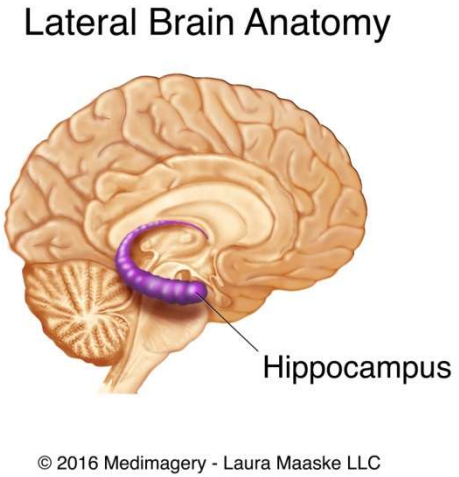
Memories of space

CS786

October 22nd 2024

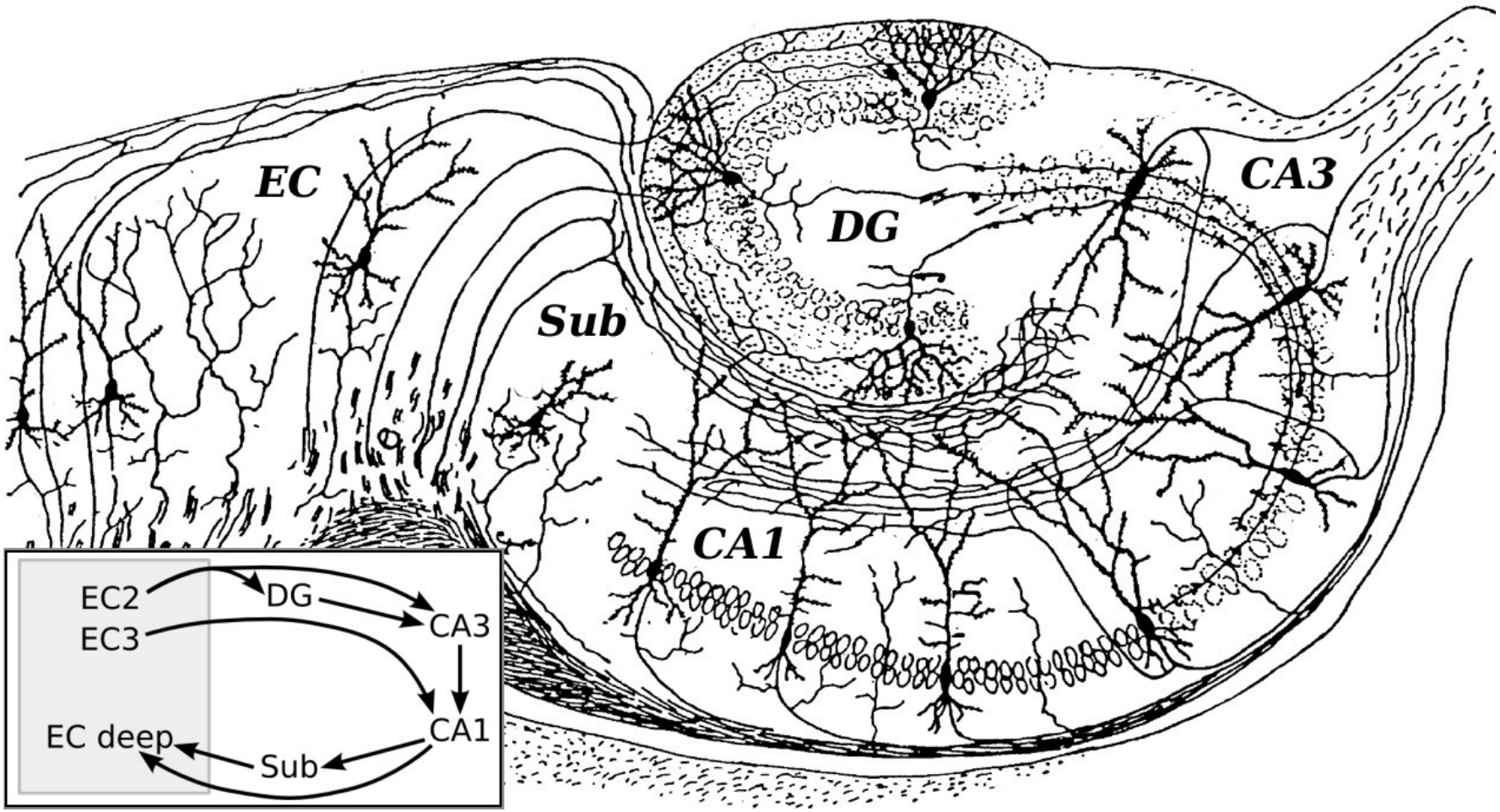
A fortuitous discovery

- H.M. suffered from epilepsy from a young age
- Operated upon for treatment
- Post-op presented with a pure case of anterograde amnesia
- Demonstrated criticality of hippocampus for memory formation



https://www.newyorker.com/books/page-turner/the-man-who-forgot-everything

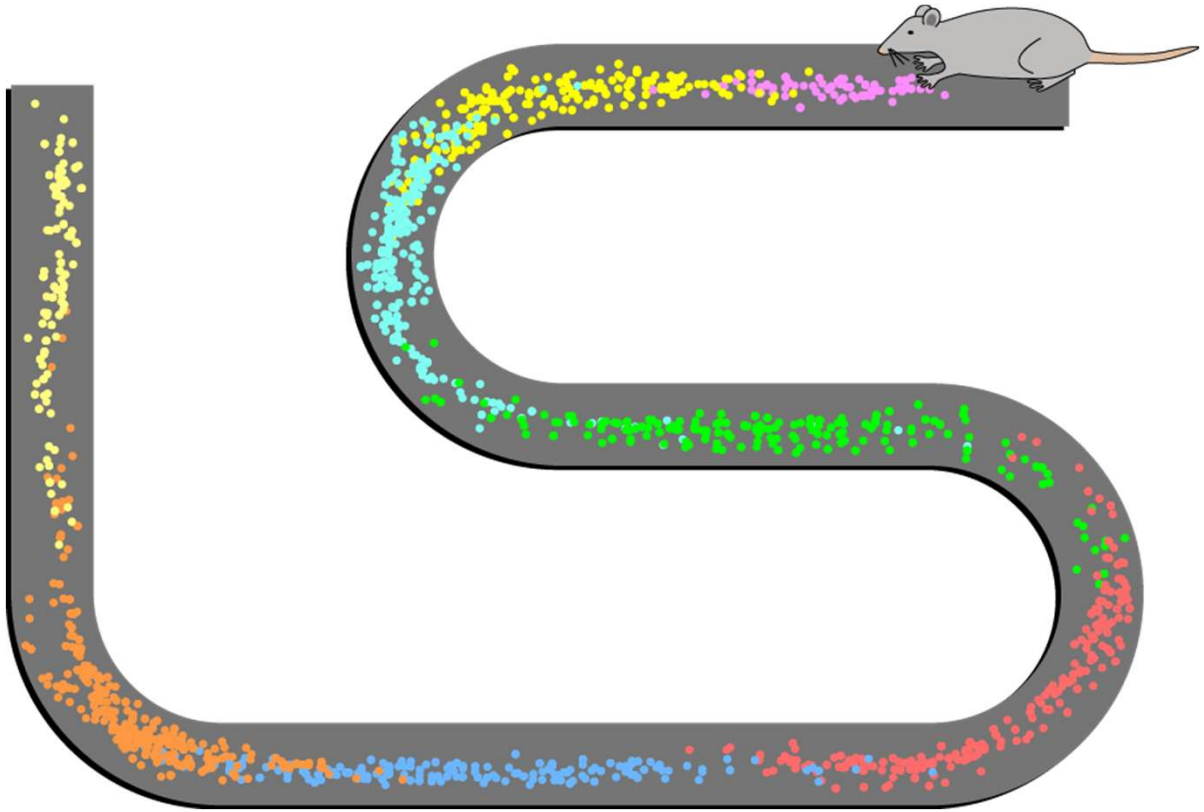
Hippocampus



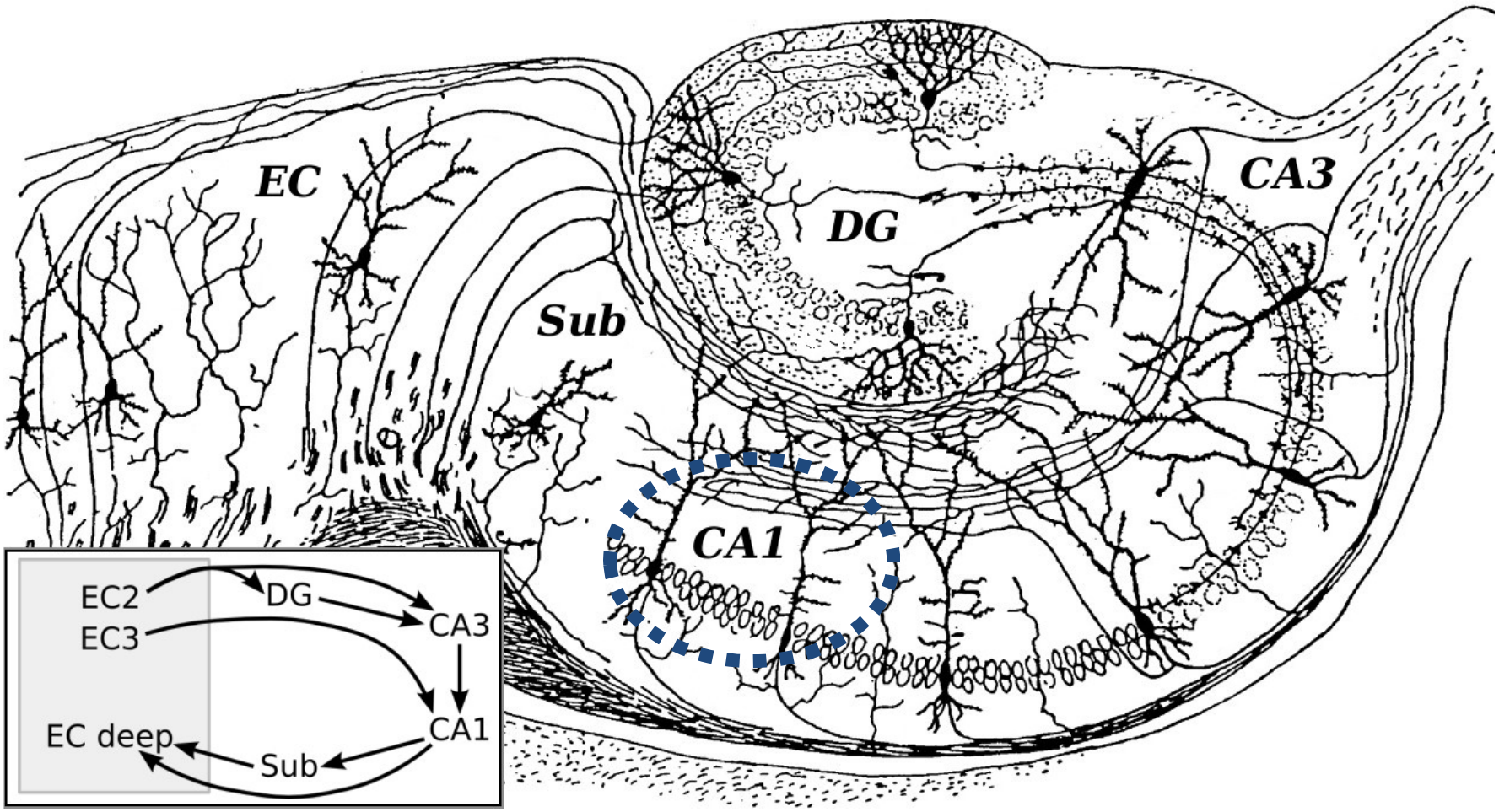
Phenomenally interesting anatomical structure; crucial for forming representations

Place cells

- Firing patterns of 8 place cells recorded from CA1 in a rat



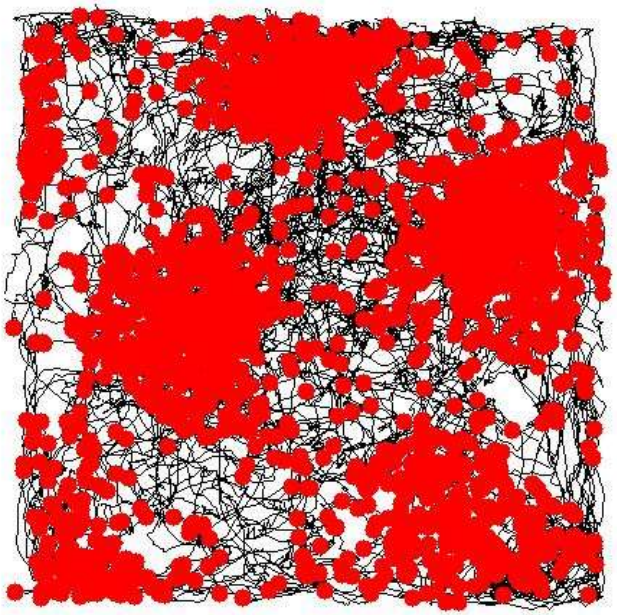
Hippocampus



Phenomenally interesting anatomical structure; crucial for forming representations

Grid cells

- Like place cells, but embed Euclidean space assumptions
- Encode spatial firing fields at equal distances from neighbors
- As if neurons are sensitive to an underlying triangular coordinate system

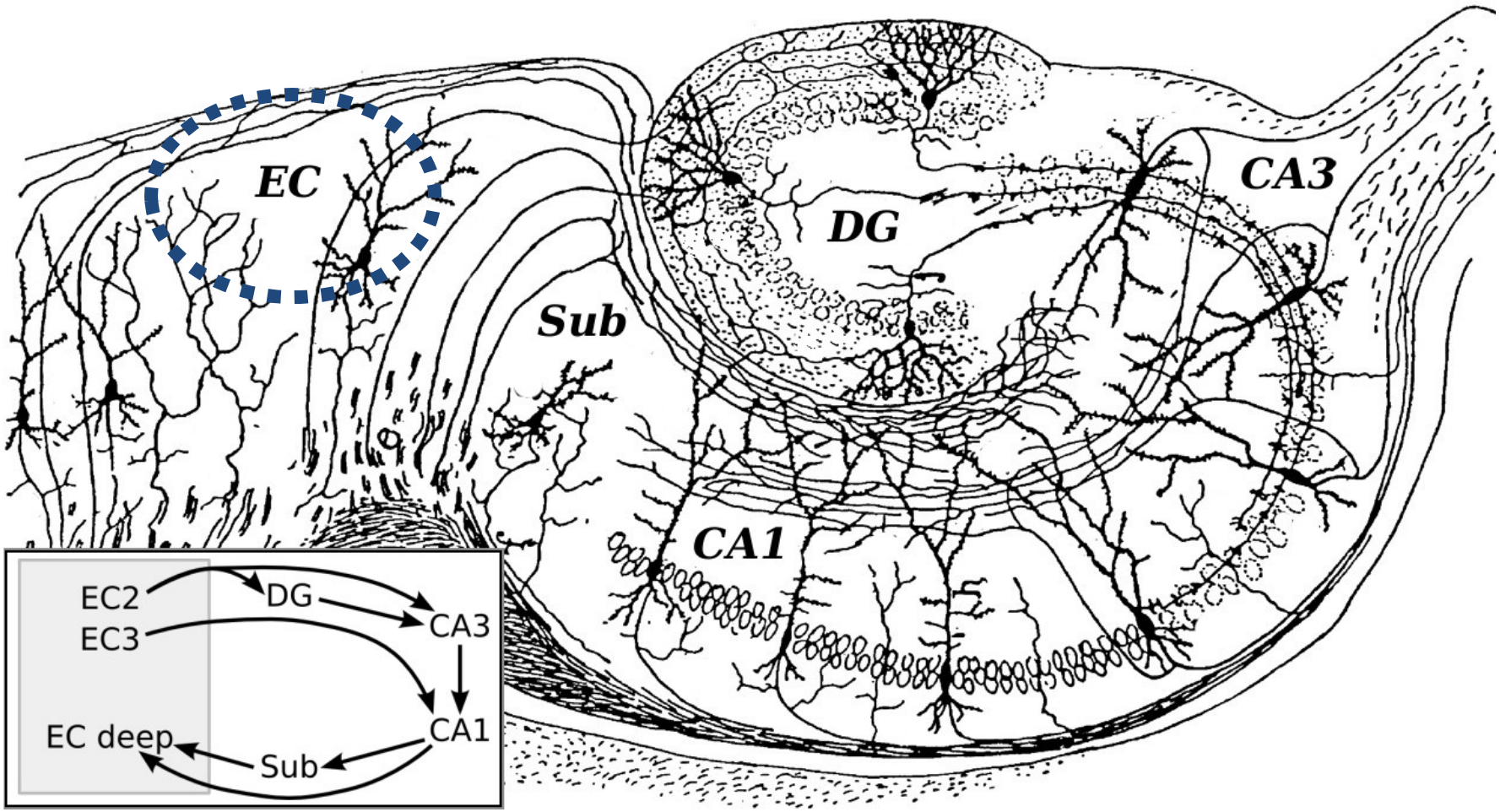


Red dots indicate location of rat in physical space when the grid cell fires

Grid cells are amazing

- https://www.youtube.com/watch?v=dgN5j16scj4
- Reflective of intrinsic coordinate geometry embedded in our hippocampus

Hippocampus

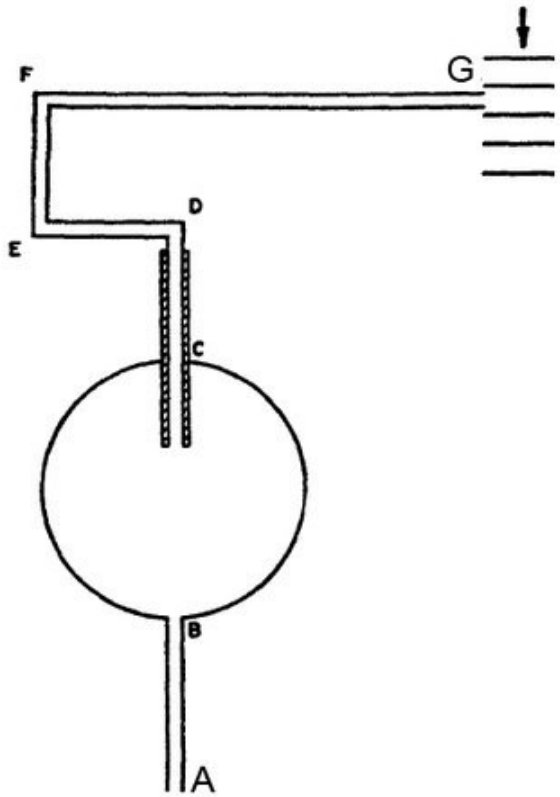


Phenomenally interesting anatomical structure; crucial for forming representations

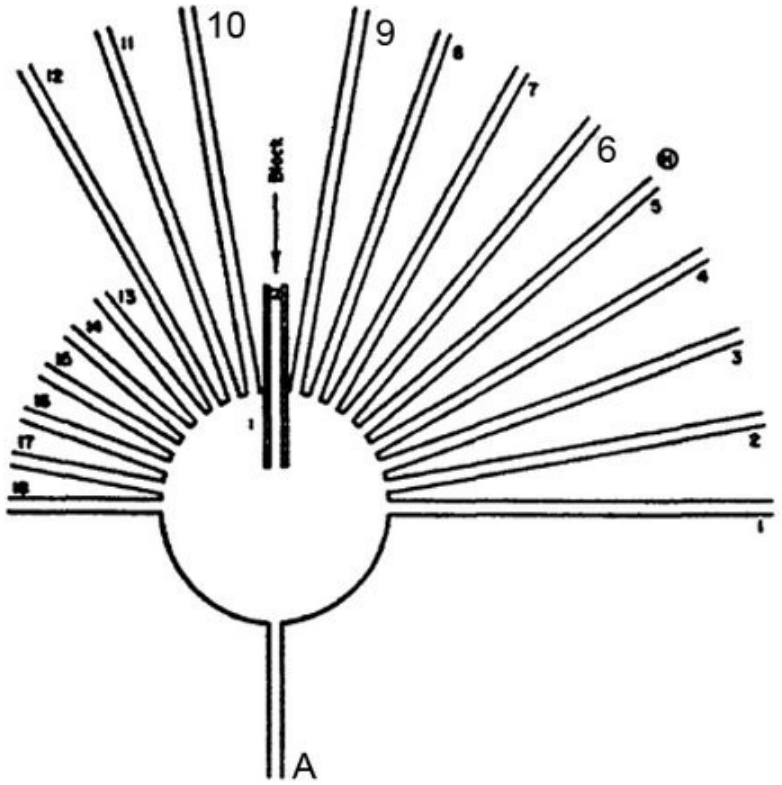
Head direction cells

- Fire when animal's head turns in a particular direction
  - Tend to lead the actual head movement by about 100 ms
- HD system interacts with place cells to generate spatial map of environment?
- Operates coherently during REM sleep

Cognitive maps in rats and men



Apparatus used in preliminary training



Apparatus used in the test trial

(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and short-cut. *J. exp. Psychol.*, 1946, **36**, p. 17.)

Revisiting an old debate

- Model free learning → learn stimulus-response mappings = habits
- What about goal-based decision-making?
  - Do animals not learn the physics of the world in making decisions?
- Model-based learning → learn what to do based on the way the world is currently set up = thoughtful responding?
- People have argued for two systems
  - Model-based systems are difficult to learn
  - Model-free systems don't generalize well

The successor representation

- Remember the value iteration equation?

$$V^{\pi}(s) = \sum_a \pi(a|s)[R(s,a) + \sum_{s'} P(s'|s,a)\gamma V^{\pi}(s')]$$

- Peter Dayan showed long ago that it could be rewritten as

$$V^{\pi}(s) = \sum_{s'} M^{\pi}(s,s') \sum_a \pi(a|s') R(s',a)$$

- Where

$$M^{\pi}(s,s') = \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t \mathbb{I}(s_t = s') \ | S_0 = s],$$

- The successor representation offers one explanation for how the TD signal could yield model-based policies

The successor representation

- The big problem with model-free RL

$$\delta_t = r_{t+1} + \gamma V^{\pi}(s_{t+1}) - V^{\pi}(s_t)$$

$$\Delta V^{\pi}(s_t) = \alpha \, \delta_t$$

- Changes in reward distributions entail propagate very slowly across action chains
- SR-based RL tries to fix this

$$\delta_t^{\mathrm{SR}} = \mathbb{I}(s_t = s') + \gamma \, M(s_{t+1}, s') - M(s_t, s')$$

$$\Delta M(s_t, s') = \alpha \, \delta_t^{\mathrm{SR}}$$

https://julien-vitay.net/post/successor_representations/

Benefits over model-free learning

- SR based learning uses a sensory prediction error to guide learning
  - Rather than look at the discrepancy between how rewarding I find a state, and how much I expected it to be
  - I look at whether I am arriving in the state more frequently than I expected, while trying to maximizing reward
- Changes in the reward distribution quickly update the state visitation matrix for all states from this state

Are humans SR RL learners?

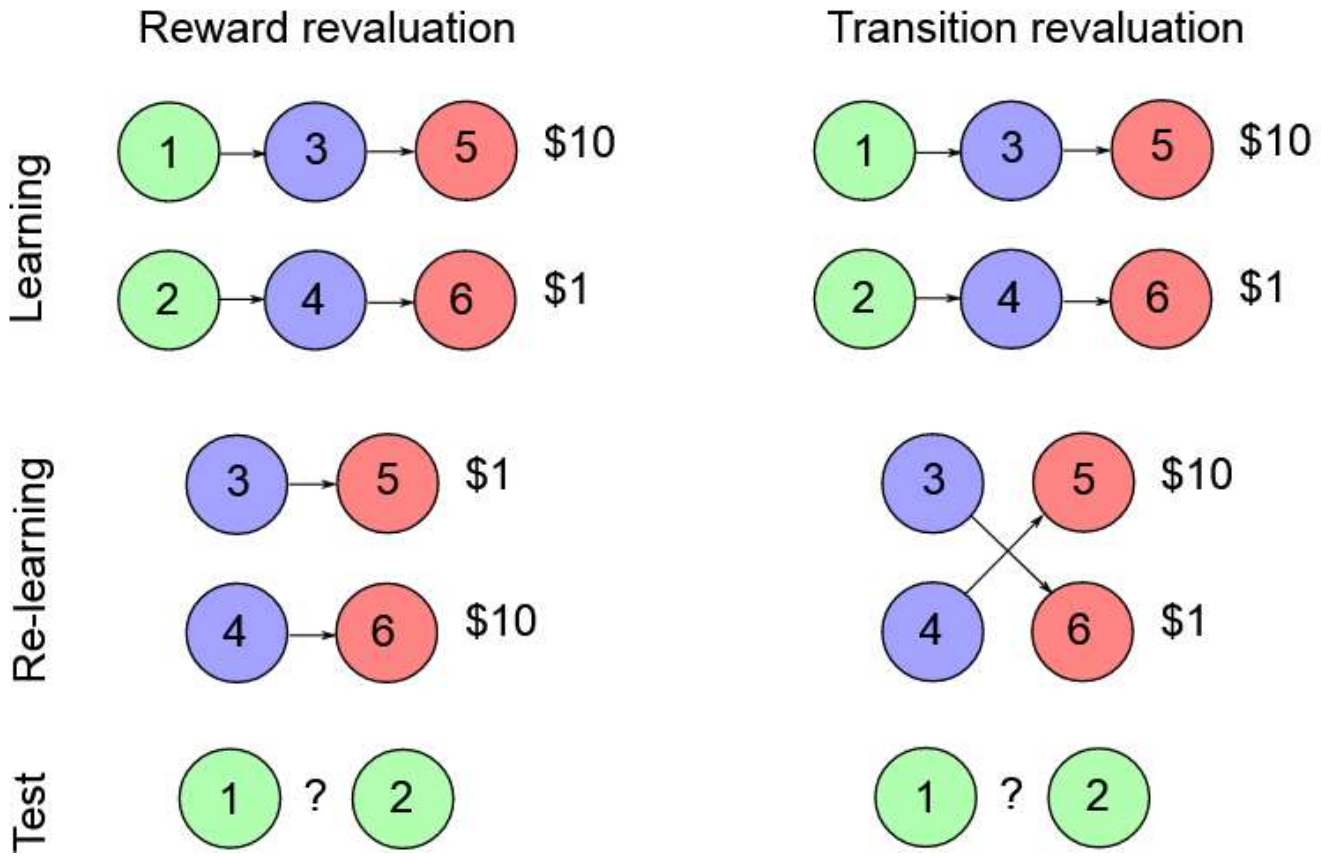


FIGURE 1: Two-steps sequential learning task of Momennejad et al. (2017).

What should happen?

- Model-free RL would predict no change after both reward and transition re-learning
- Model-based RL would predict a change in both conditions
- SR-based RL would predict a change in the reward relearning condition, but not in the transition relearning condition

What actually happens

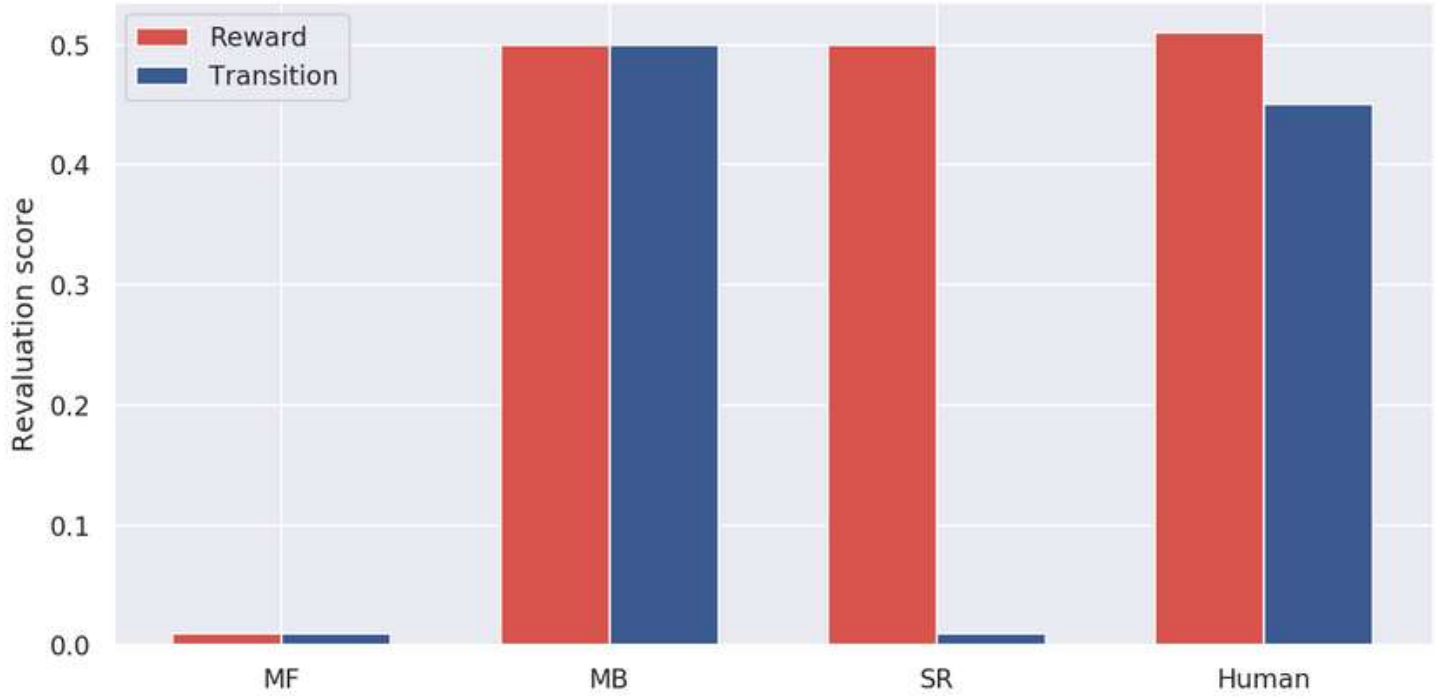
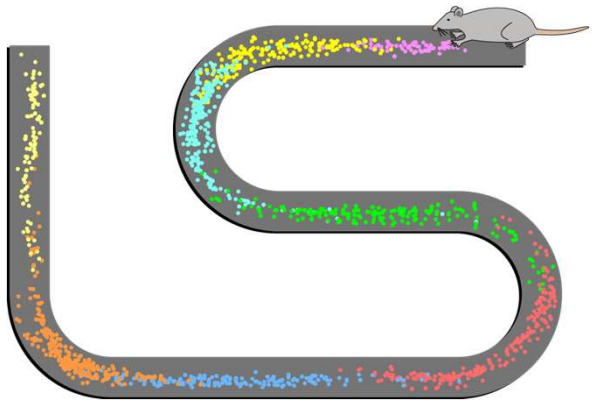


FIGURE 2: Revaluation score in the reward (red) and transition (blue) revaluation conditions for the model-free (MF), model-based (MB), successor representation (SR) and human data as reported in Momennejad et al. (2017).

Humans seem pretty model-based in this task, though SR seems to be mildly implicated

A neuroscientific hypothesis

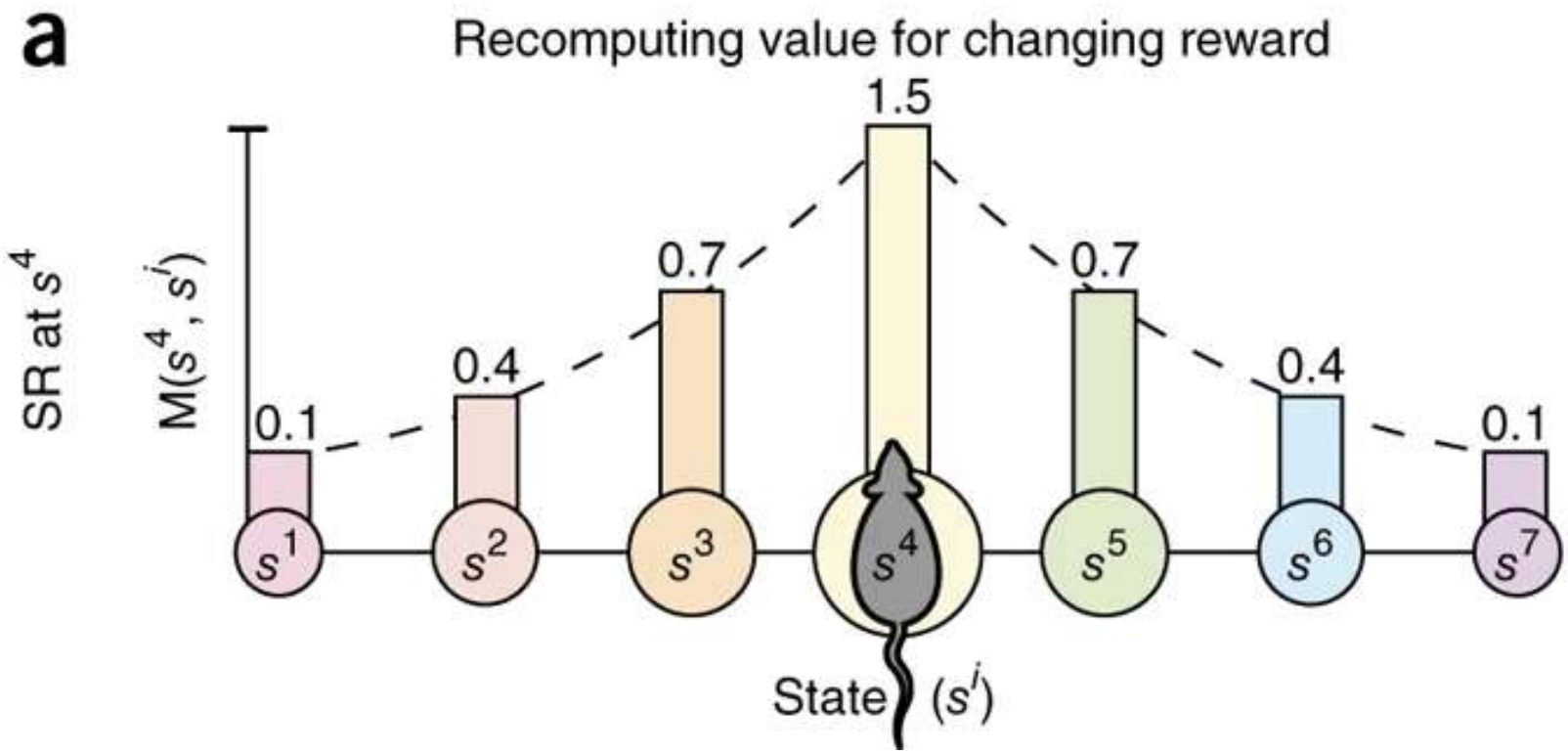


- Place cells have historically been seen to represent an animal's sense of place
- Stachenfeld et al (2017) propose that they encode the SR matrix for a task
- Representing a sense of where they are likely to go next

https://www.nature.com/articles/nn.4650

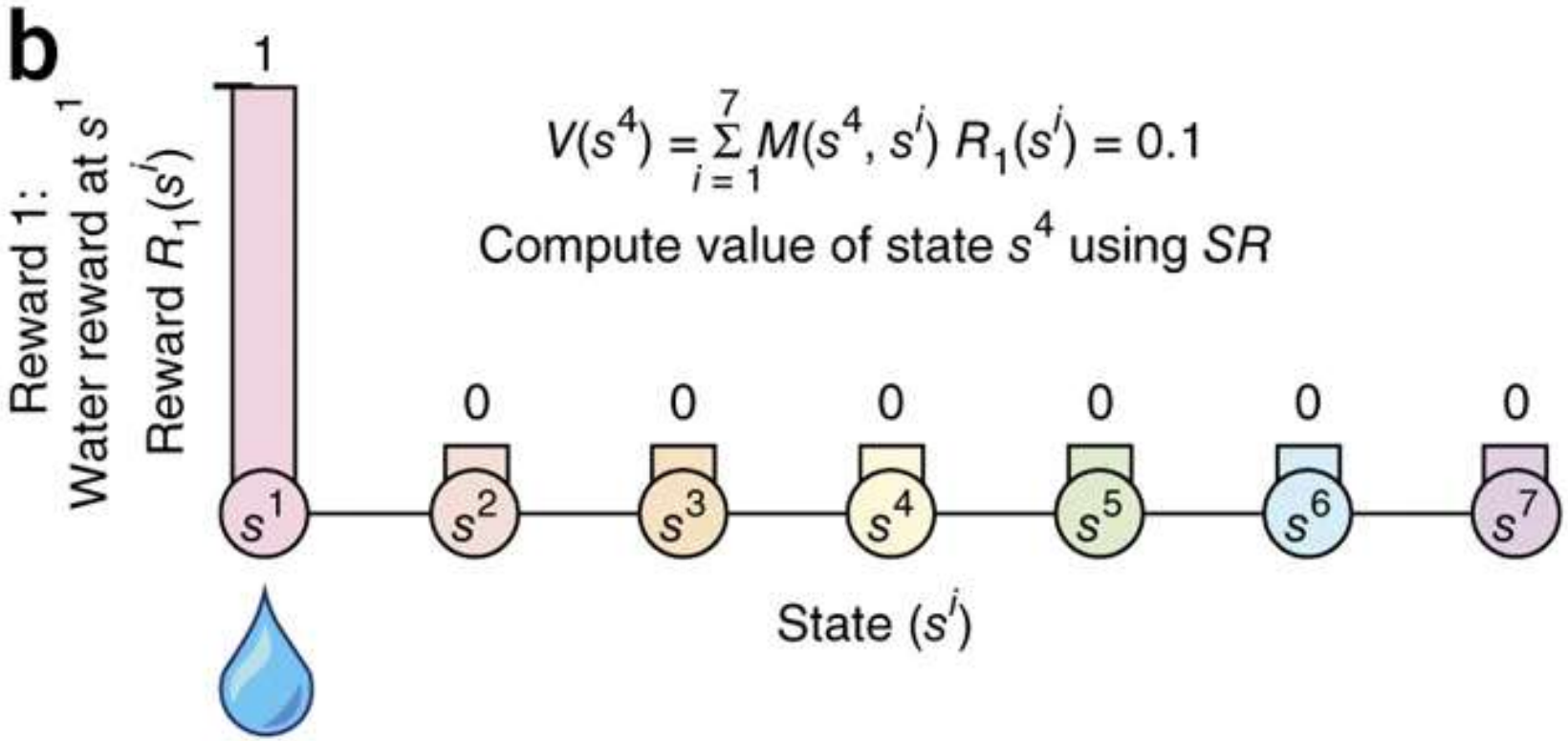
The model

- We assume that value is updated with the SR when reward changes



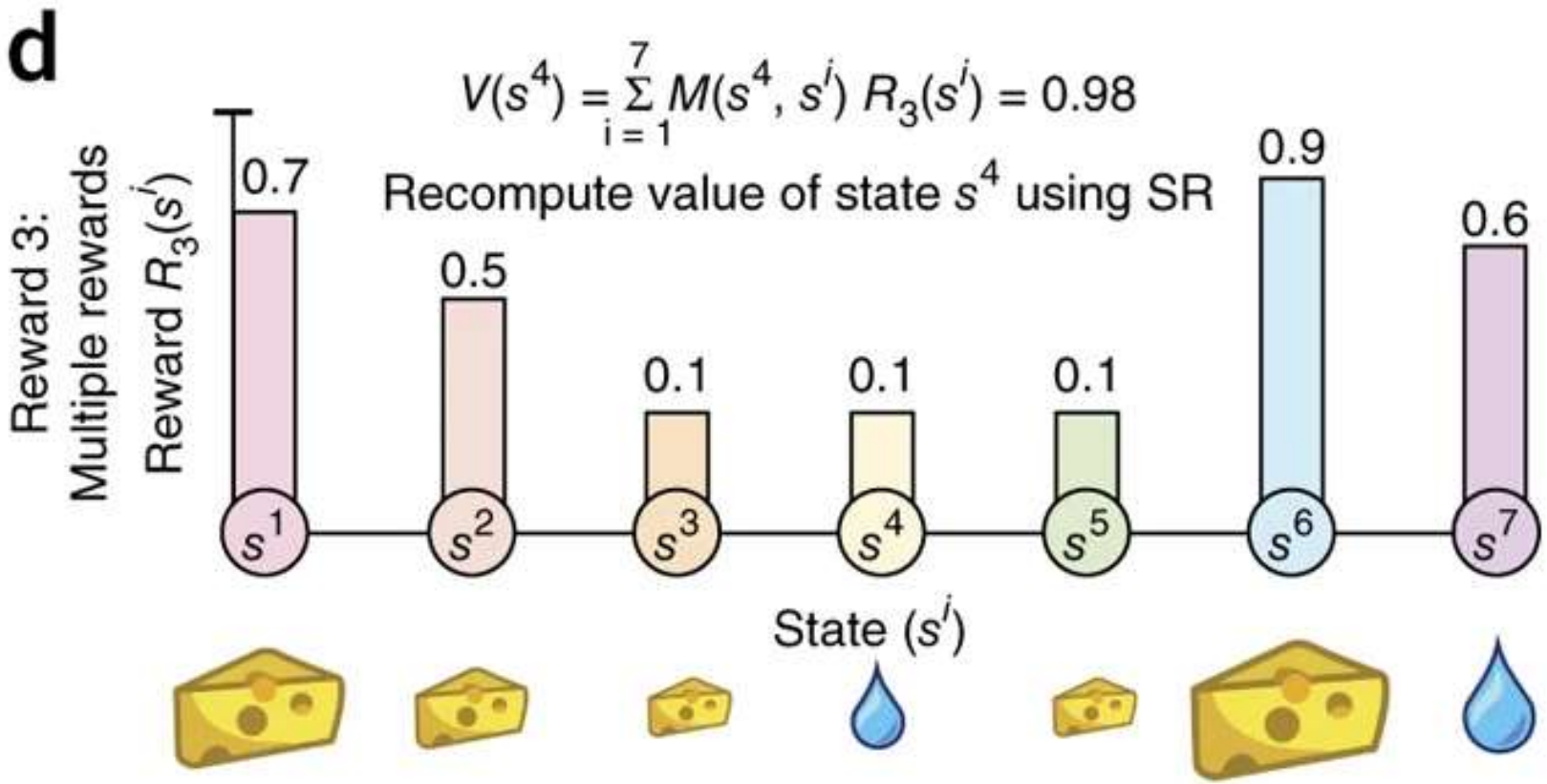
The model

- Halfway house between model free and model-based learning

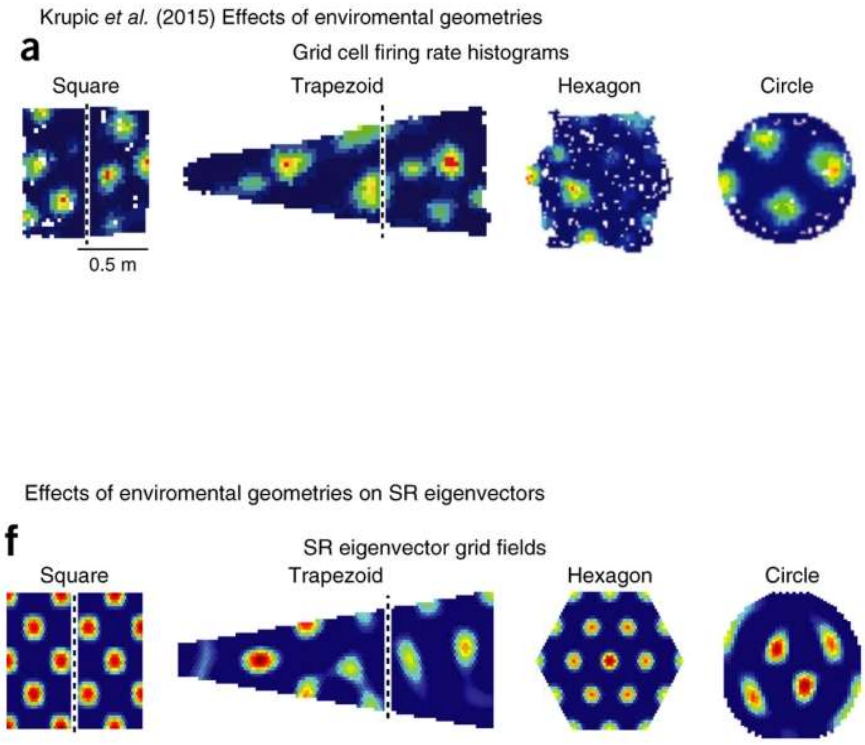


The model

- Value functions can be recomputed without changing the SR



Consistent with existing data on place cell behavior



- Model accommodates deviations from Euclidean geometry seen in actual rat experiments
- Also explains several non-spatial task results drawn from fMRI imaging
- A subtle hypothesis that is consistent with a primarily predictive role for hippocampus