Rational thoughts in neural codes

Zhengwei Wu (吴争蔚)[a,b,c], Minhae Kwon[a,b,c], Saurabh Daptardar[b,d], Paul Schrater[e,f], and Xaq Pitkow[a,b,c,1]

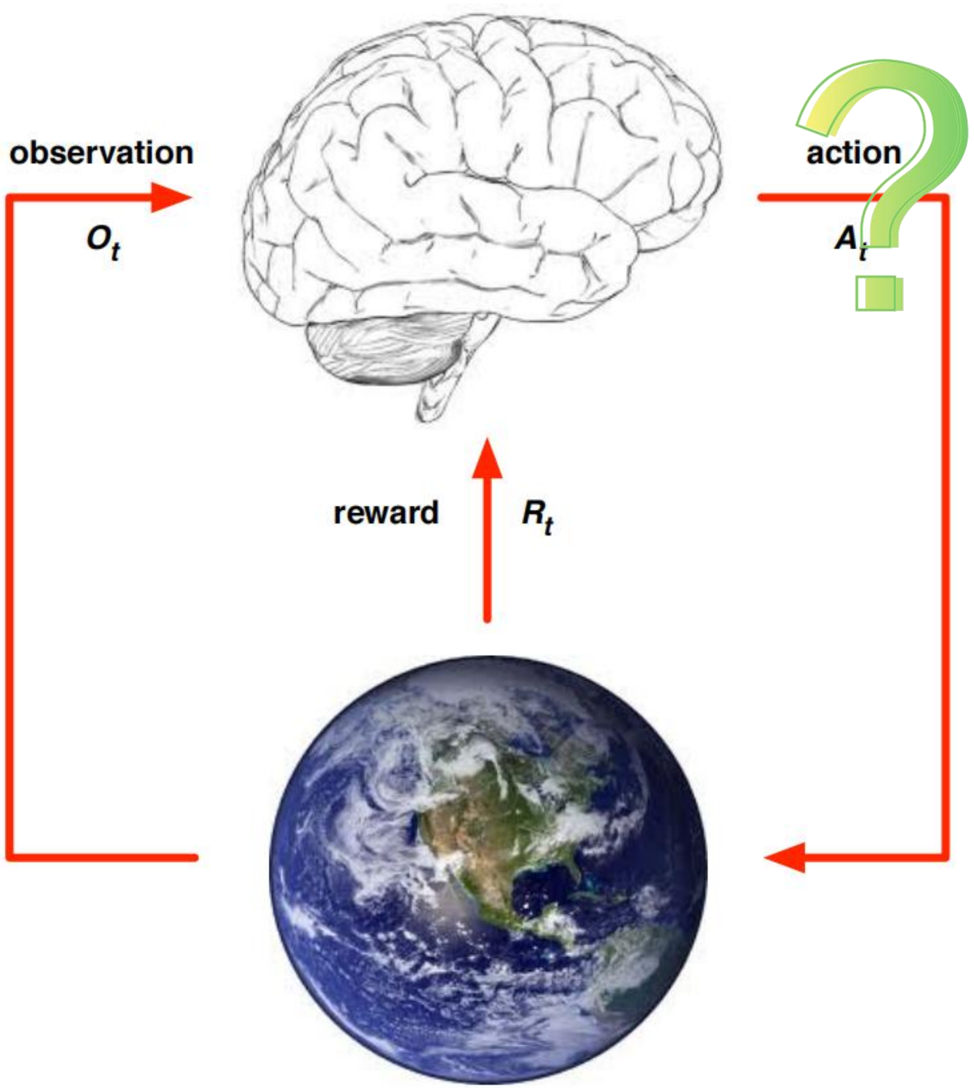
[a]Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030; [b]Department of Electrical and Computer Engineering, Rice University, Houston, TX 77030; [c]Center for Neuroscience and Artificial Intelligence, Baylor College of Medicine, Houston, TX 77030; [d]Geo Core Data, Atomic Maps, Google, Mountain View, CA 94043; [e]Department of Psychology, University of Minnesota, Minneapolis, MN 55455; and [f]Department of Computer Science, University of Minnesota, Minneapolis, MN 55455

Xaq Pitkow @ CMU, Neuroscience Institute, with affliation to ML Department

lab site: https://xaqlab.com/

presented by 陆杨帆，张博涛

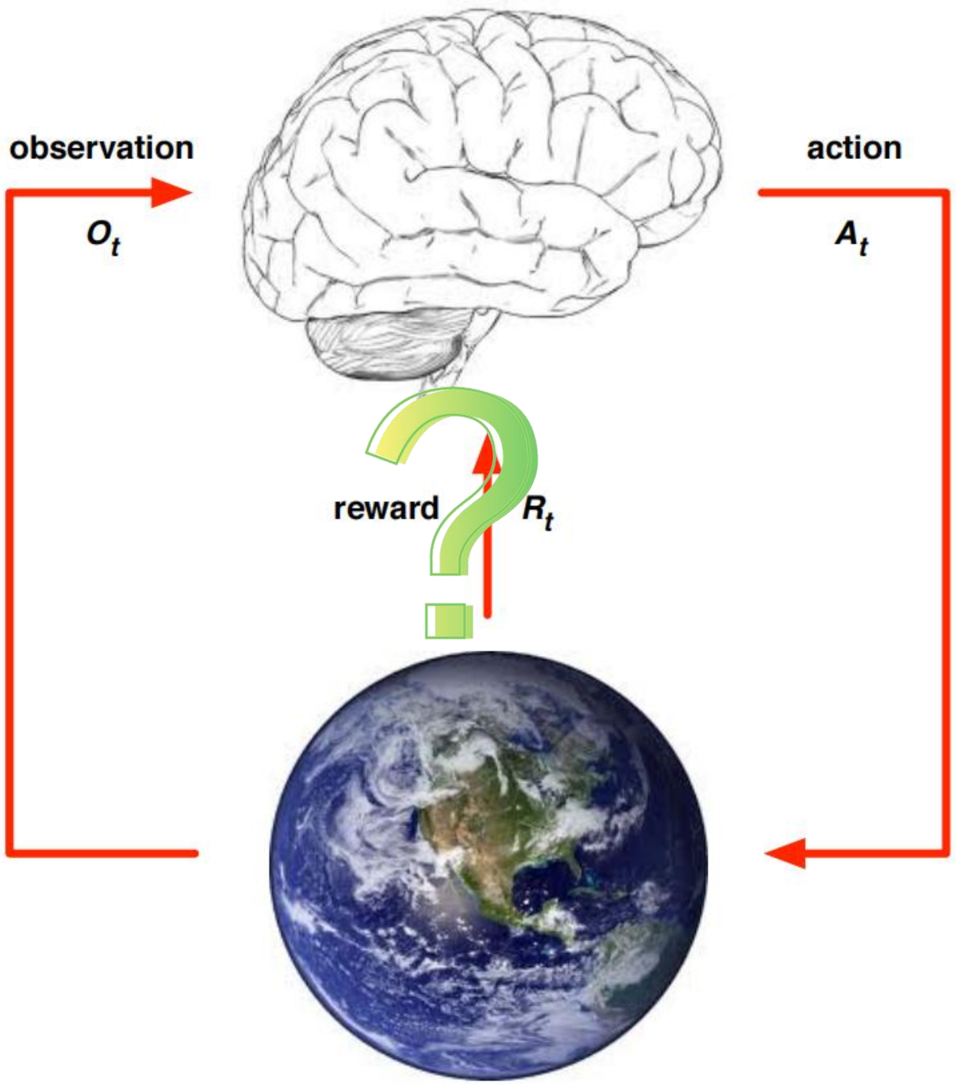
Reinforcement Learning



Given: reward(, state)

What is the optimal action?

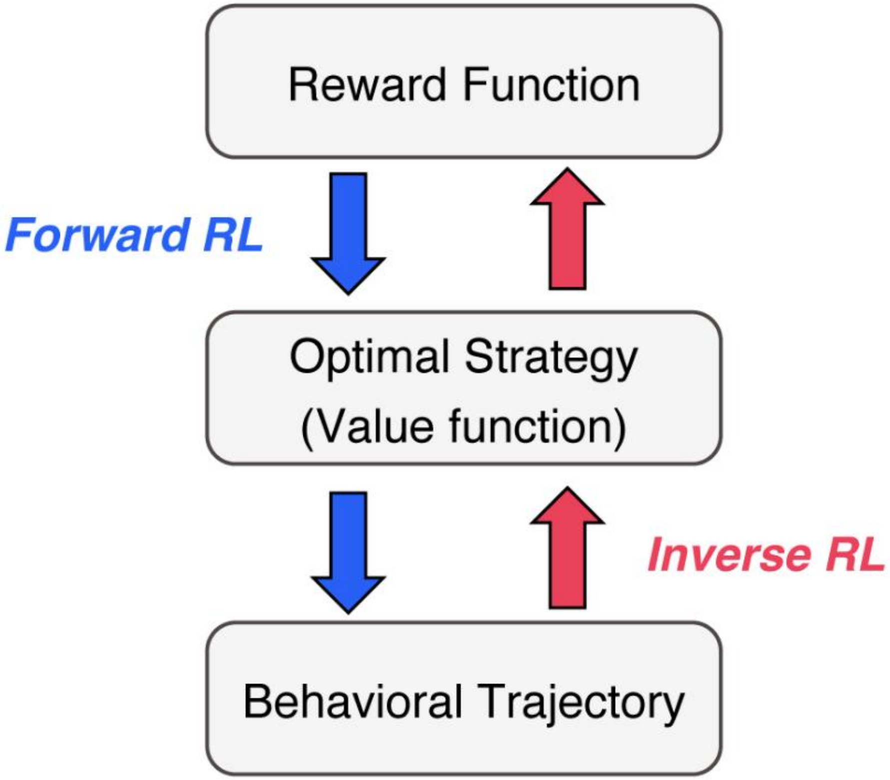
Inverse Reinforcement Learning



Given: Action trajectory of an agent(mostly, an expert)

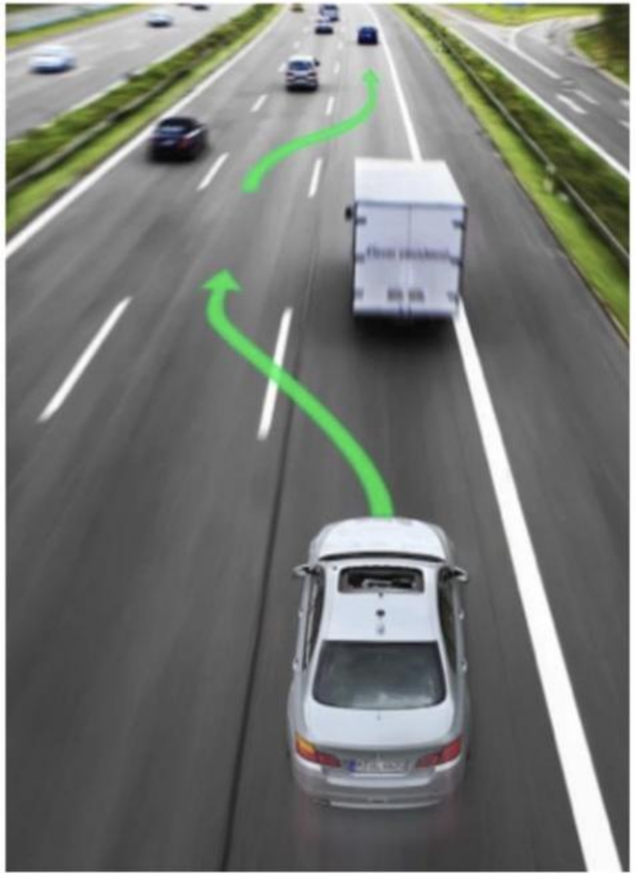
What is the reward(function that the agent is trying to maximize)?

Inverse Reinforcement Learning



I saw him walking this way, what is his intention?
i.e. What is the reward at each state(-action) that makes him act like this?

Inverse Reinforcement Learning



maintaining safe following distance?

keeping away from the curb?

staying far from any pedestrians?

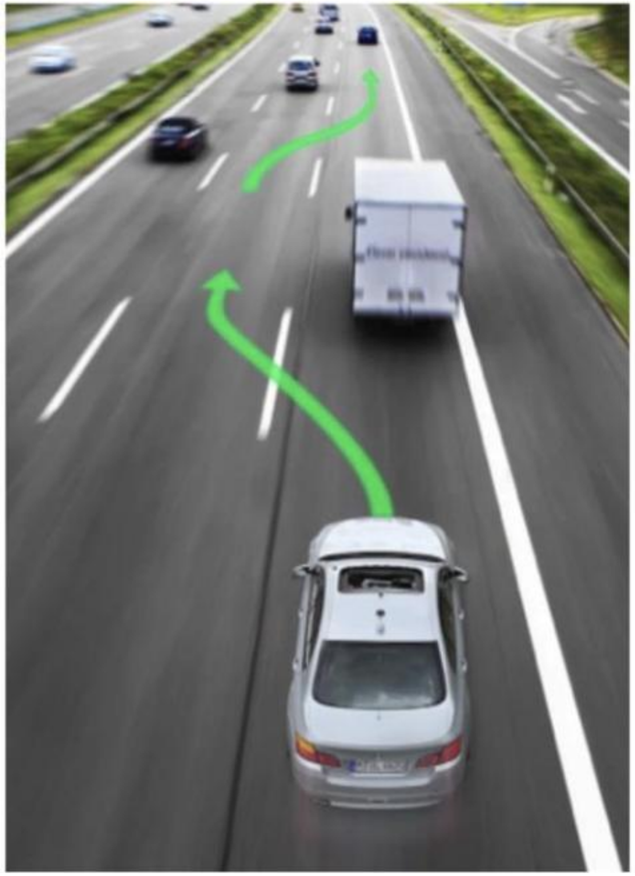
maintaining a reasonable speed?

a slight preference for driving in the
middle lane?

not changing lanes too often?
......

reward function is too complicated to hand-tune

Inverse Reinforcement Learning



apprenticeship learning

but we can see what an expert does ☺

(and thus infer the reward function that give rise to a good policy)

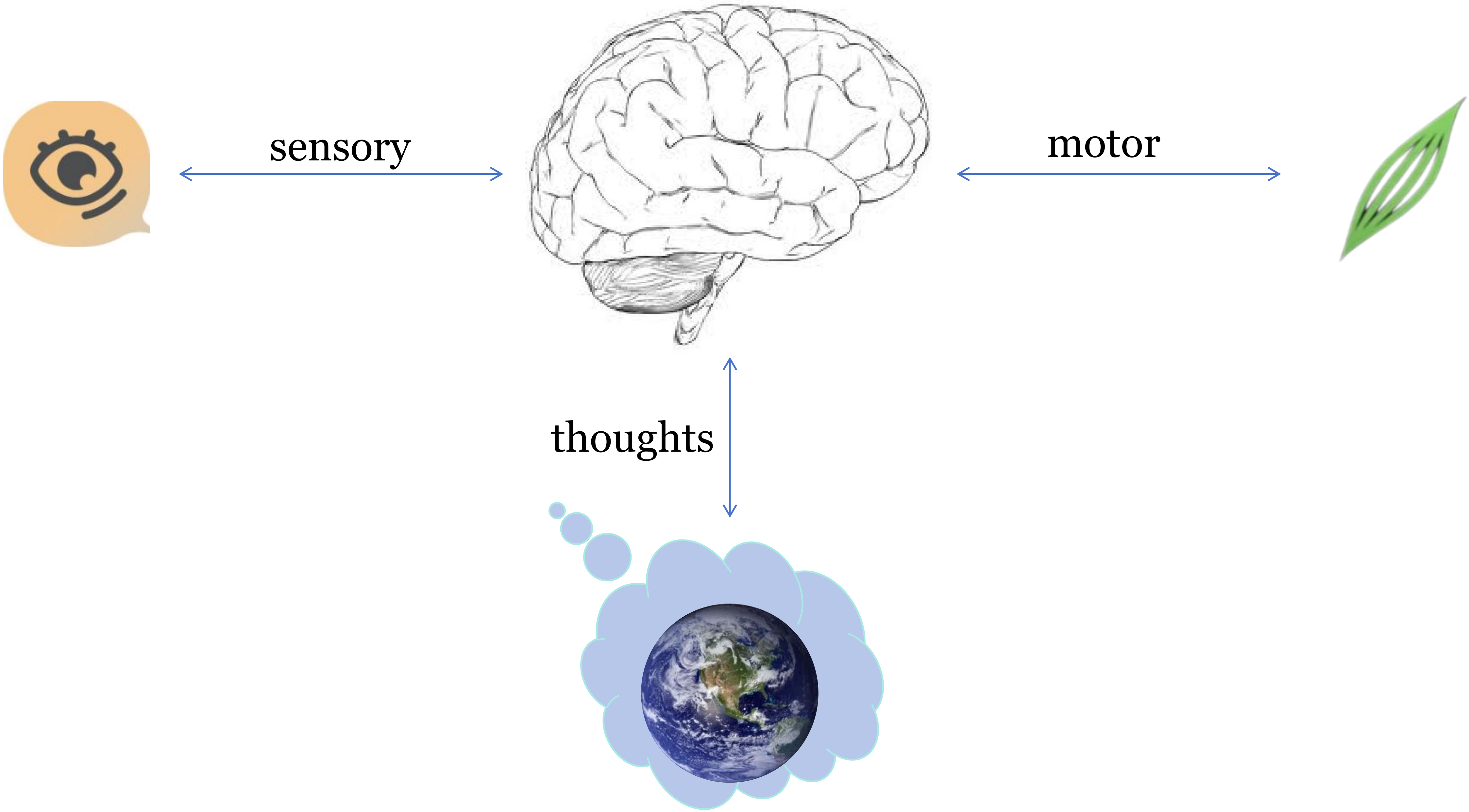
reward function is too complicated to hand-tune

Inverse Reinforcement Learning

Formal Definition:
Given
1) measurements of an agent's behaviour over time, in a variety of circumstances,
2) measurements of the sensory inputs to that agent;
3) a model of the physical environment (including the agent's body).

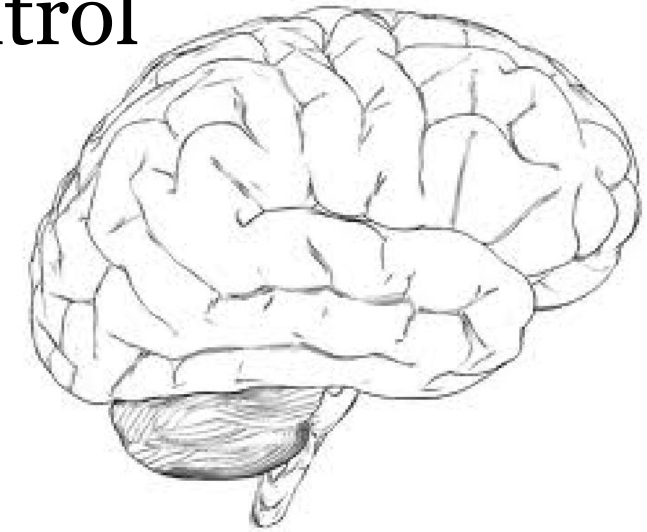
Determine the reward function that the agent is optimizing.

(Stuart Russell, 1998)


sensory
motor
thoughts

Inverse Rational Control

sensory



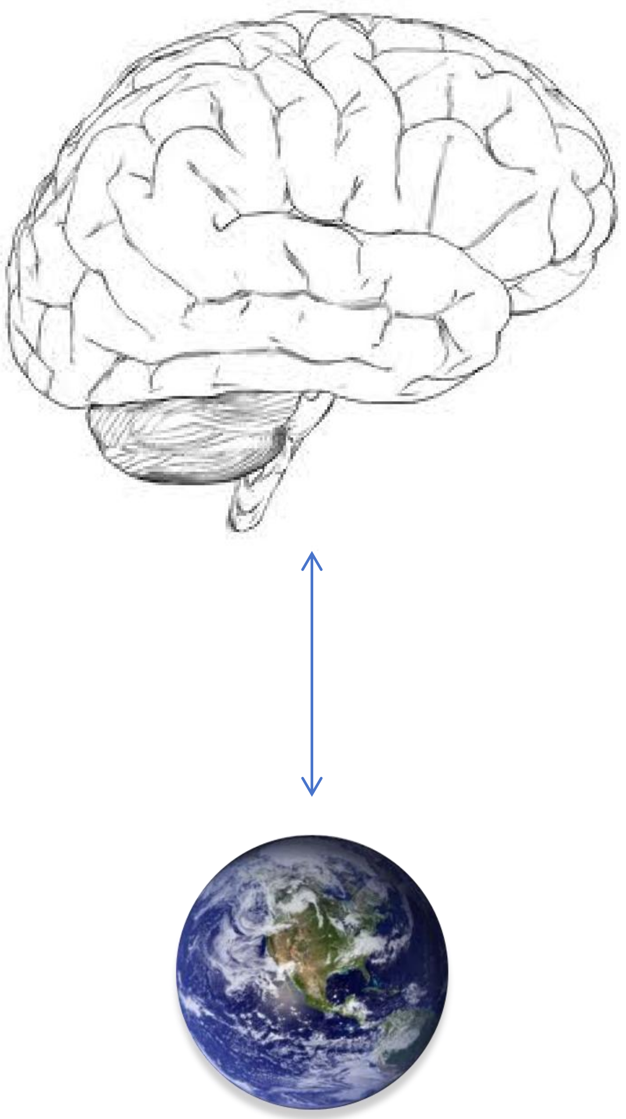
motor

We see what the agent observes
and how he acts
Then we try to infer:

What's the agent's mental model?

What's his belief of
the world, e.g.
transition prob,
reward..

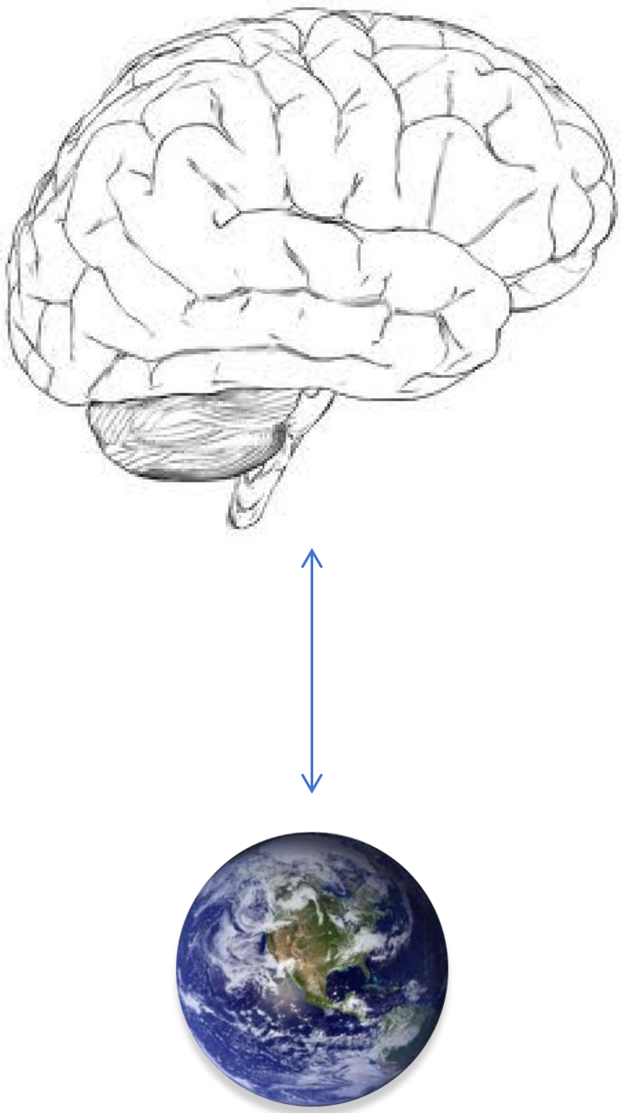
Inverse Rational Control



What if the agent is wrong… but

rational: doing wrong things for
the right reasons

Inverse Rational Control

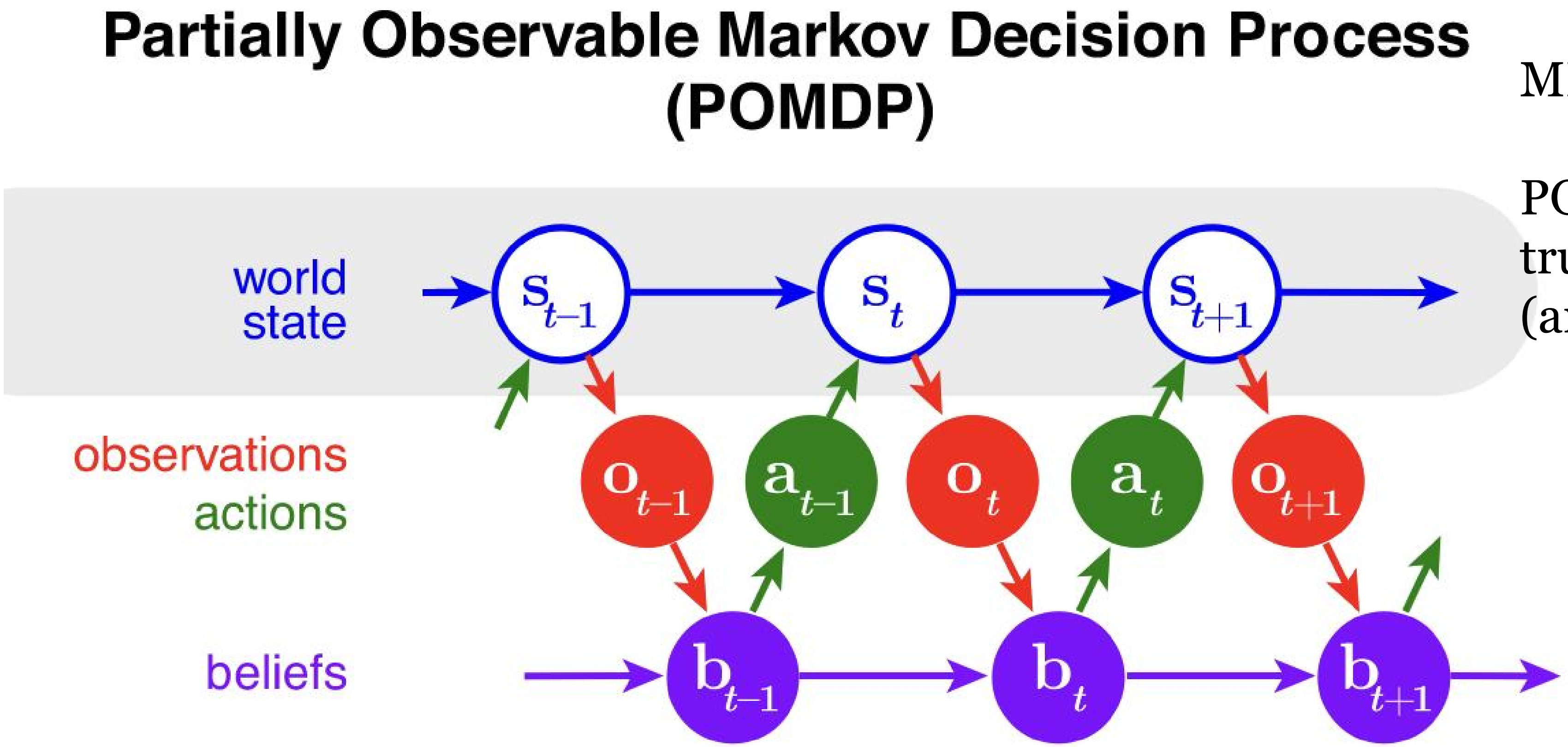


What if the agent is wrong... but

rational: doing wrong things for
the right reasons

The agent's belief about the world is wrong, but it acts optimally according to the belief (giving rise to suboptimal behavior)

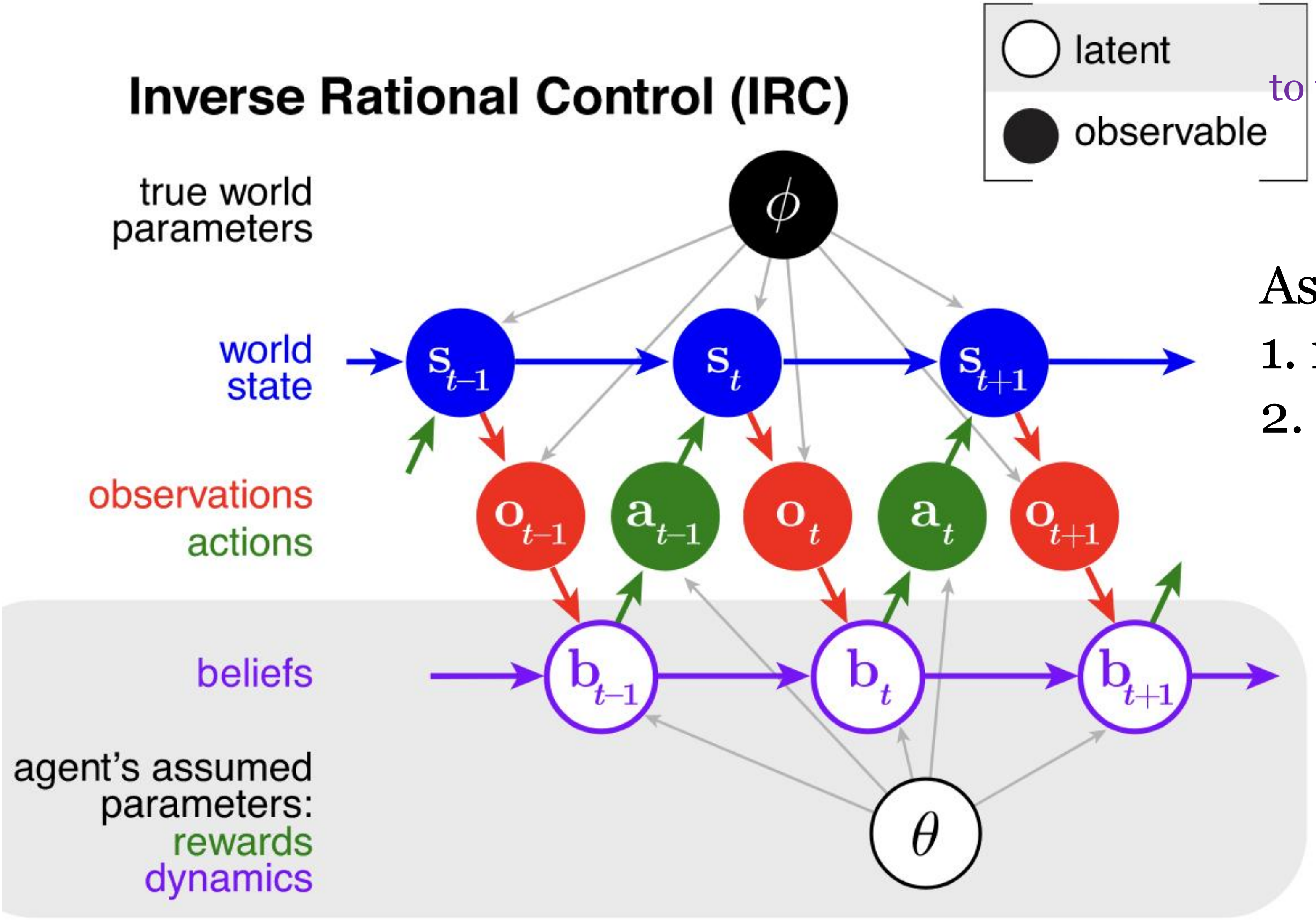
Modelling Behavior as Rational



MDP: agent knows the real state

POMDP: agent has to infer the true state from observations (and actions)

Modelling Behavior as Rational

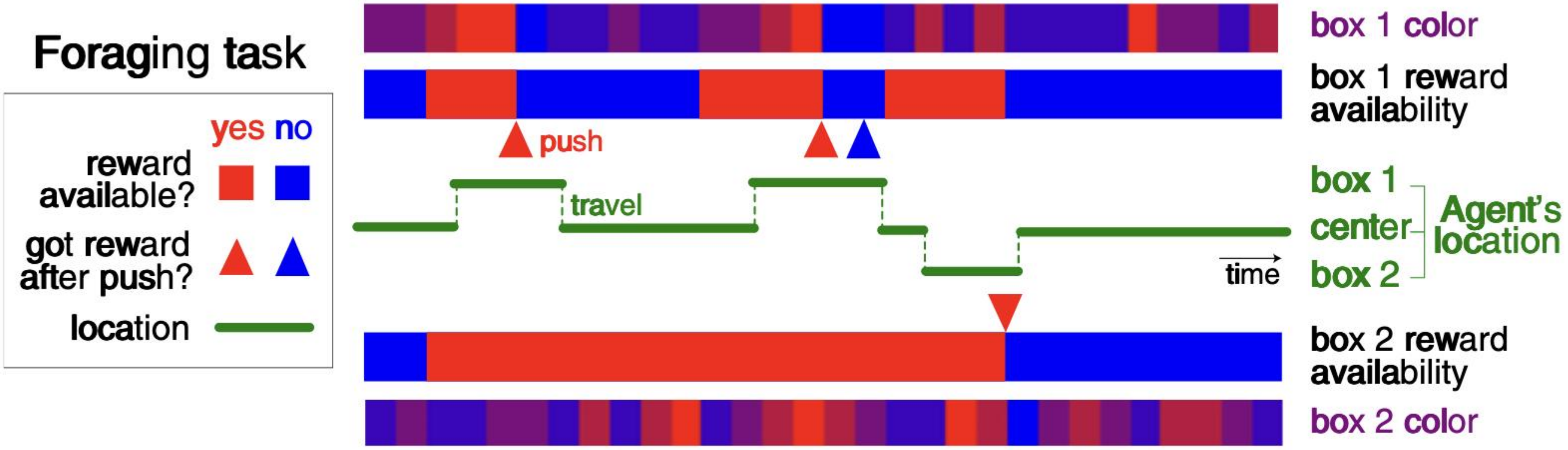


to the experimenter

Assumptions:

1. maximize total utility
2. stationary policy, no learning

Modelling Behavior as Rational



agent's parameters:

appearance rate(box1, 2),
disappearance rate(box1, 2)

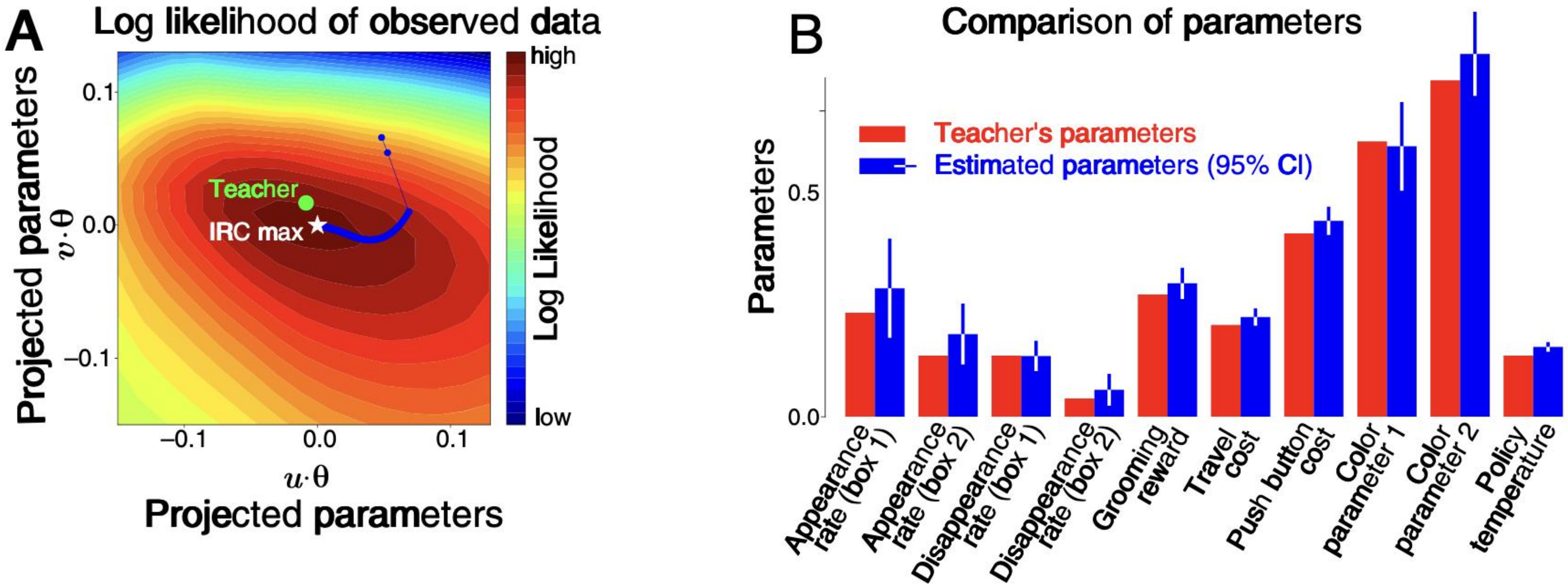
(reward availability as a telegraph process)
(these parameters are very useful when you're far away from the box)

color parameters(box1, 2)
(color drawn from a binomial)
(useful when you're right at the box)

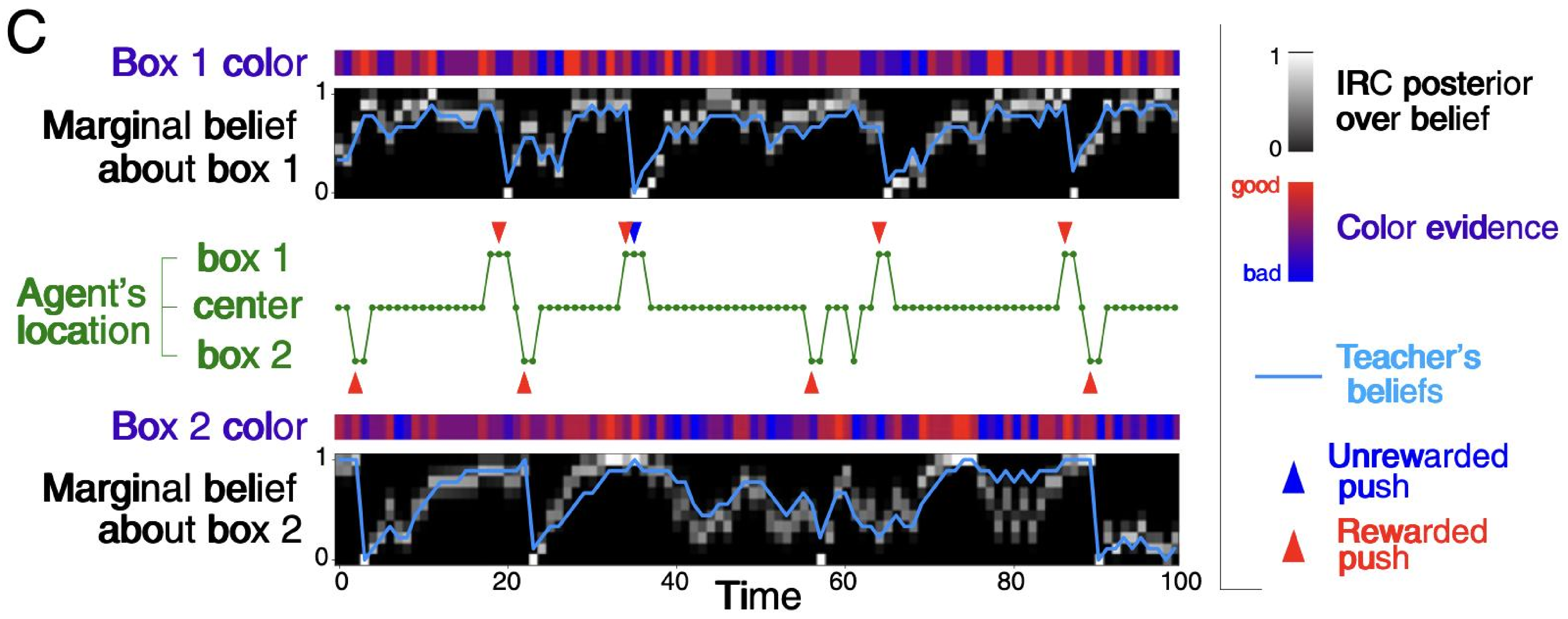
utility:
grooming reward
travel cost
push button cost

Maths behind modelling

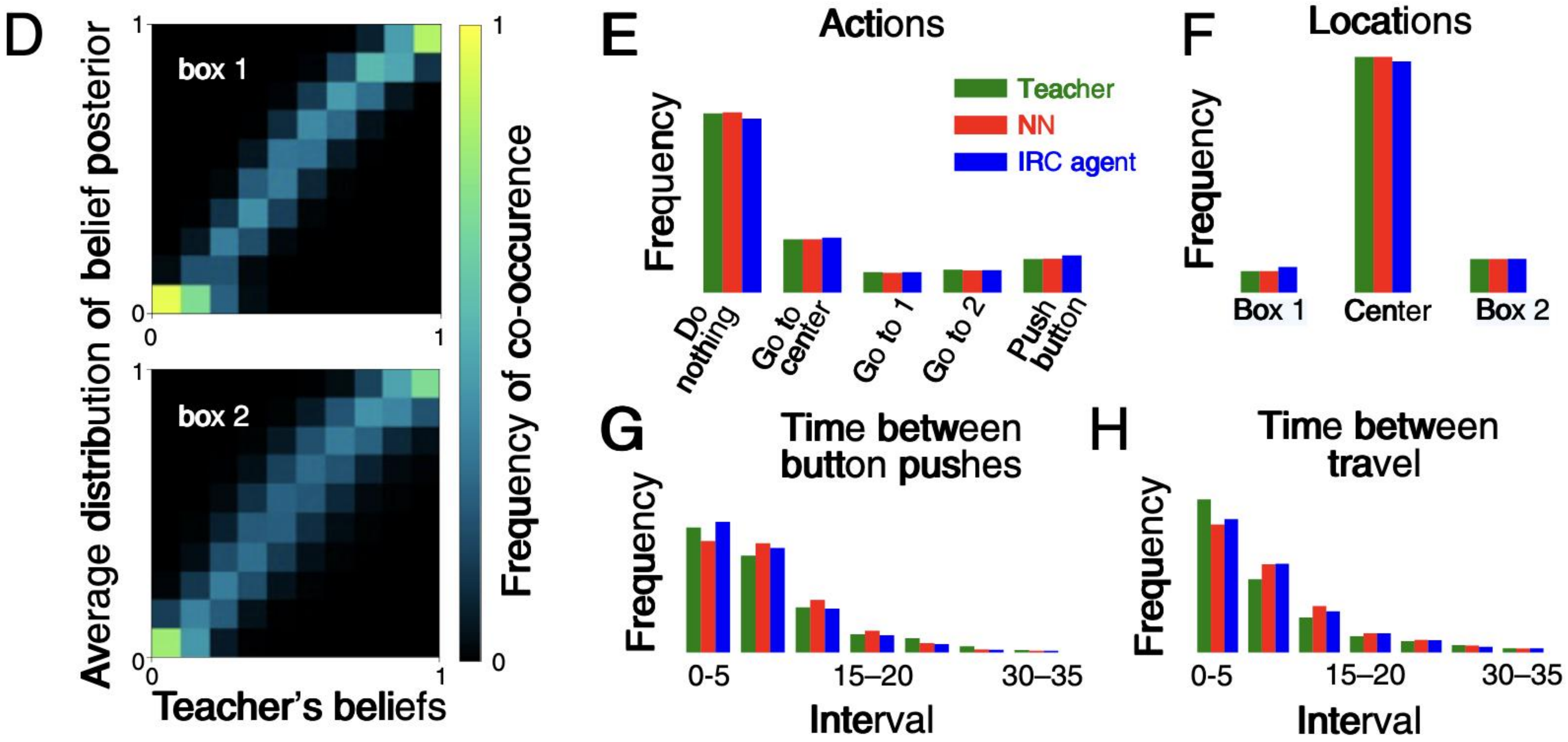
Modelling Behavior as Rational



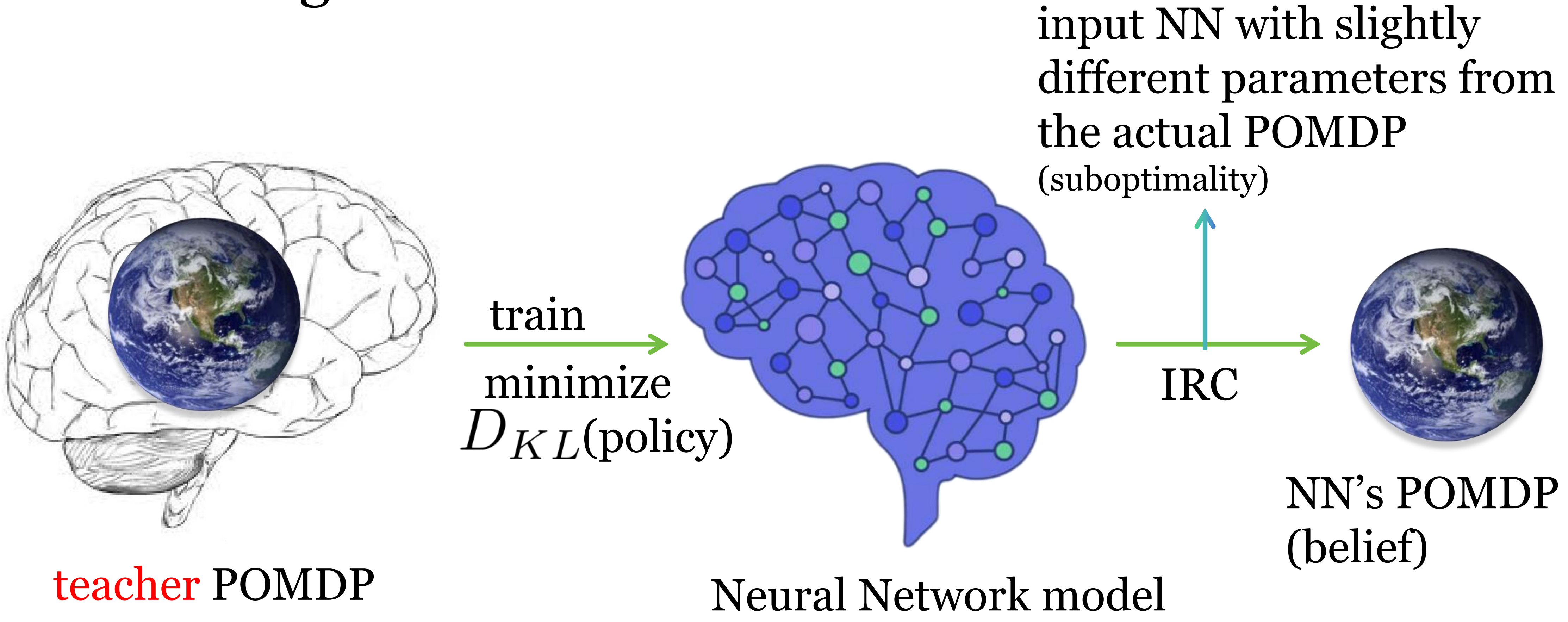
Modelling Behavior as Rational



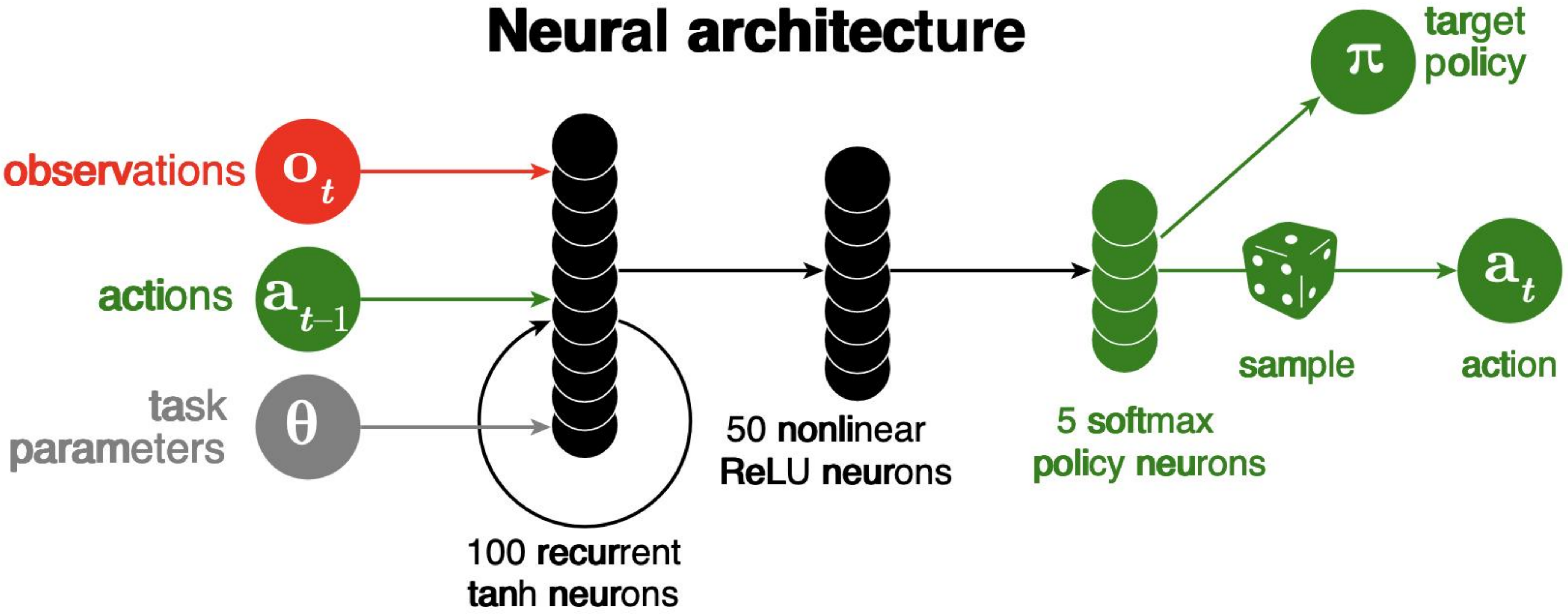
Modelling Behavior as Rational



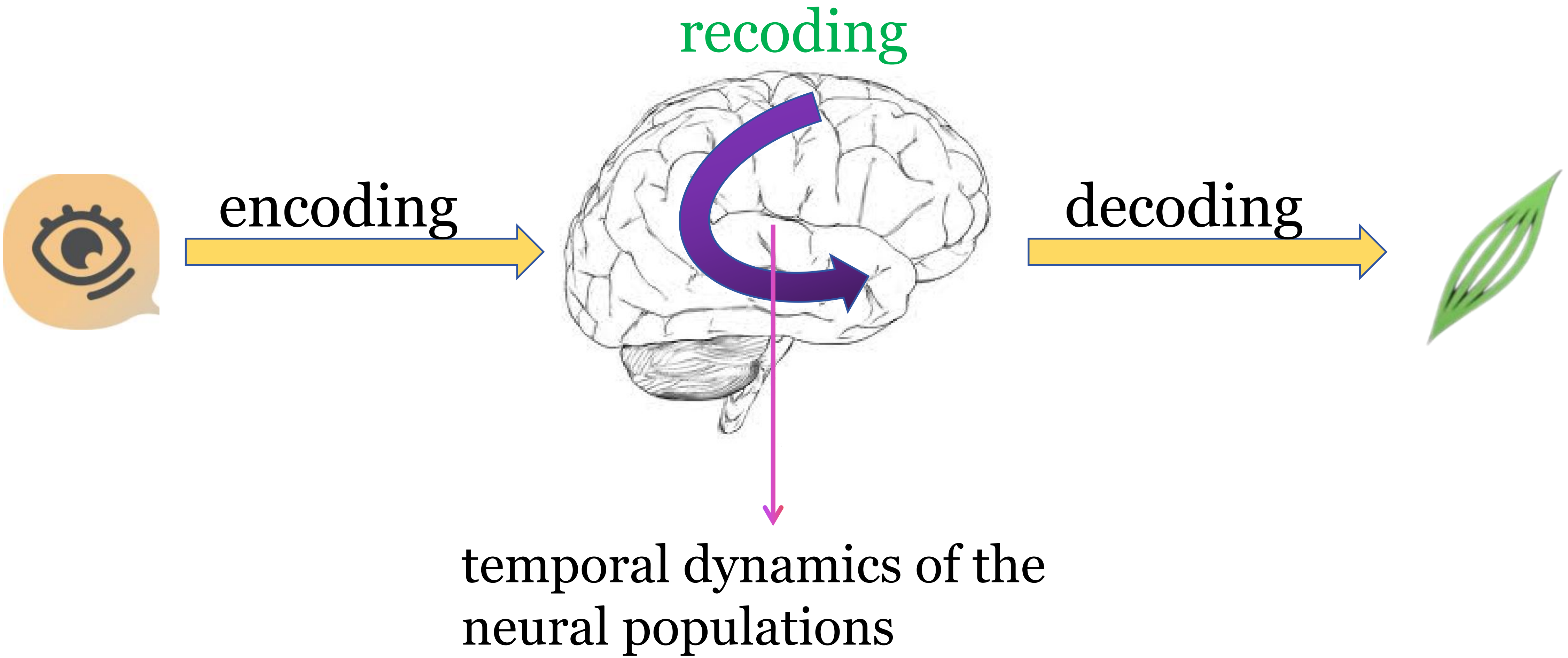
Modelling Behavior as Rational



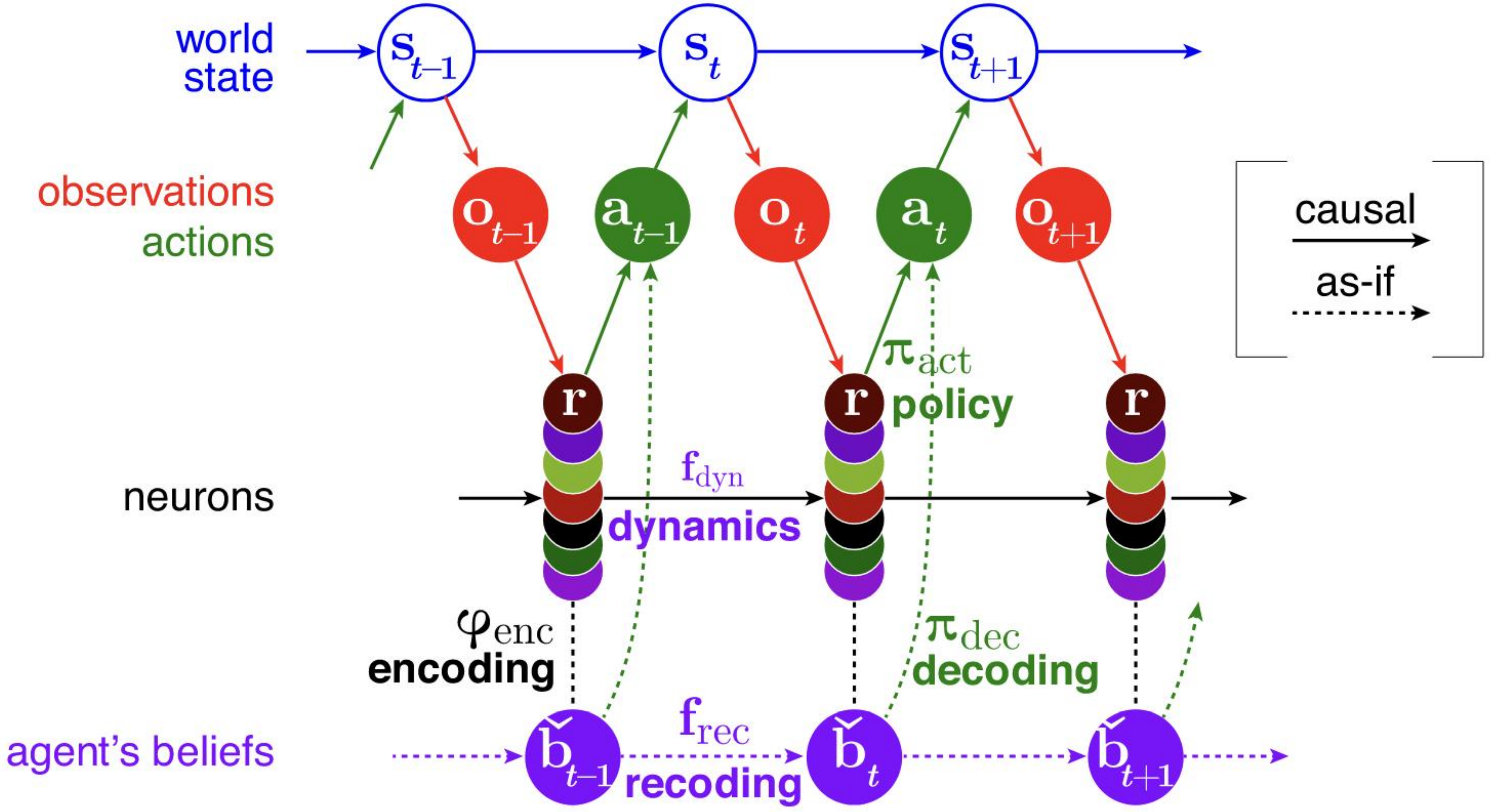
Modelling Behavior as Rational



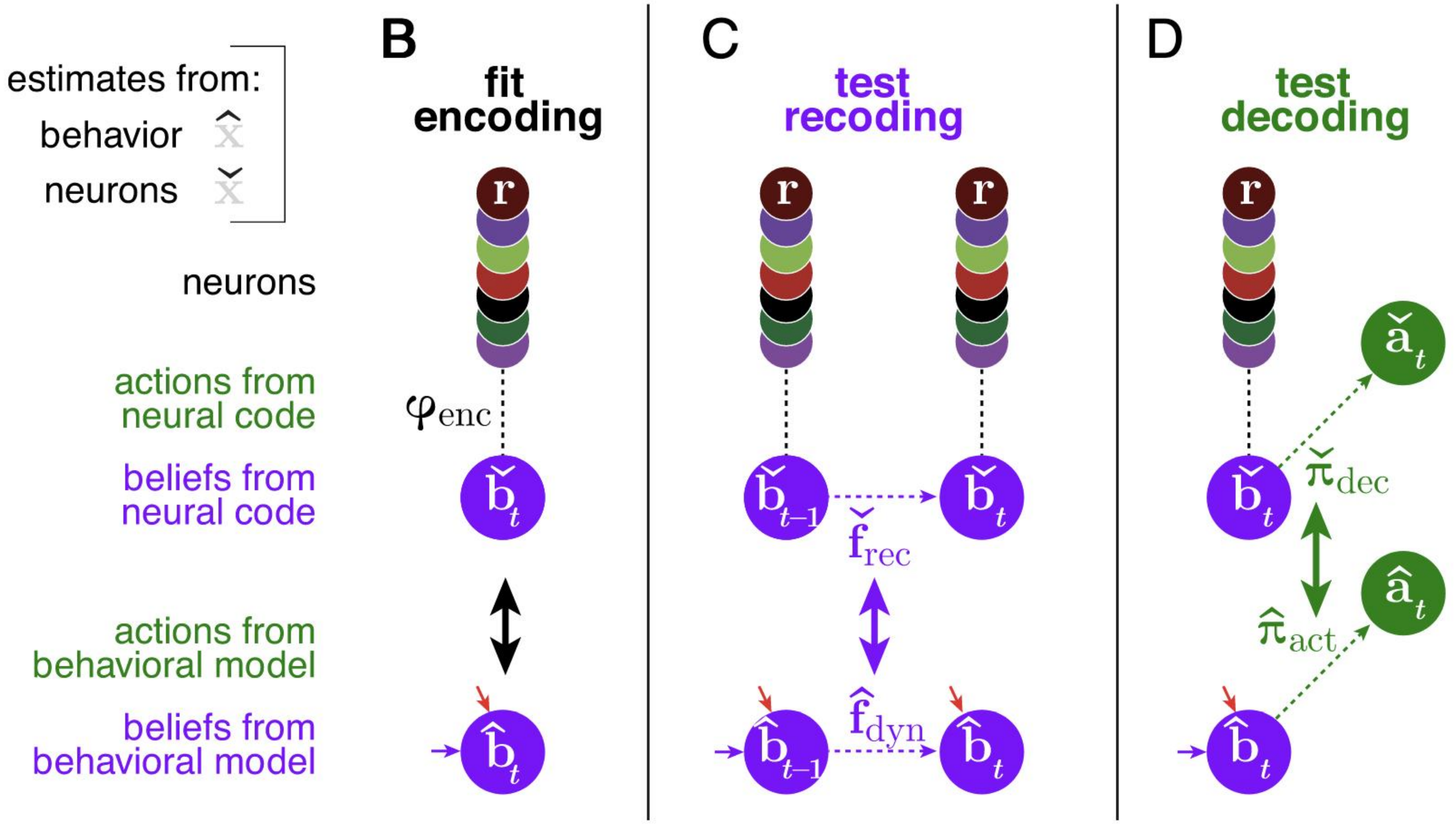
Neural Coding



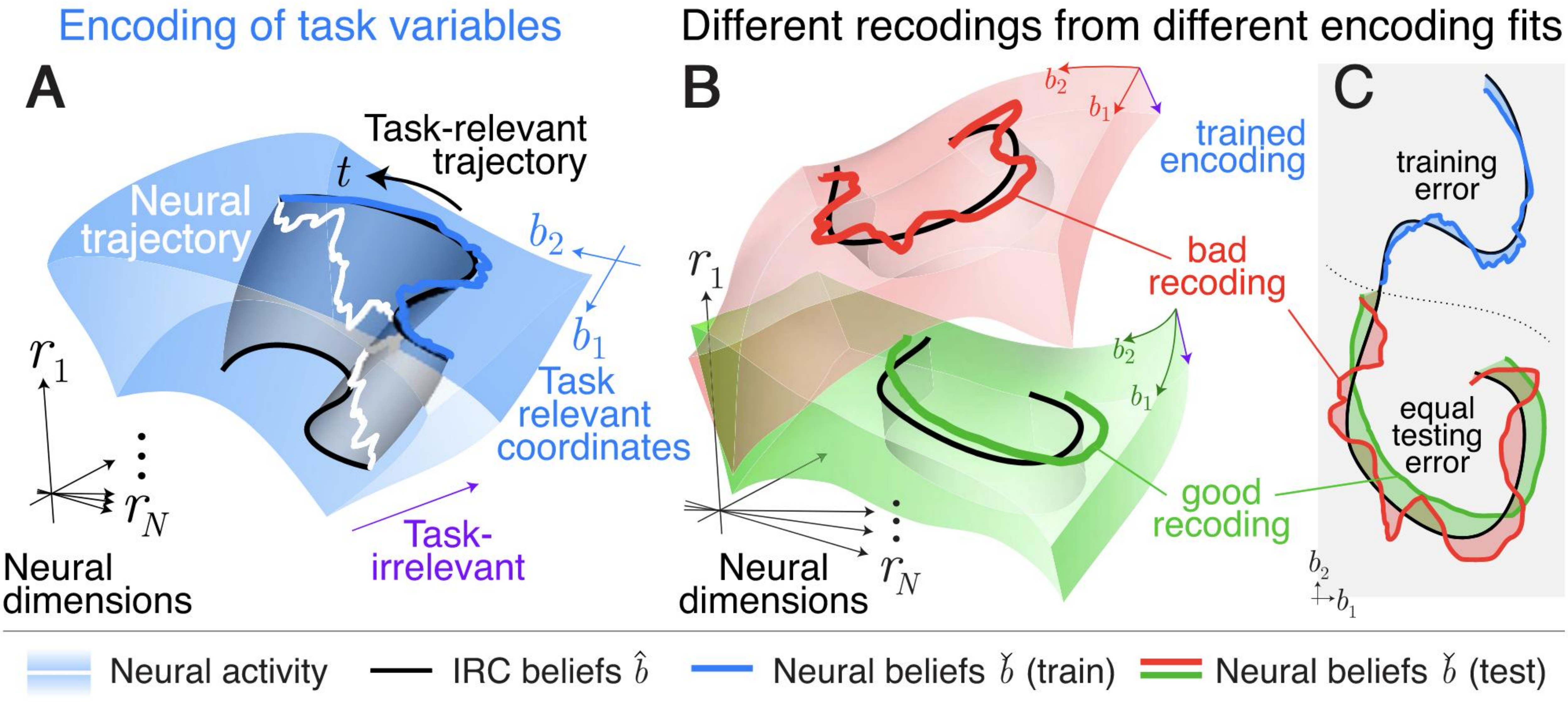
Neural Coding: neural implementation of POMDP

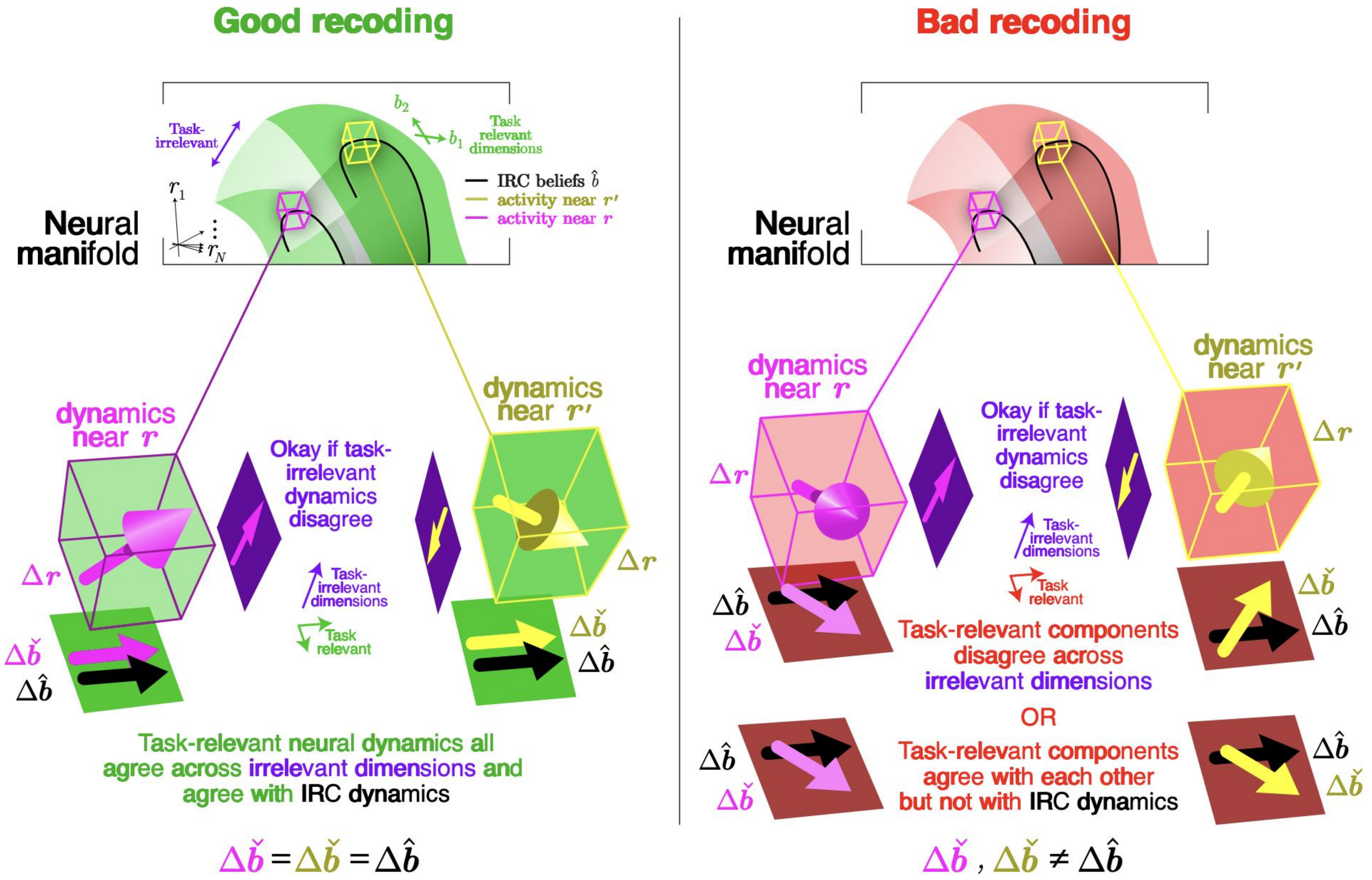


Neural Coding of rational thoughts



Neural Coding

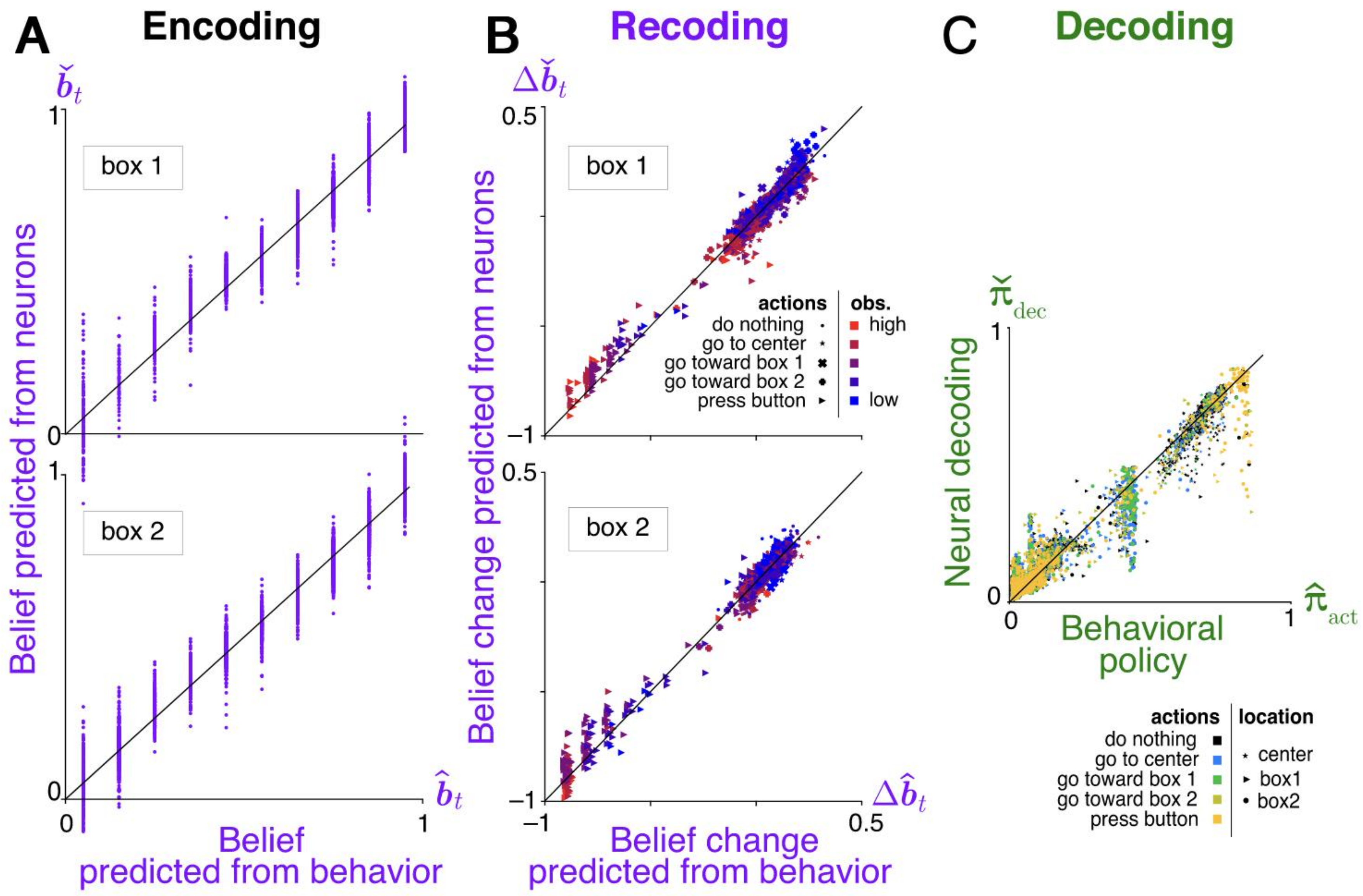




Good recoding
Bad recoding
Task-irrelevant
b_2
b_1
Task relevant dimensions
r_1
r_N
IRC beliefs $\hat{b}$
activity near r'
activity near r
Neural manifold
Neural manifold
dynamics near r
dynamics near r'
Okay if task-irrelevant dynamics disagree
Δr
Δr
Task-irrelevant dimensions
Task relevant
$\Delta\check{b}$
$\Delta\hat{b}$
$\Delta\check{b}$
$\Delta\hat{b}$
Task-relevant neural dynamics all agree across irrelevant dimensions and agree with IRC dynamics
$\Delta\check{b} = \Delta\check{b} = \Delta\hat{b}$
dynamics near r
dynamics near r'
Okay if task-irrelevant dynamics disagree
Δr
Δr
Task-irrelevant dimensions
Task relevant
$\Delta\hat{b}$
$\Delta\check{b}$
$\Delta\check{b}$
$\Delta\hat{b}$
Task-relevant components disagree across irrelevant dimensions
OR
Task-relevant components agree with each other but not with IRC dynamics
$\Delta\hat{b}$
$\Delta\check{b}$
$\Delta\hat{b}$
$\Delta\check{b}$
$\Delta\check{b}, \Delta\check{b} \neq \Delta\hat{b}$

Neural Coding

Neural Coding



Problems

Problems

THE END

unless there’s still time left

(?)Representation learning is the bridge from planning to problem solving

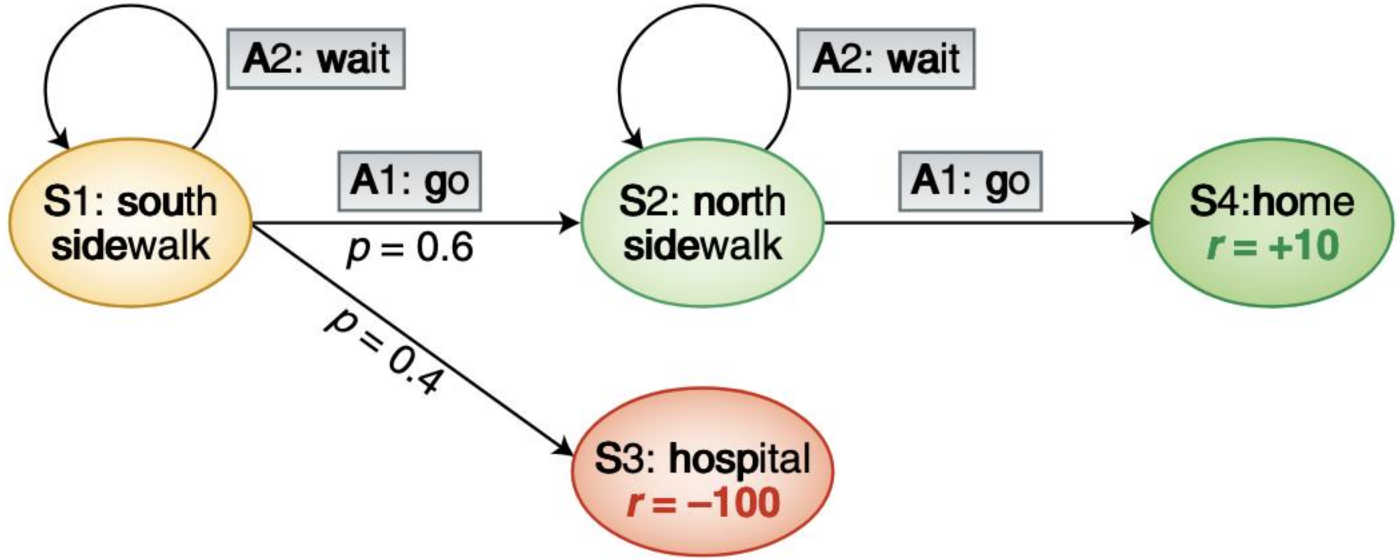
PERSPECTIVE | FOCUS
https://doi.org/10.1038/s41593-019-0470-8

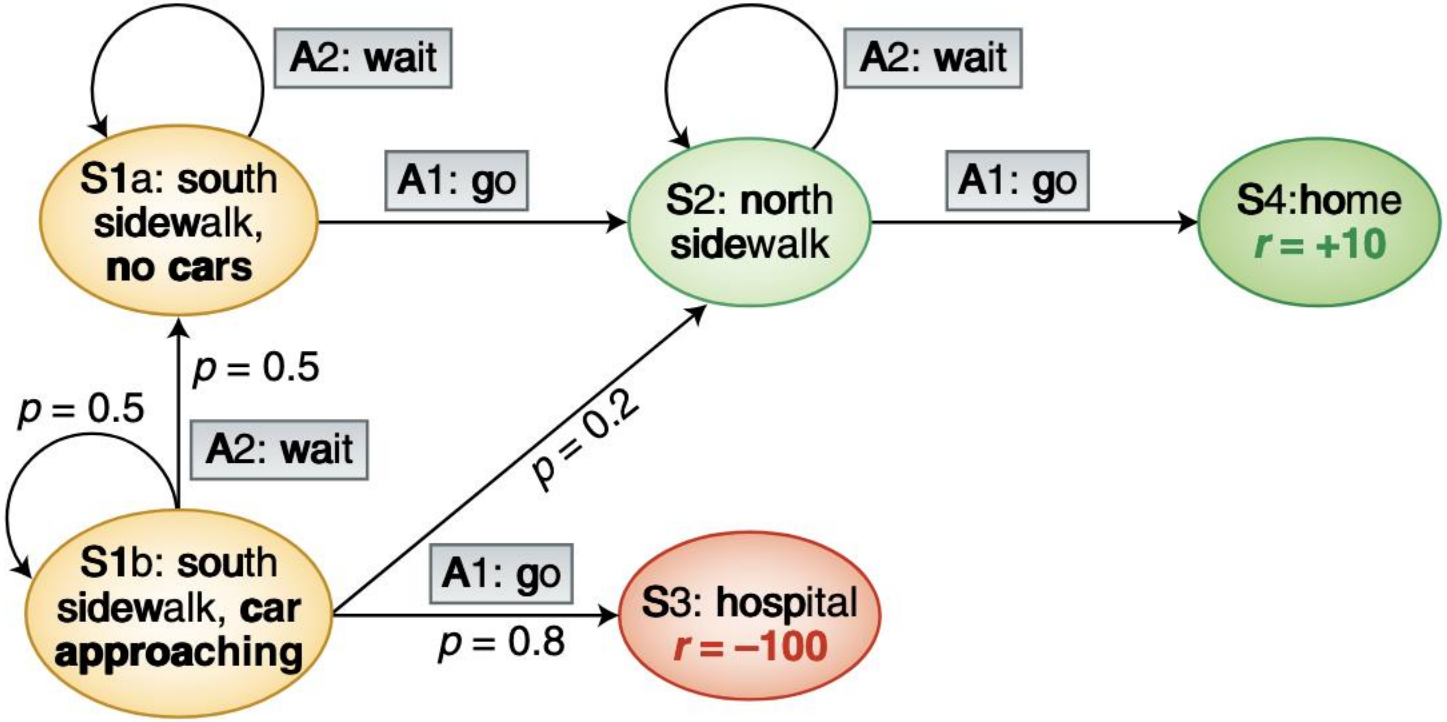
nature neuroscience

Learning task-state representations

Yael Niv

Arguably, the most difficult part of learning is deciding what to learn about. Should I associate the positive outcome of safely completing a street-crossing with the situation 'the car approaching the crosswalk was red' or with 'the approaching car was slowing down'? In this Perspective, we summarize our recent research into the computational and neural underpinnings of 'representation learning'—how humans (and other animals) construct task representations that allow efficient learning and decision-making. We first discuss the problem of learning what to ignore when confronted with too much information, so that experience can properly generalize across situations. We then turn to the problem of augmenting perceptual information with inferred latent causes that embody unobservable task-relevant information, such as contextual knowledge. Finally, we discuss recent findings regarding the neural substrates of task representations that suggest the orbitofrontal cortex represents 'task states', deploying them for decision-making and learning elsewhere in the brain.


A2: wait
A2: wait
S1: south sidewalk
A1: go
$p = 0.6$
S2: north sidewalk
A1: go
S4:home
$r = +10$
$p = 0.4$
S3: hospital
$r = -100$


A2: wait
A2: wait
S1a: south sidewalk, no cars
A1: go
S2: north sidewalk
A1: go
S4:home
$r = +10$
$p = 0.5$
$p = 0.5$
A2: wait
$p = 0.2$
S1b: south sidewalk, car approaching
A1: go
$p = 0.8$
S3: hospital
$r = -100$


a
N
Cafe Dalsace

Problems in Mathematics

10 people A1-A10 attend a meeting, some of them shake hands with each other.

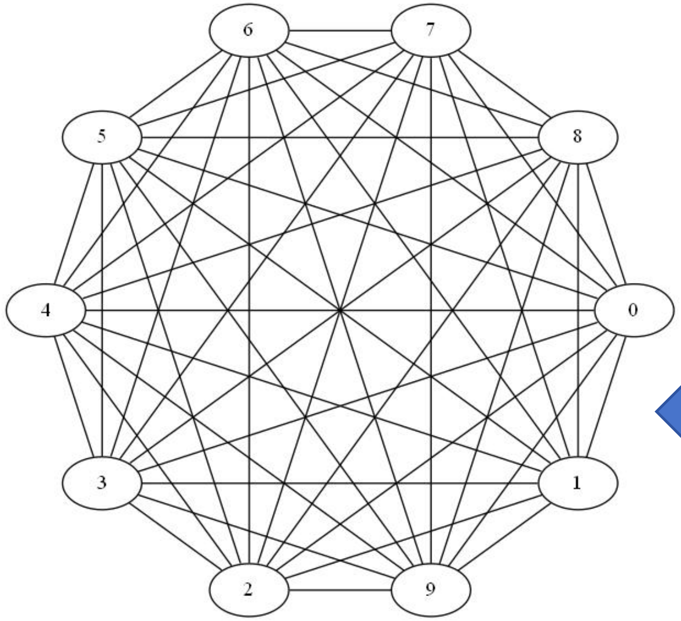
After the meeting, they report how many people they’ve shaked hands with: respectively 2, 5, 6, 3, 10, 5, 7, 1, 4, 8.

Is it possible?

Problems in Mathematics

10 people A1-A10 attend a meeting, some of them shake hands with each other.

After the meeting, they report how many people they've shaked hands with: respectively 2, 5, 6, 3, 10, 5, 7, 1, 4, 8.

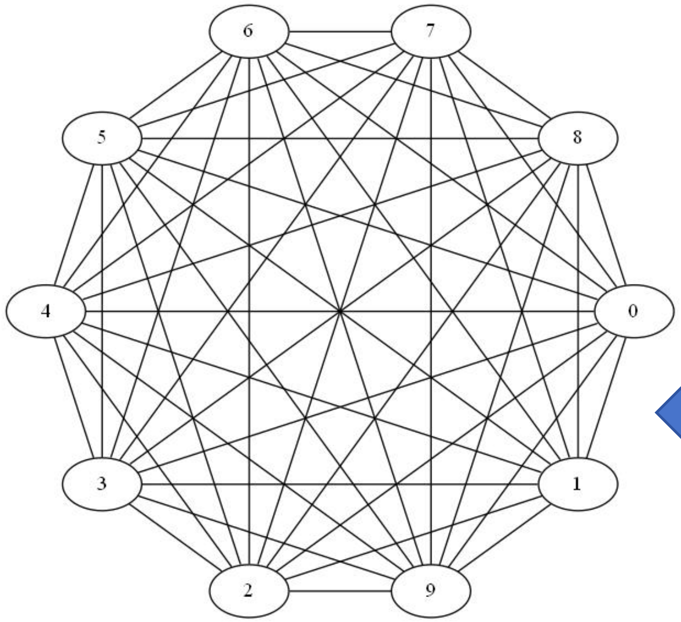


If you try to do a tree search on the graph, tremendous computation…

Problems in Mathematics

10 people A1-A10 attend a meeting, some of them shake hands with each other.

After the meeting, they report how many people they've shaked hands with: respectively 2, 5, 6, 3, 10, 5, 7, 1, 4, 8.

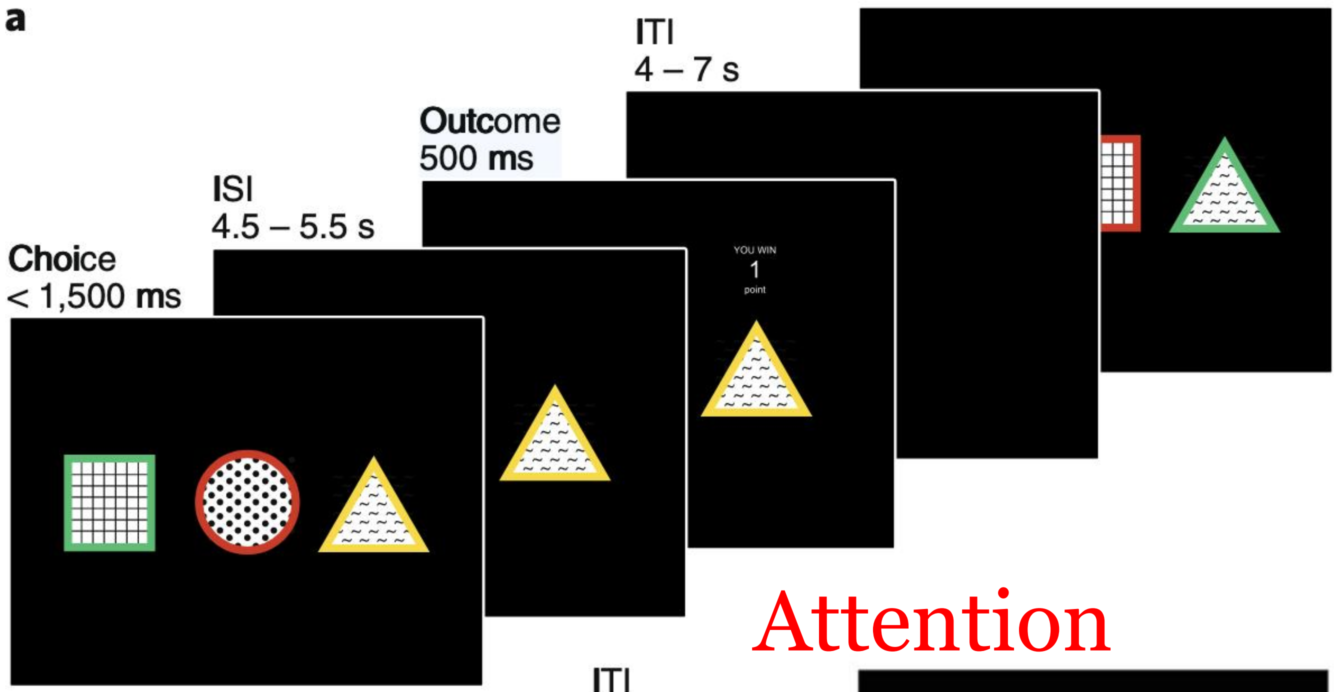


If you try to do a tree search on the graph, tremendous computation…

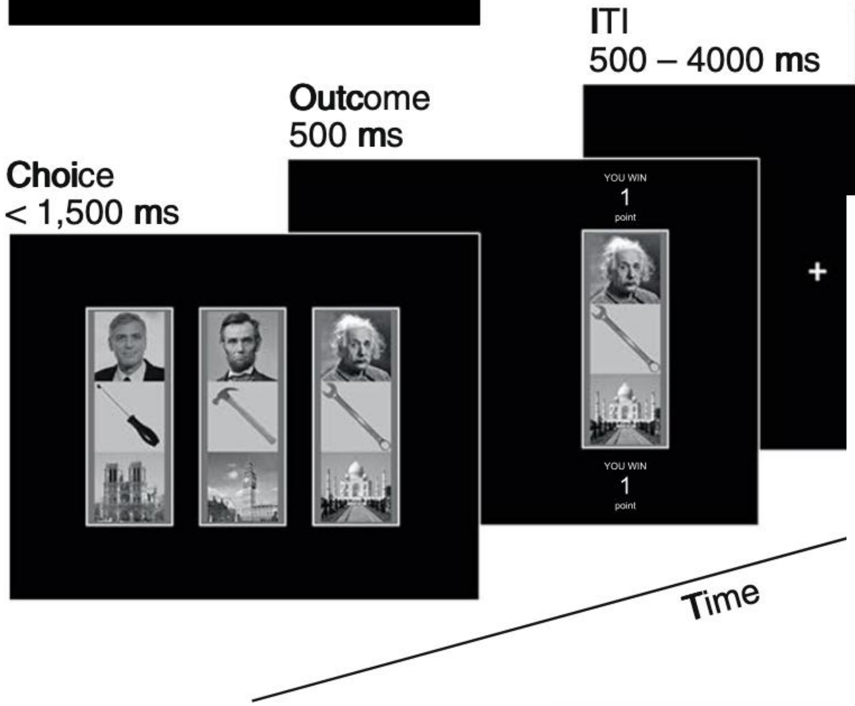
…but if you find an efficient representation of the task:
Is the sum of the degrees even?

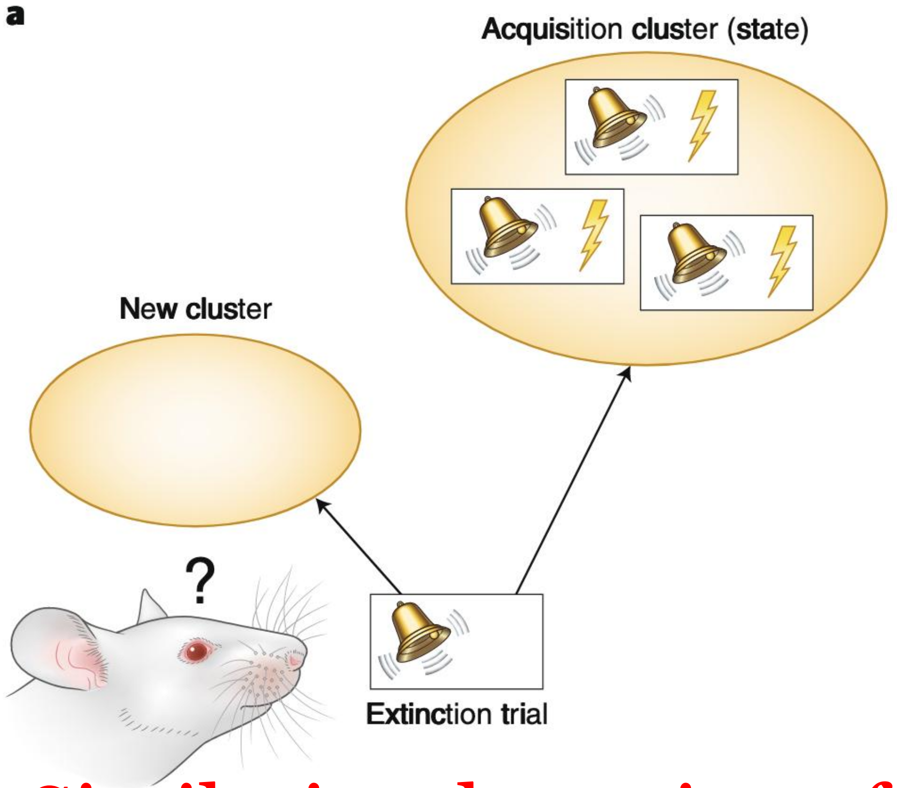
2+5+…+8 = odd, so someone's lying!

- tasks do not have unique state representations
- the brain solves seemingly complex tasks by learning efficient, low-dimensional representations that simplify these tasks
- efficient representations are task-specific
- how living agents know what to represent in order to use neural RL to solve tasks?

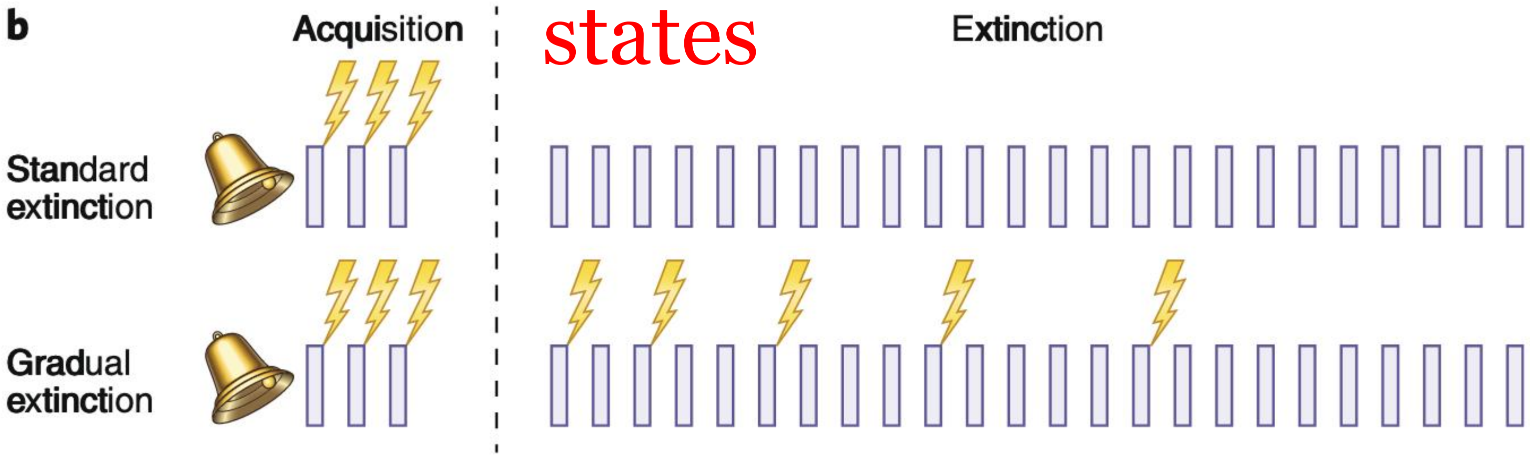


Attention





Similarity clustering of states



Can we devise complex tasks where representation (learning) really is the game changer?

THE END