Grounding Artificial Cognition

Rich Sutton
Reinforcement Learning and Artificial Intelligence Laboratory
University of Alberta

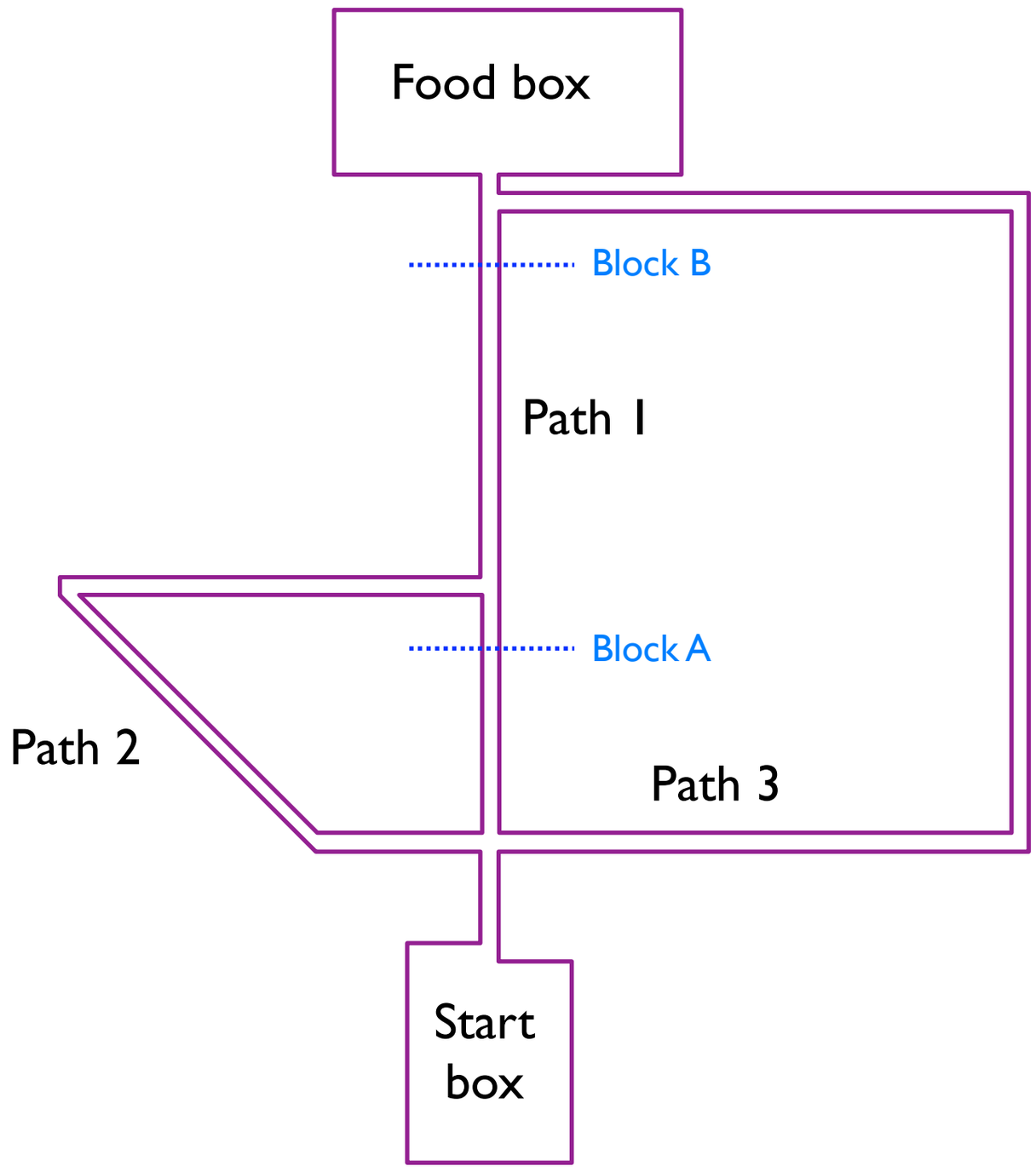
Reinforcement learning

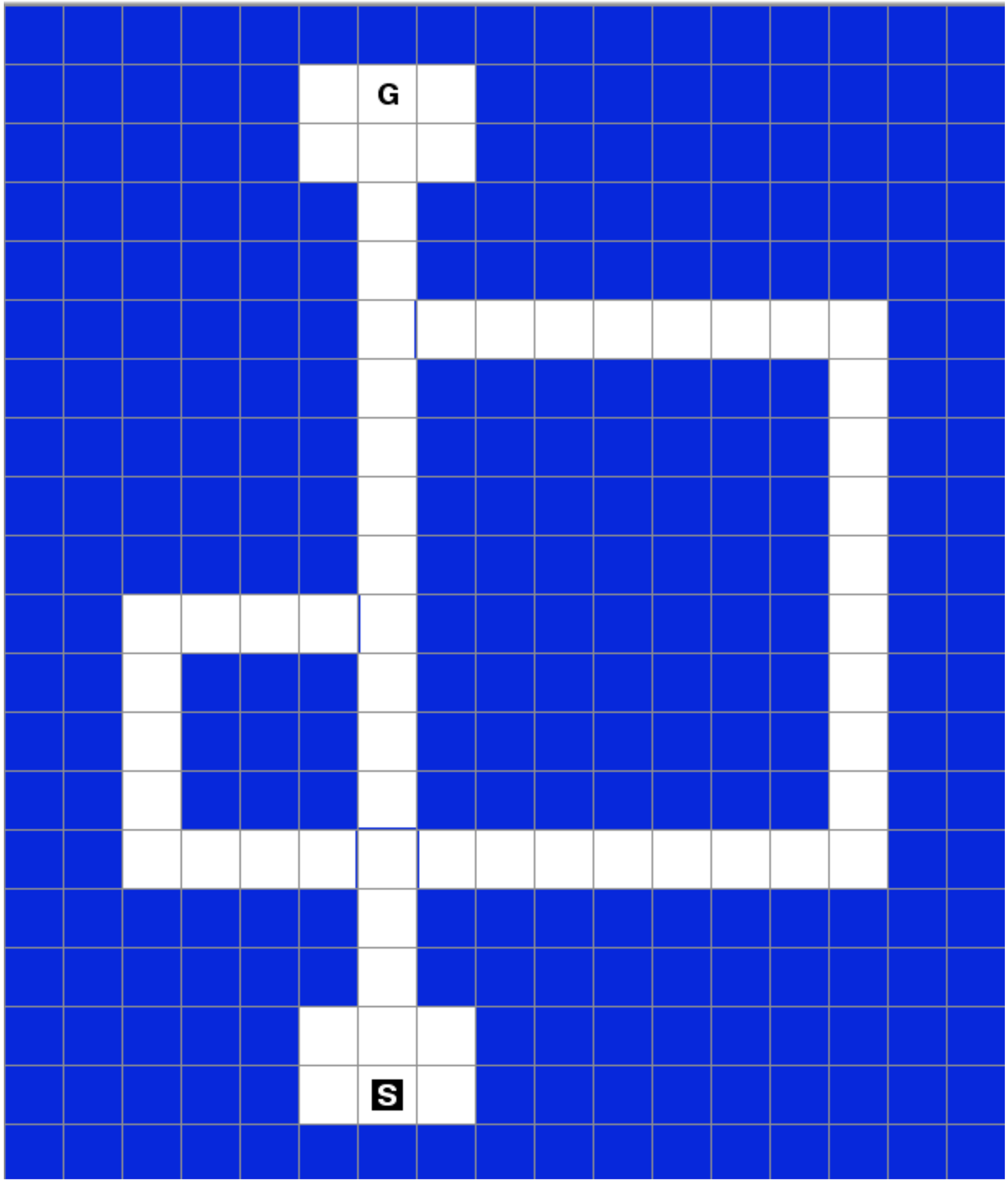
- A computational approach to learning in action selection, valuation, prediction, planning...
- Popular and successful in artificial intelligence, 1990–
  - Compact theory, impressive applications
- Standard model of dopamine reward system, 1997–
- Influential model of Pavlovian conditioning, 1981–
- Trendy topic in operations research, 1996–

Things to expect in this talk

- An extreme position on grounding knowledge in low-level experience
- A concern with abstraction (grand challenge)
- An AI perspective
- Big ideas, small systems
- Some intuition-building demos
- The PEAK-complete hypothesis
- PEAK =
  Predictive Empirical Abstract Knowledge

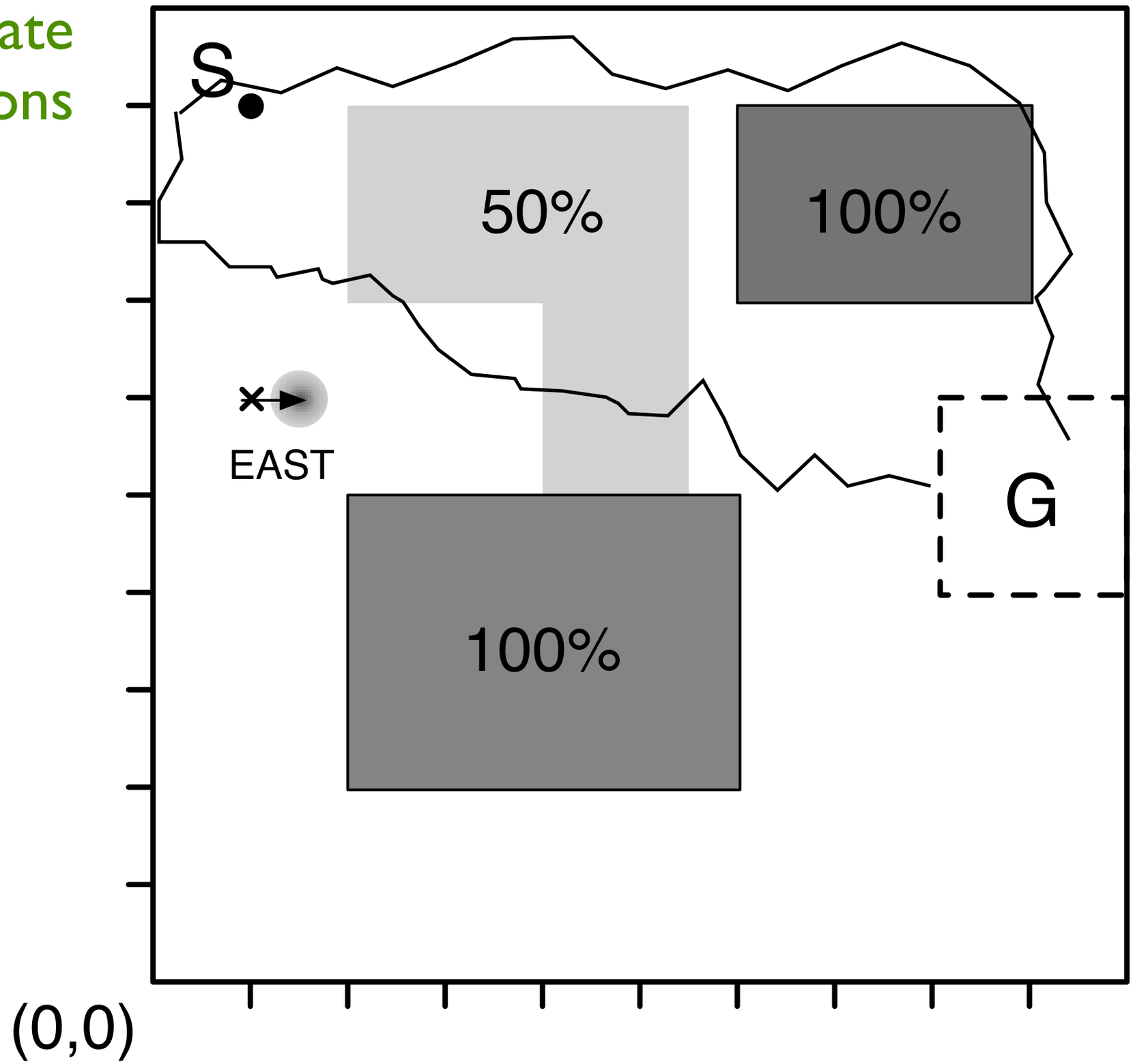
Tolman & Honzik, 1930
"Reasoning in Rats"



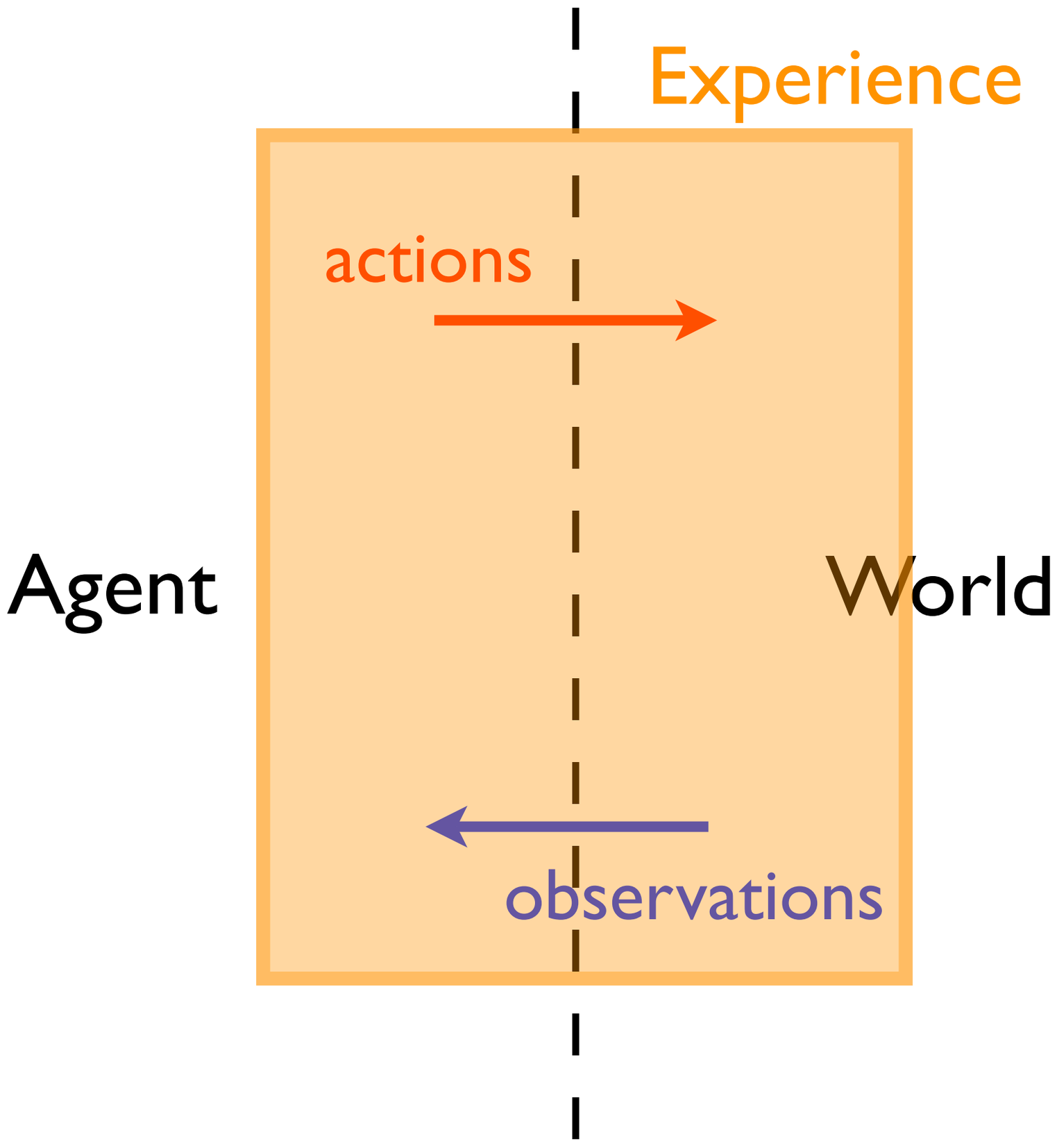


Dyna arch.
Sutton, 1990

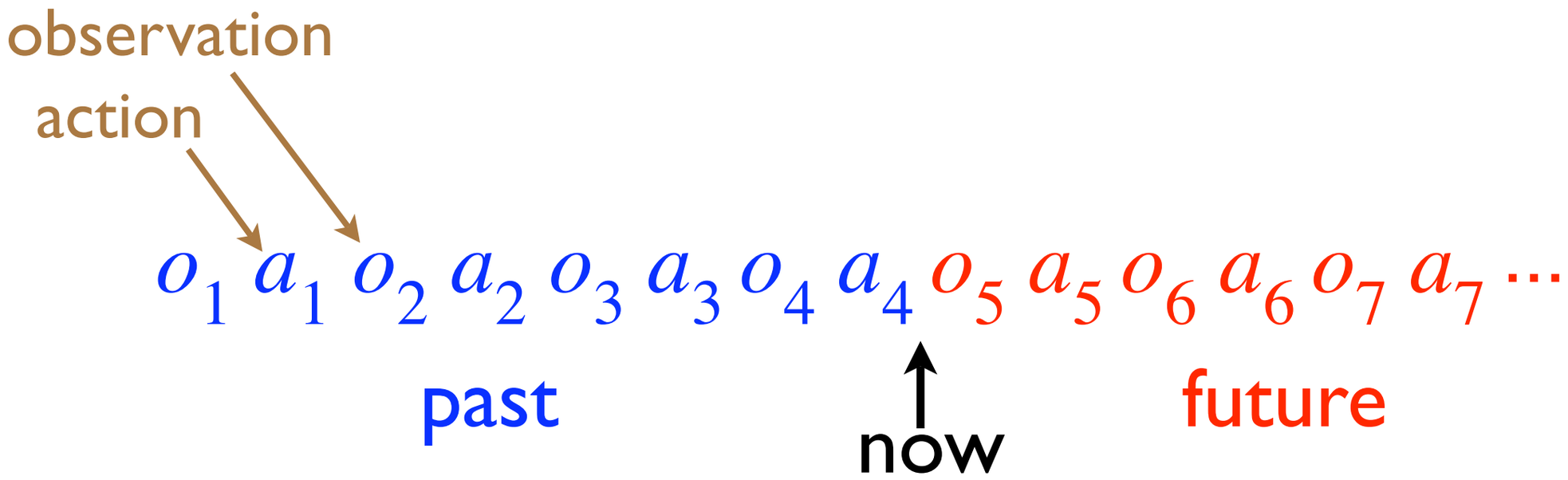
Continuous state
Probabilistic actions



Cosmin Paduraru, 2006


Experience
actions
Agent
World
observations

Experience is the data of the mind



low-level sensori-motor experience, e.g., 100 Hz

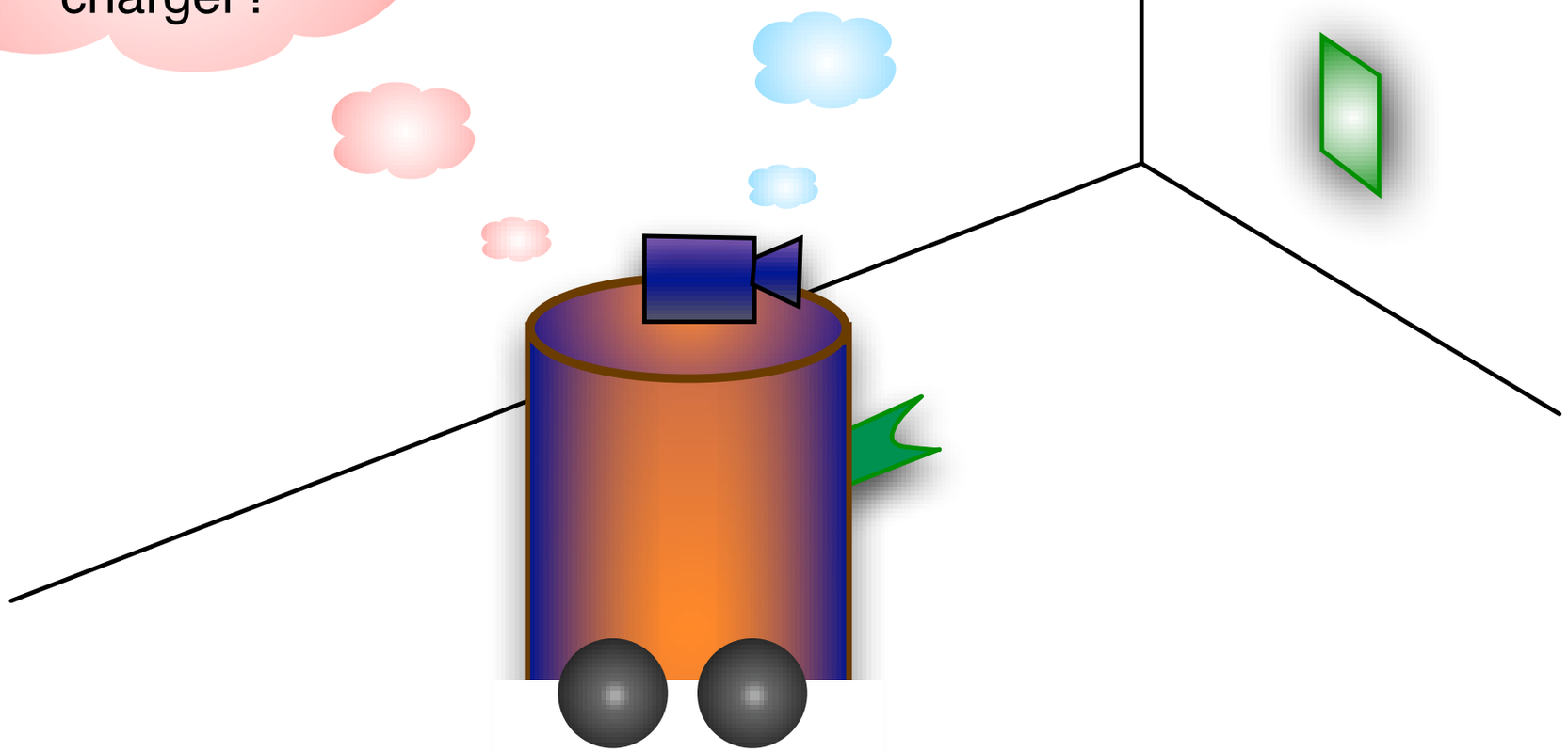
Grounding, verification are key to AI

semantics grounded in human judgments

Would a person say that's a battery charger?

Would `try-to-dock` succeed?

semantics grounded in prediction



Experience matters

- Experience is the most prominant feature of the computational problem we call AI
- It's the central data structure
- It has a definite temporal structure
  - revealed and chosen over time
  - speed of decision is important
  - order is important
- This has unavoidable implications for AI

Experience in AI

Many, many AI systems have no experience

They don't have a life!

- Expert Systems
- Knowledge bases like CYC
- Question-answering systems
- Puzzle solvers,
  - or any planner that is designed to receive problem descriptions and emit solutions

Part of the new popularity of agent-oriented AI is that it highlights experience

Other AI systems have experience, but don't focus on it

A new empiricism

- Knowledge is about experience
  - knowledge is prediction of experience
  - enables verification, self-maintenance
- But not necessarily from experience
  - takes no position on the nature/nurture debate
- Does not require public agreement
  - specific to agent's sensory and motor capabilities
  - subjective is primary; objective secondary

Philosophical and Psychological Roots

- Like classical british empiricism (1650–1800)
- Like logical positivism (Ayer, Peirce)
- But not anti-nativist, not tabula rasa
- Subjective rather than objective
- Emphasizing sequential rather than simultaneous events
- Close to Tolman's "Expectancy Theory" (1932–1950)
  - Cognitive maps, vicarious trial and error

A Grand Challenge:
Grounding knowledge in experience

- To represent human-level world knowledge solely in terms of lowest-level experience
  - sensations
  - actions
  - time
- A minimal ontology
  - no state
  - no objects, no people, no space, no self, no chickens...
  - all these are "just" patterns in sensation & action

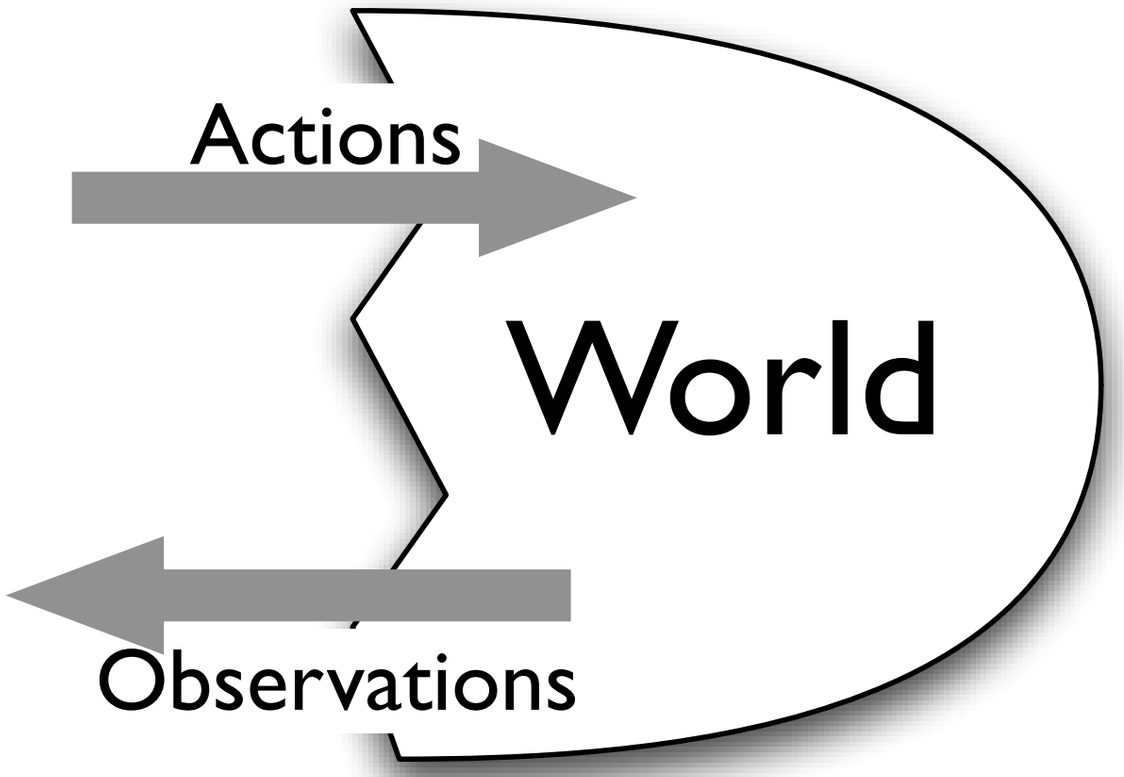
World knowledge must be predictions

actions $a_t \in \mathbf{A}$

observations $o_t \in \mathbf{O}$

experience $e_t \in \{\mathbf{O} \times \mathbf{A}\}^t$

The world is completely described by the probability distribution



$\omega(o \mid e) = \text{Prob}(o_{t+1} = o \mid e_t = e)$

To know something about the world at time t

is to know something about $\omega(o \mid e_t e)$ for $e \in \{\mathbf{O} \times \mathbf{A}\}^*$

There is nothing else to know

What would it be like to accept the challenge?

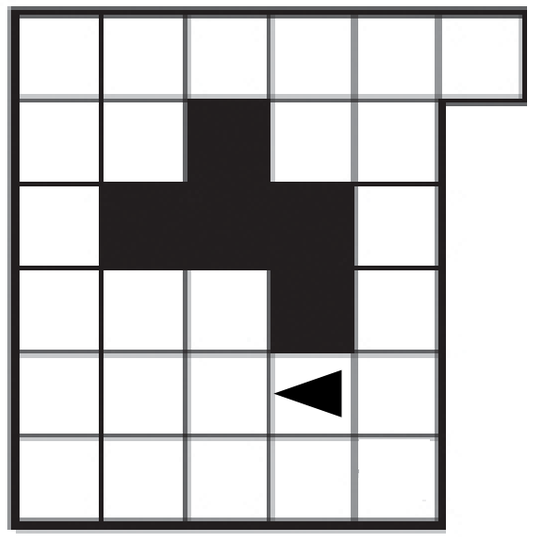
- Abstraction is key
  - abstract states (eg, predictive representations)
  - abstract actions/transitions (eg, options)
- Need to think in unfamiliar ways
- Microworlds, robotics
- Indexical (deictic) representations
  - sequence instead of symbols

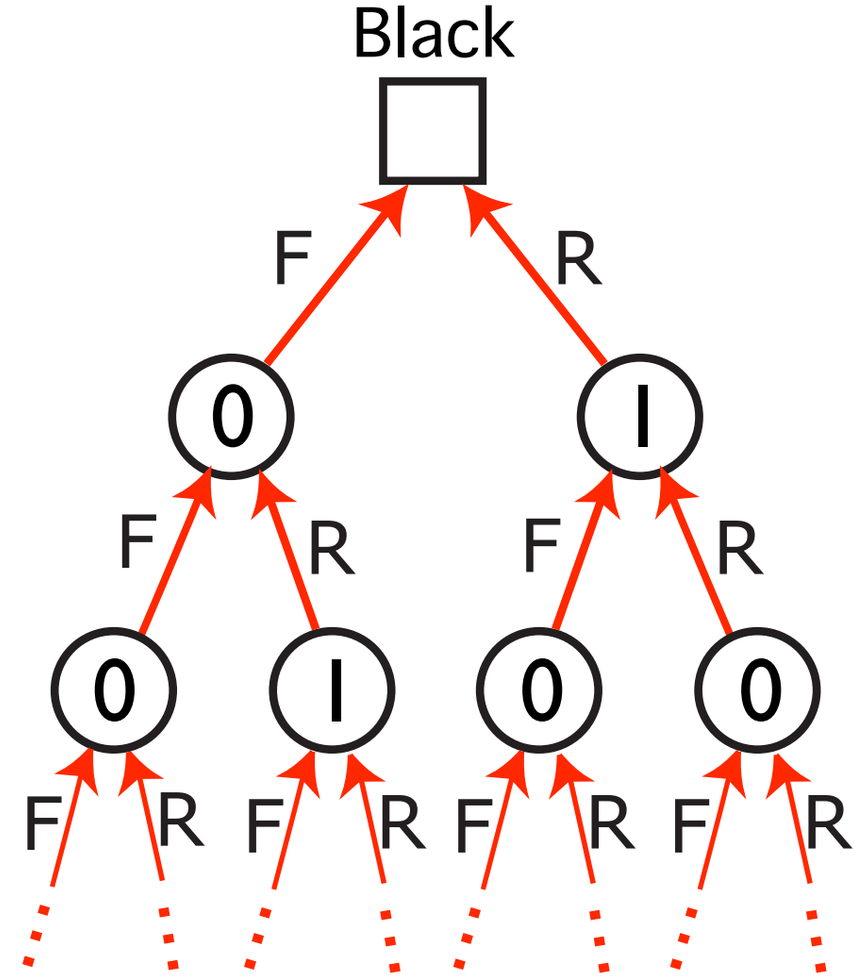
In experiential terms,

- What is space?
  - regularities in sensation change with eye movement
- What are objects?
  - subsets of sensations
  - that tend to occur together temporally
  - and can be in arbitrary relative spatial arrangements

- What is my body, my hands?
  - objects that are always present
  - and can be controlled
- What are people?
  - objects that may move on their own
  - that have a particular subset of sensations
  - whose presence may change my sensations for the better
  - eventually:
    - ✦ that are best predicted with respect to goals
    - ✦ that are analogous to me

Bit-to-bit gridworld





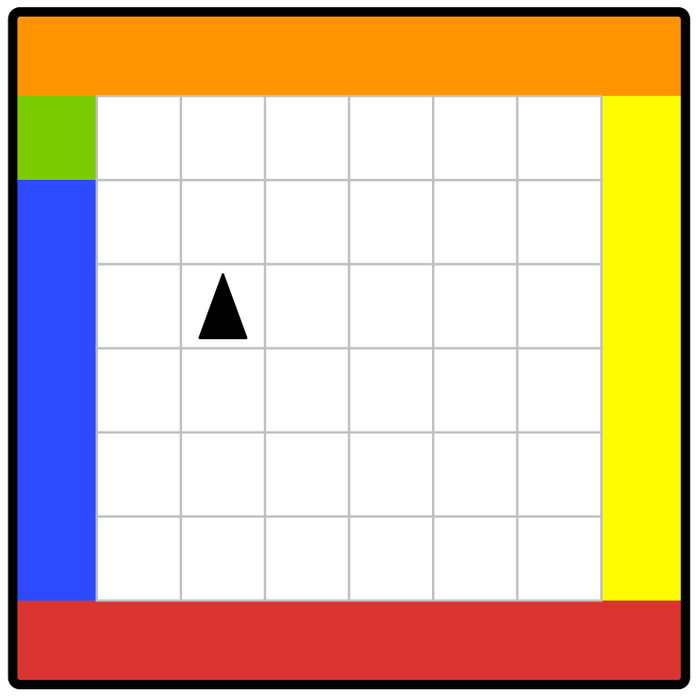
sensations: black/white ahead
actions: F(orward), R(ight)

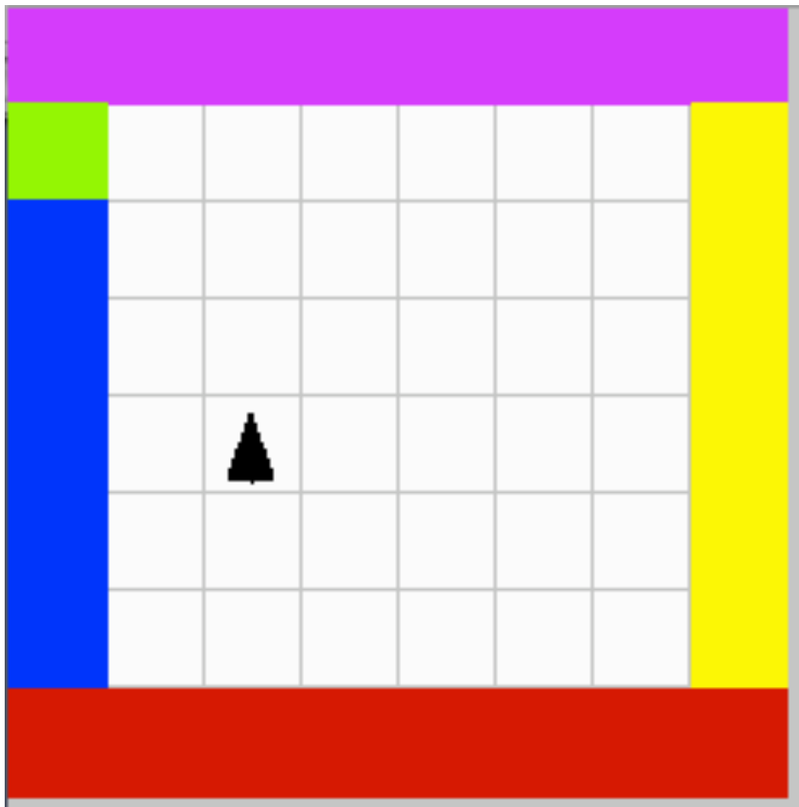
one bit each of sensation and action!

Brian Tanner, 2005

Compass world

- sensation: color ahead
- actions:
  - L(eft)
  - R(ight)
  - F(orward)
- options:
  - Leap (to wall)
  - Wander (randomly)



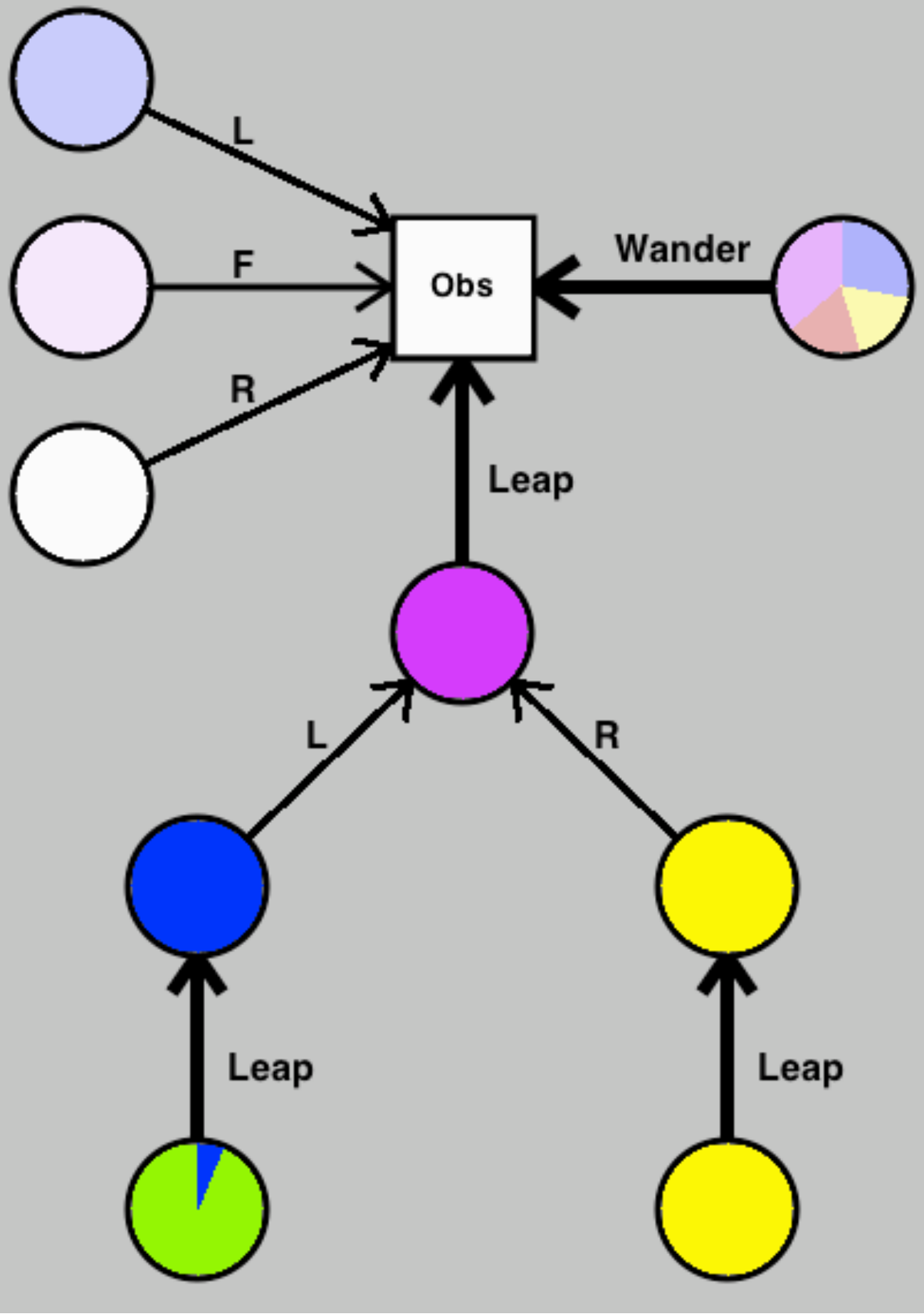


Step Forward

Turn Left Turn Right

Wander 20 steps

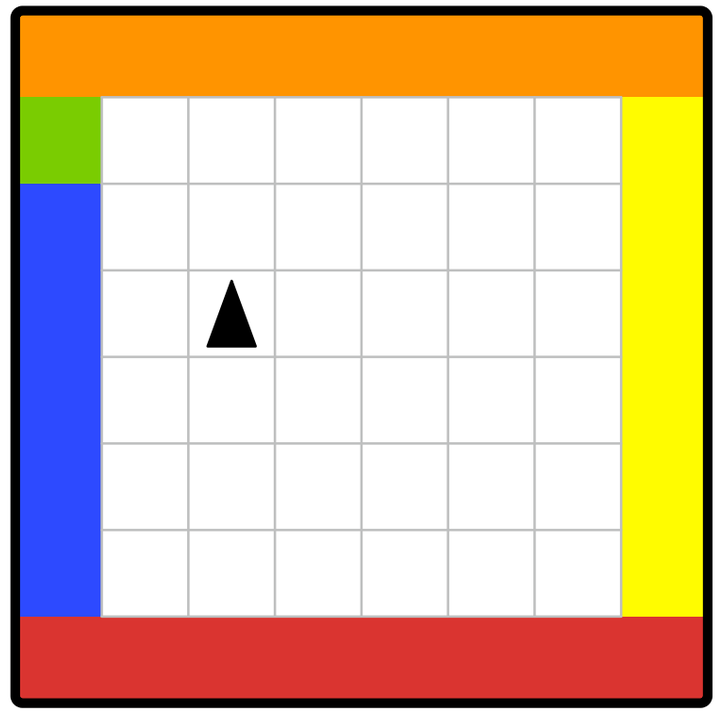
Time step = 200003



Examples in compass world

If I were to...

...step forward till I hit a wall,
would it be orange?



"facing an orange wall"

not compositional

If I were to...

...step forward till I hit a wall, then turn right,
would I be "facing a yellow wall?"

compositional

Sutton, Precup & Singh, 1999

Key machinery 1: options

- options are a generalization of actions
  - a way of behaving (policy), $\pi : \mathbf{S} \times \mathbf{A} \rightarrow [0,1]$
  - a way of stopping (term. cond.), $\beta : \mathbf{S} \rightarrow [0,1]$
- for the robot and the battery charger:
  - behave according to some try-to-dock policy
  - stop when docked or timed out

Why options?

- they are very simple and general
  - a minimalist, least-commitment form of macro-action
  - allow arbitrary closed-loop policies
  - support action-independent temporal abstraction
- they are compatible with planning methods based on dynamic programming

Key machinery 2: option models

- an option model is a prediction of the option's outcome
  - what state you will end up in: $p : \mathbf{S} \times \mathbf{S} \rightarrow [0,1]$
  - how much reward you'll get along the way: $r : \mathbf{S} \rightarrow \mathbf{R}$
- for the robot and the battery charger:
  - will I end up docked?
  - will it hurt along the way, or take a long time?
- These are *subjunctive* predictions – "If I were to..."

Examples of subjunctive, compositional predictions

If I were to...

...follow this hallway to its end,
would I find a restroom?

...look in the fridge,
would I see a beer?

...open the box,
would I see an apple?

...turn over the glass,
would the carpet be wet?

Outcomes are not primitive observations

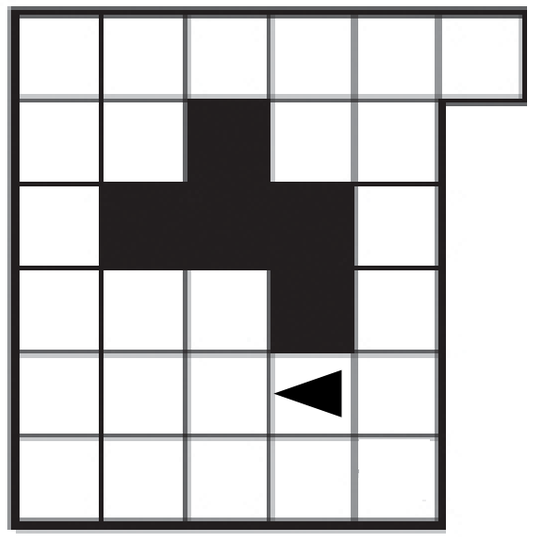
They are sets of predictions

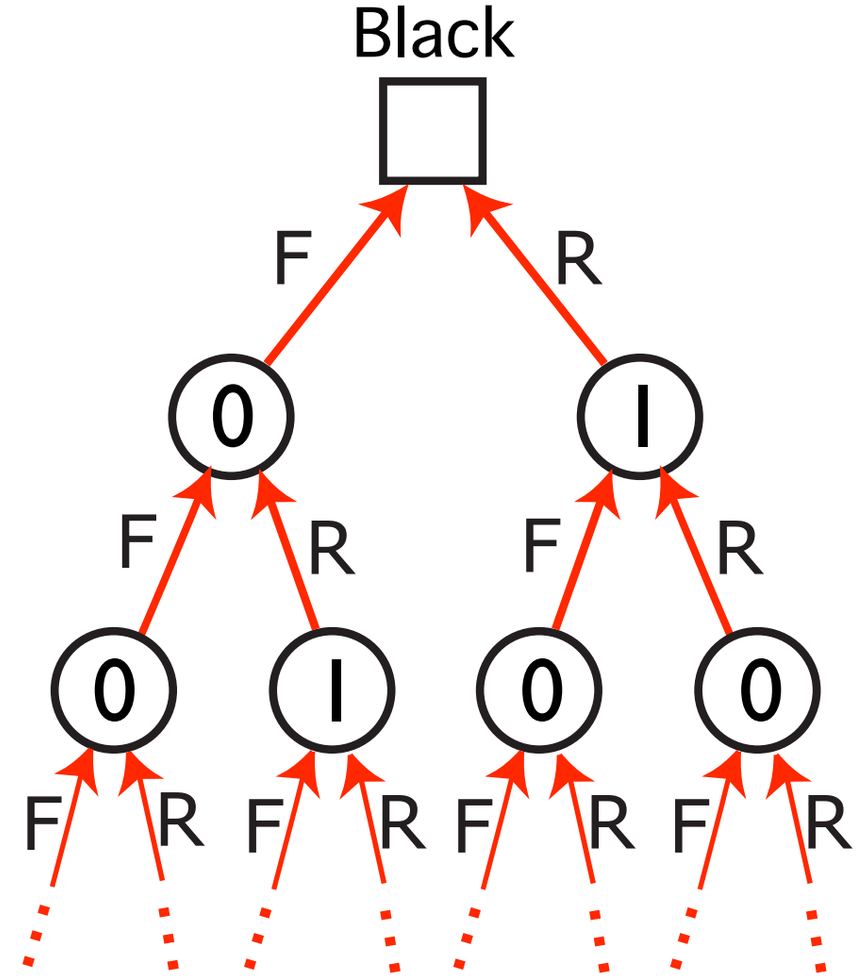
Littman, Sutton, Singh, 2001

Key machinery 3: Predictive representations of state

- Use predictions of option outcomes as state variables
- for the robot and the battery charger:
  - is this a state where try-to-dock will succeed? a.k.a. is there a battery charger here?
  - is this a state where roll-backwards will trigger my bump sensor? a.k.a. is there an obstacle behind me?

Bit-to-bit gridworld





sensations: black/white ahead
actions: F(orward), R(ight)

one bit each of sensation and action!

Brian Tanner, 2005

State is thus exorcised

- State is reduced to predictions of experience
- Option models are usually state to state
- Now they are state variable to state variable
- And the state variables are predictions
  - may be direct predictions of experience
  - or may be predictions of other predictions – *compositionality*

I want to say

"All world knowledge is predictions"

But words fail me

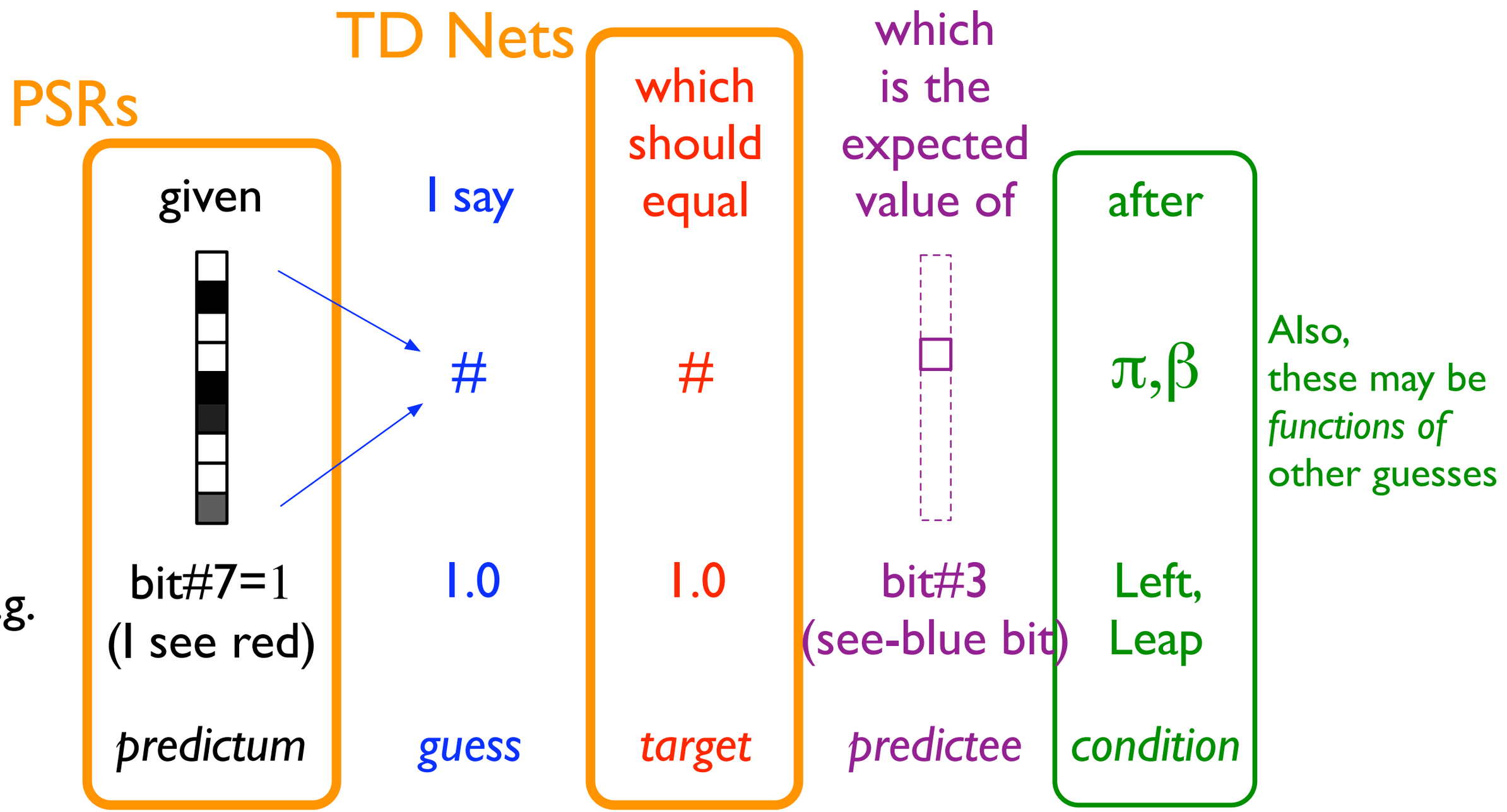
Predictions?

Questions?

Answers?

There are actually *five* different things we need to name.

Anatomy of a prediction



May be the guesses of other, unrelated predictions

PEAK-complete hypothesis:

Predictions suffice to represent all empirical knowledge

Given:

- Predictums can be guesses
- Targets can be guesses
- Conditions can depend on guesses
- Options can be used as conditions
- Exact function approximation

Theory

- Pick a subset of:
  - predictums are guesses
  - targets are guesses

PSR results, learning and discovery

  - option/test conditions
  - off-policy and TD
  - linear function approximation

New: convergence of learning

- Convergence of learning?
- Completeness of discovery?

Steps toward a predictive AI

1. Representation
2. Verification
3. Learning
4. Planning
5. Exploration
6. Discovery
7. Scaling

Take-home messages

- AI should be oriented around experience
  - but it hasn't been
- Knowledge must be predictions
  - but that's nearly unimaginable
- Predictions can be really complex, abstract, expressive and compositional
  - while their machinery is simple and uniform
- Run-time verification may enable big AI
  - although I have shown you just small AI

Thank you for your attention