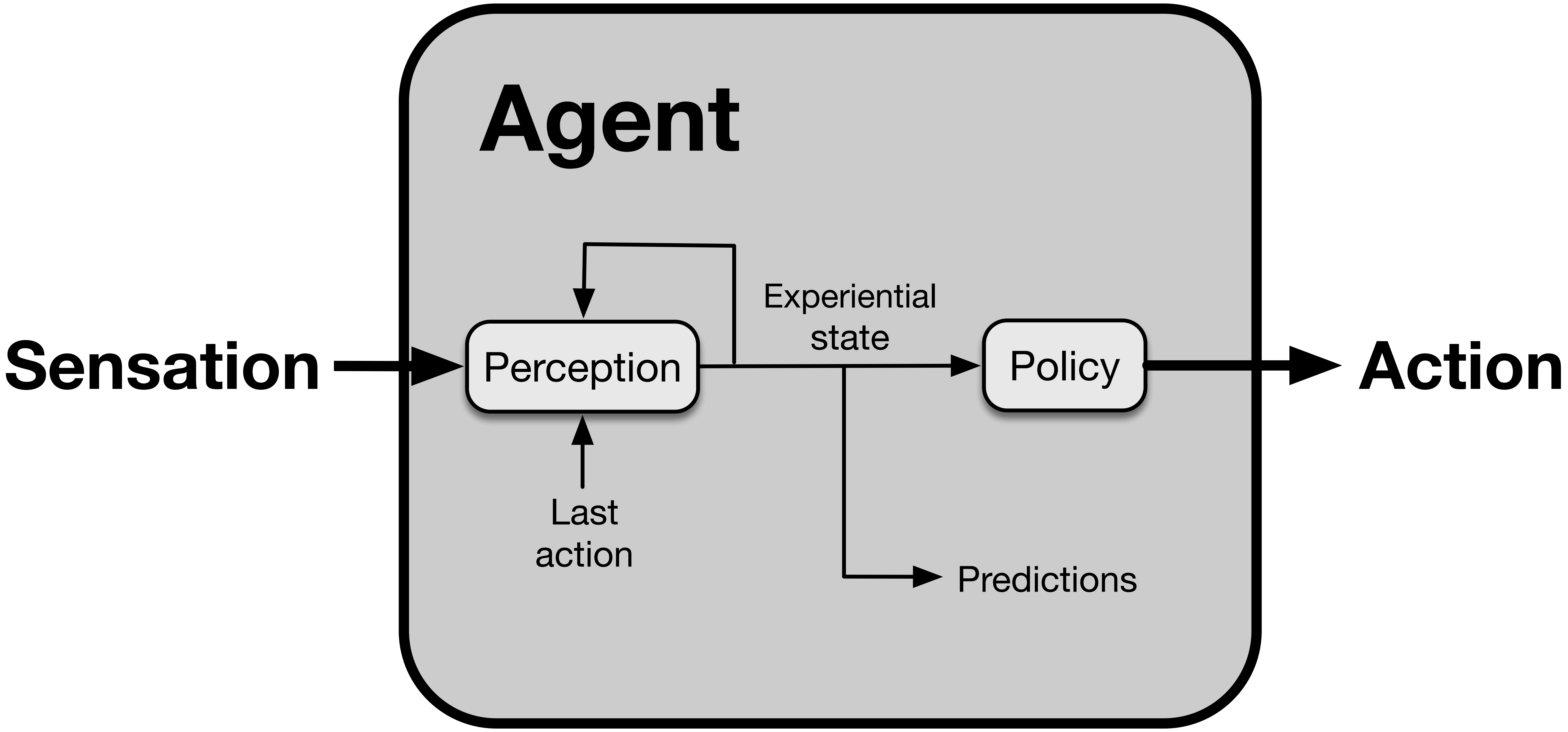



*Experiential state is a summary of past experience*
*that is useful for predicting and controlling future experience*

# 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 *a standard (basic) model of the experiential agent*

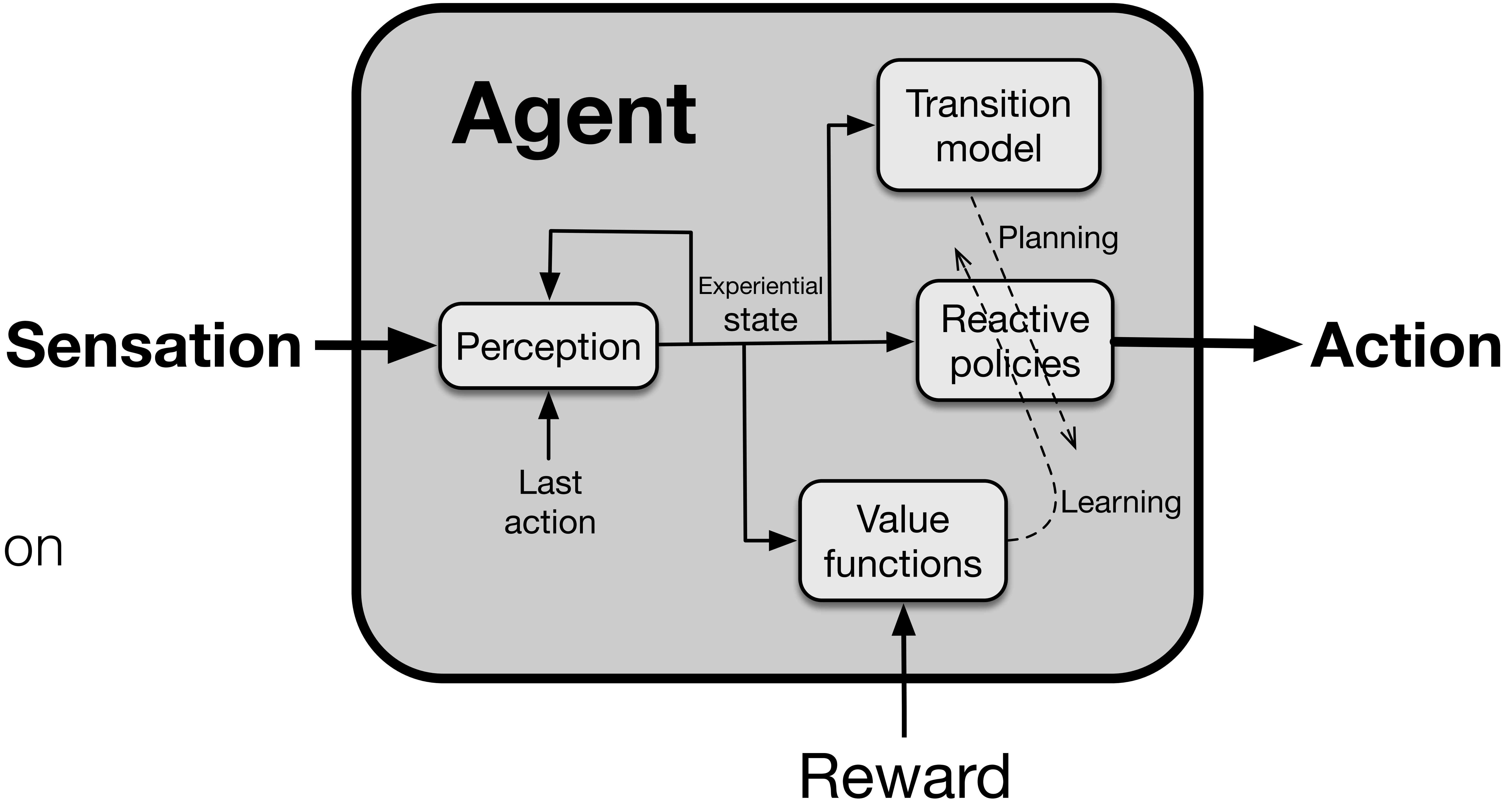
Step 1: **Agenthood**
(sensation & action)

Step 2: **Reward**

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

Step 4: **Predictive knowledge**

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



# Step 4: Predictive knowledge

(to know is to predict experience)

# Much world knowledge seems to be about the external world independent of 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

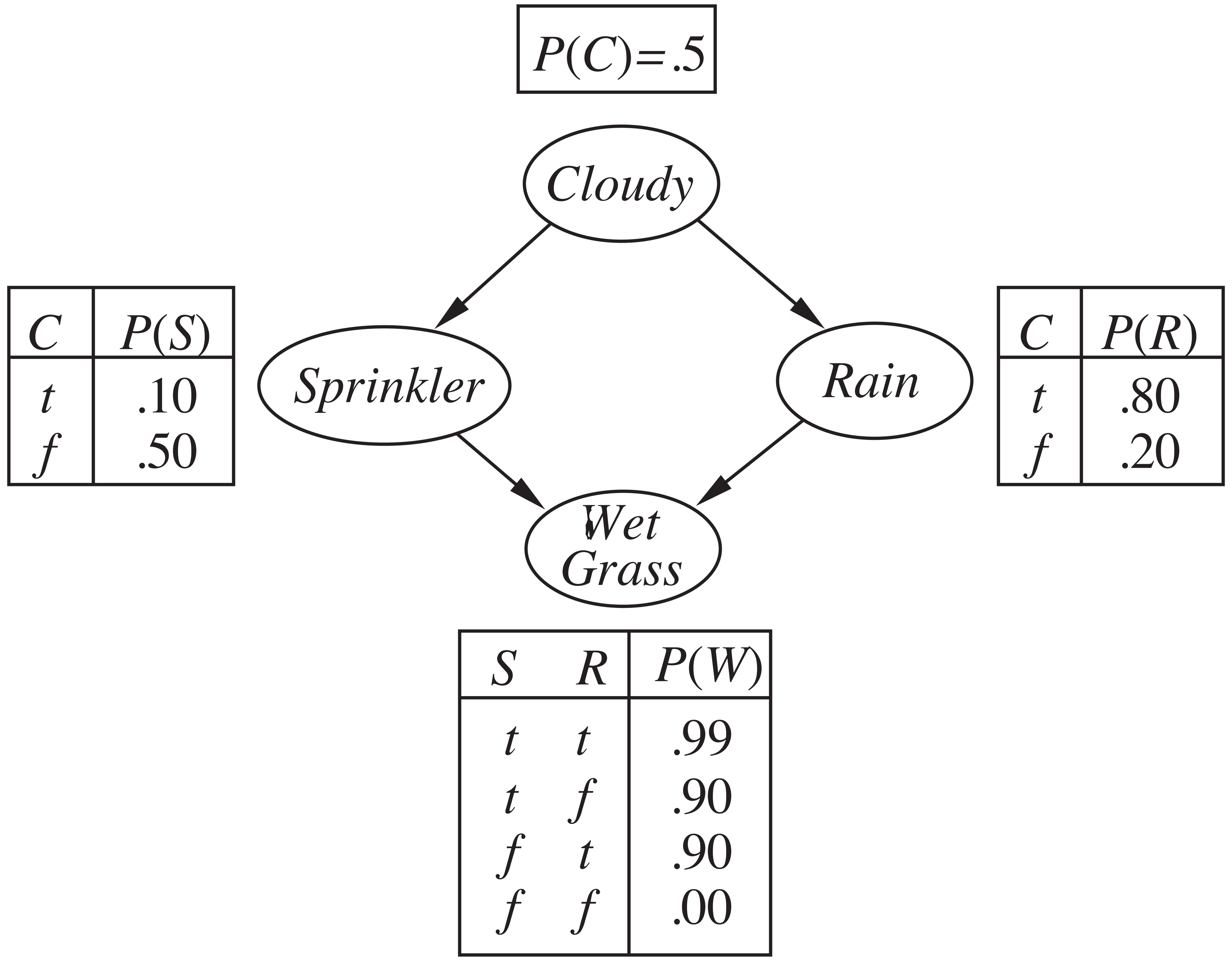
## Other knowledge seems more like predictions of experience

- 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

# Knowledge is becoming more predictive

*President*(*US*) = *Joe_Biden*

*Capital_of*(*France*) = *Paris*




—Figure from Russell & Norvig, 2008

- Early AI systems, lacking experience, could not predict
- Much modern AI still treats knowledge as database entries
- Much modern AI (e.g., probabilistic graphical models) has knowledge only about *simultaneous* events
- Prediction of sequential events is a kind of knowledge with a clear semantics
- A *predictive model of the world* is AI's upcoming new view of world knowledge
- The cutting edge of predictive knowledge (IMO) is *general value functions* (GVFs) and *option models*

# Types of 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)
  - predictive knowledge about *state transitions*, i.e., a predictive model of the world

# A state-to-state predictive model need not be low level

- A model need not be differential equations or a Markov decision processes
- A model can be abstract in state (e.g., experiential state)
- A 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, can be learned off-policy
- Are there limits to the expressiveness of option models and experiential state?
- Can we bridge the abstraction gap between experience and knowledge?

# Experience is fundamental to world knowledge

- By definition,

  we (agents) gain information about the world only thru our sensors

  and we affect the world only thru our actions

⇒ We know the world only through our experience

  ⇒ Everything we know about the world is a fact about our experience

- This perspective seems inescapable to me…

  and in the long run it is good for the science of AI

- But… still we don’t *like* to think about experience

# Why we dislike experience

- Experience is unfamiliar, strange, unintuitive, temporal, complex, and so darn low-level

- Experience is hard to talk about — subjective and private, impossible to communicate to others or to be verified by others

- Public, external terms are clearly superior to experiential terms for everything humans do… *except perhaps for creating AI*

# Why we should like experience

- Experience comes from the ordinary operation of the AI; it is “free” data; it enables autonomous learning that scales with computation
- Experience offers a path to knowing the world:
  - If any fact about the world is a fact about experience, then it can be learned and verified from experience

# In summary…

- I have discussed four major steps in the increasing role of sensorimotor experience in AI:

  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, I have shown
  - That AI has chosen first to work in objective, non-experiential terms
  - But there is less-familiar approach, based on experience, growing in importance, with advantages in *grounding*, *learnability*, and *scaling*
- The trend toward experience in AI may have much further to go
  - Steps 3 and 4 are far from complete; there are research opportunities
- 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!