The hippocampus as a predictive map

Kimberly L Stachenfeld[1,2], Matthew M Botvinick[1,3] & Samuel J Gershman[4]

Noémi Éltető

Cognitive Maps Seminar

30.11.2022

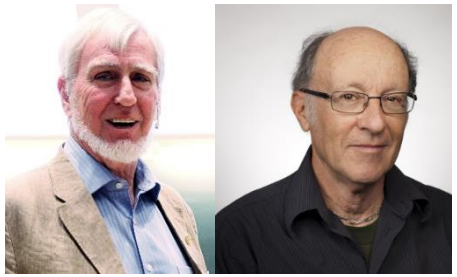

SCYTHICVS OCEANVS
EUROPA
ASIA
AFRICA
SARMATIA
GERMANIA
GALLIA
HISPANIA
ITALIA
ARMENIA
SYRIA
PERSIA
INDIA
SERES
SCYTHIA
ARABIA
AETHIOPES
MARE MEDITERRANEUM
MARE ATLANTICVM
MARE RVBRVM
PALVS MAEOTIS
EOUS OCEANS
INDICVS OCEANVS

It is clear that whatever the status of relative space in physics and mathematics, it cannot be given ontological priority within psychology. [...] there is a clear need for the concept of **unitary space**. Further, it appears that this framework **cannot be acquired through experience**; it must be available soon after birth, for the processes of localization, identification and the coherent organization of experience depend on it.

The Hippocampus as a Cognitive Map (1978)

Paraphrased:
Cognitive maps reflect **predictions** that are based on **prior experience**.

The Hippocampus as a Predictive Map (2018)



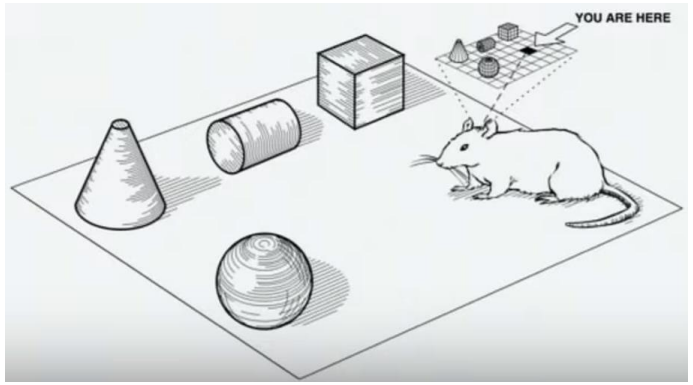


Image taken from Sam Gershman’s talk on youtube

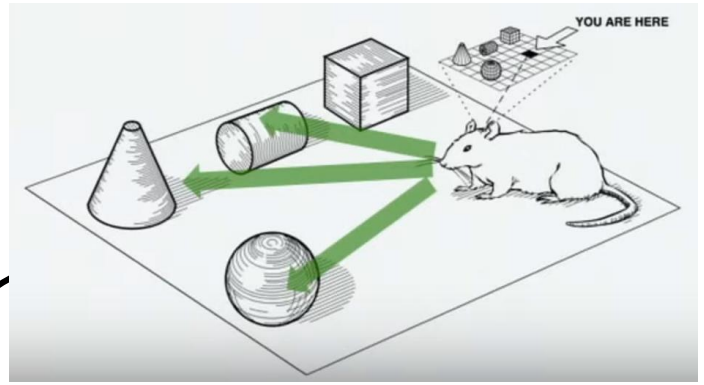
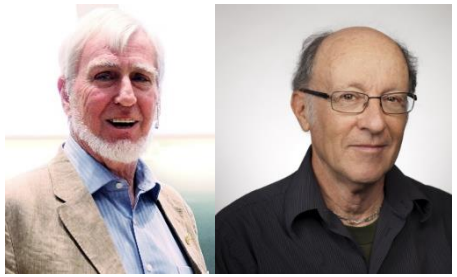
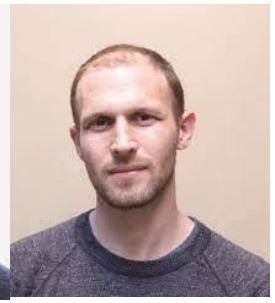


Image taken from Sam Gershman’s talk on youtube



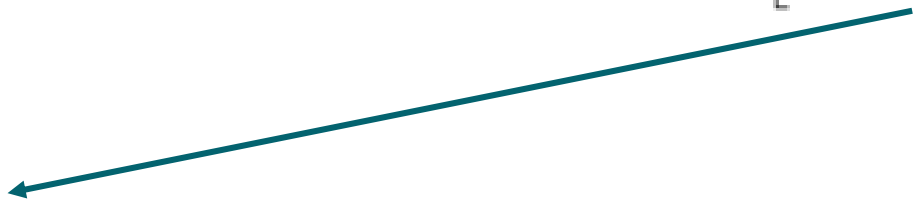


How to compute value for navigation?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$

How to compute value for navigation?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$

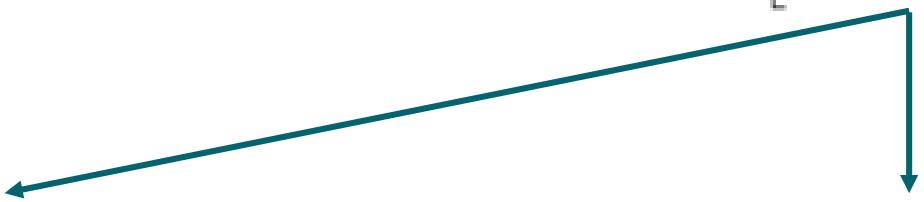


Model-free

$Q(s,a)$

How to compute value for navigation?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$



Model-free

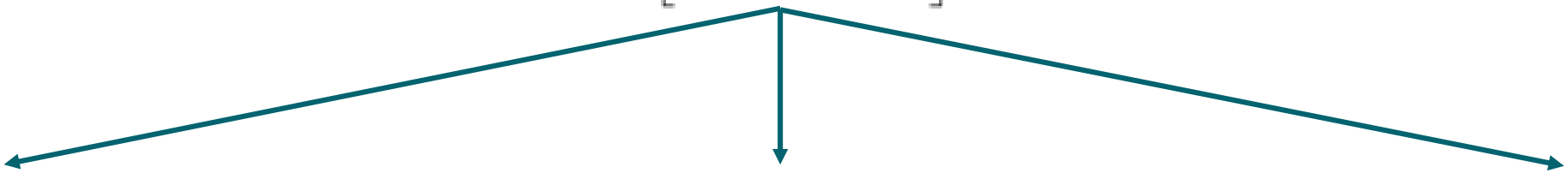
$Q(s,a)$

Model-based

$$v_{\pi}(s)=\sum_{a}\pi(a|s)\sum_{s',r}p(s',r\,|\,s,a)\Big[r+\gamma v_{\pi}(s')\Big]$$

How to compute value for navigation?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)|s_0=s\right]$



Model-free

$Q(s,a)$

Model-based

$$v_{\pi}(s)=\sum_{a}\pi(a|s)\sum_{s',r}p(s',r|s,a)\left[r+\gamma v_{\pi}(s')\right]$$

Successor representation (SR)
(Dayan, 1993)

$$v_{\pi}(s)=\sum_{s'}M(s,s')R(s')$$

$$M(s,s')=\mathbb{E}\left[\sum_{t=0}\gamma^{t}\mathbb{I}\left(s_t=s'\right)|s_0=s\right]$$

How to compute value for navigation?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)|s_0=s\right]$

Model-free

$Q(s,a)$

Model-based

$$v_{\pi}(s)=\sum_{a}\pi(a|s)\sum_{s',r}p(s',r|s,a)\left[r+\gamma v_{\pi}(s')\right]$$

Successor representation (SR)
(Dayan, 1993)

$$v_{\pi}(s)=\sum_{s'}M(s,s')R(s')$$

$$M(s,s')=\mathbb{E}\left[\sum_{t=0}\gamma^{t}\mathbb{I}\left(s_t=s'\right)|s_0=s\right]$$

If we think further, is it possible that any information in us is saved with a "value", i.e. that we only remember things with respect to their possible benefit/harm for us?

How to compute value for navigation?

Maximize this: $\mathbb{E}_\pi\left[\sum_{t=0}^{\infty}\gamma^t R(s_t)\,|\,s_0=s\right]$

Model-free

$Q(s,a)$

value-based

Model-based

$$v_\pi(s)=\sum_a \pi(a|s)\sum_{s',r} p(s',r|s,a)\left[r+\gamma v_\pi(s')\right]$$

value-free

Successor representation (SR)
(Dayan, 1993)

$$v_\pi(s)=\sum_{s'} M(s,s')R(s')$$

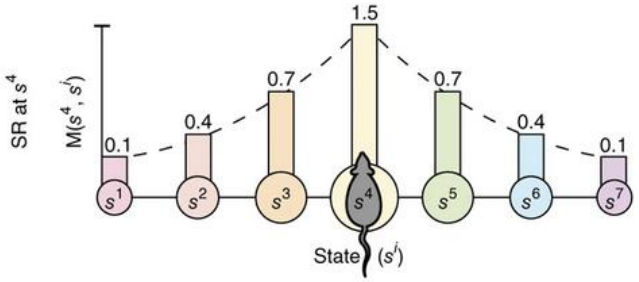
$$M(s,s')=\mathbb{E}\left[\sum_{t=0}^{}\gamma^t \mathbb{I}\left(s_t=s'\right)|\,s_0=s\right]$$

value-free

If we think further, is it possible that any information in us is saved with a "value", i.e. that we only remember things with respect to their possible benefit/harm for us?

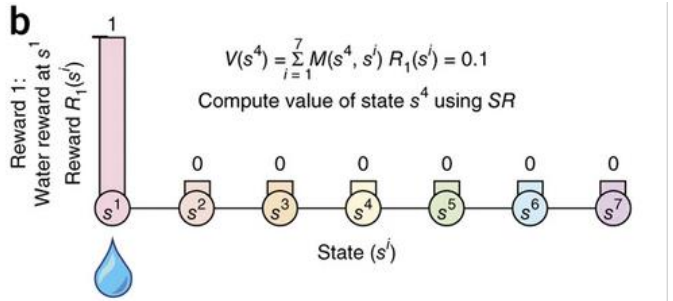
The successor representation (SR): expected discounted future state occupancy

The SR

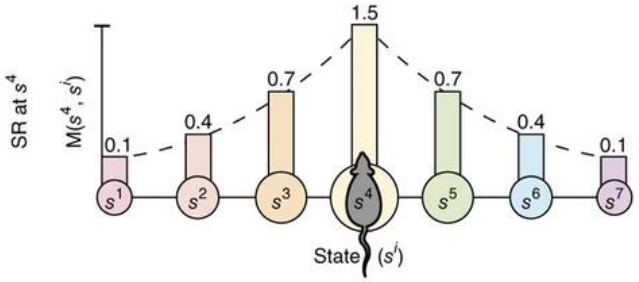


The successor representation (SR): expected discounted future state occupancy

Recomputing value for changing reward

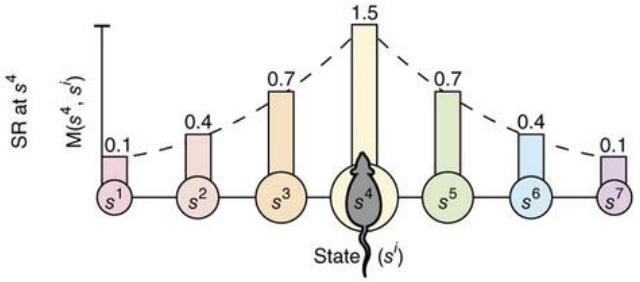


The SR

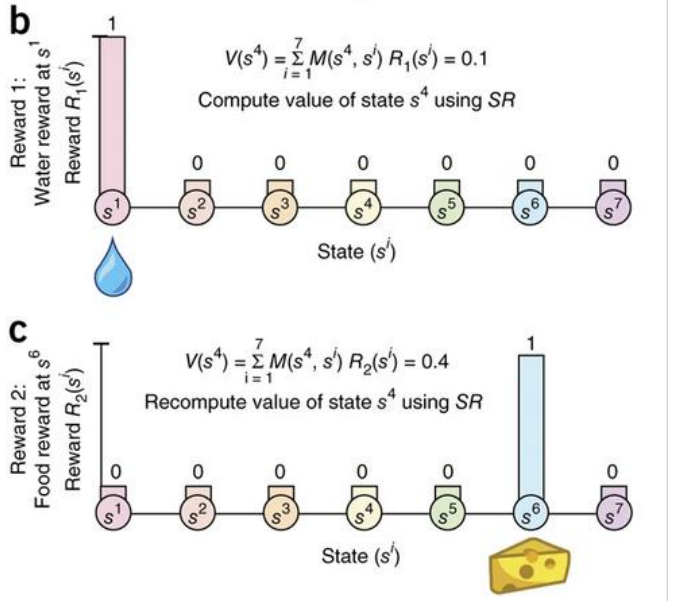


The successor representation (SR): expected discounted future state occupancy

The SR

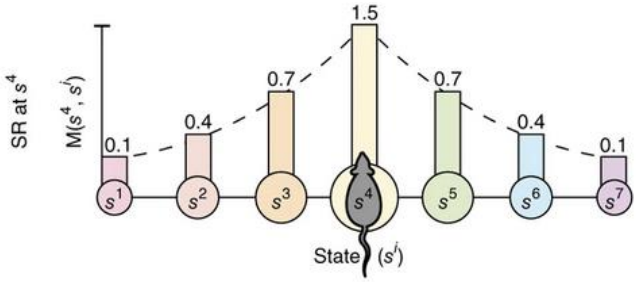


Recomputing value for changing reward

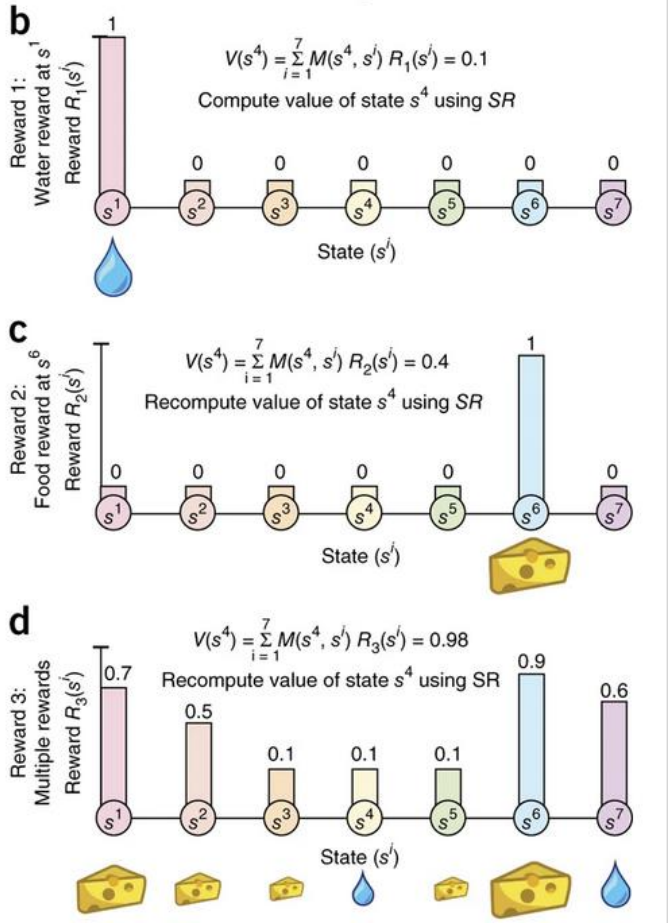


The successor representation (SR): expected discounted future state occupancy

The SR

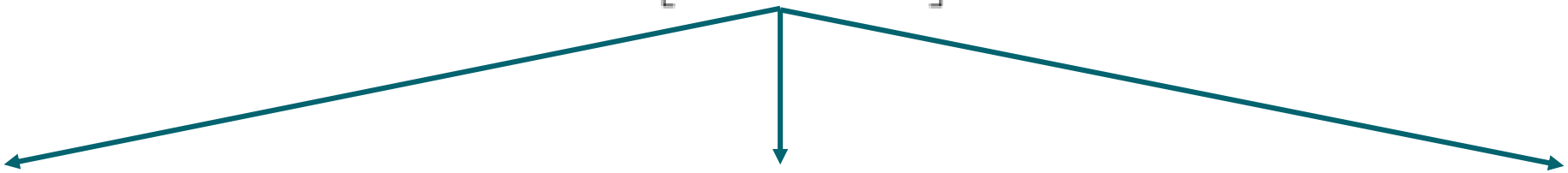


Recomputing value for changing reward



How to compute value for decisions?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$



Model-free

Model-based

Successor representation (SR)

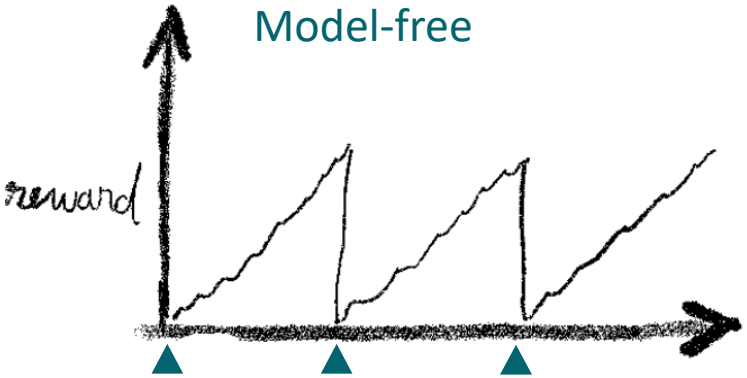
How to compute value for decisions?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$

Model-free

Model-based

Successor representation (SR)

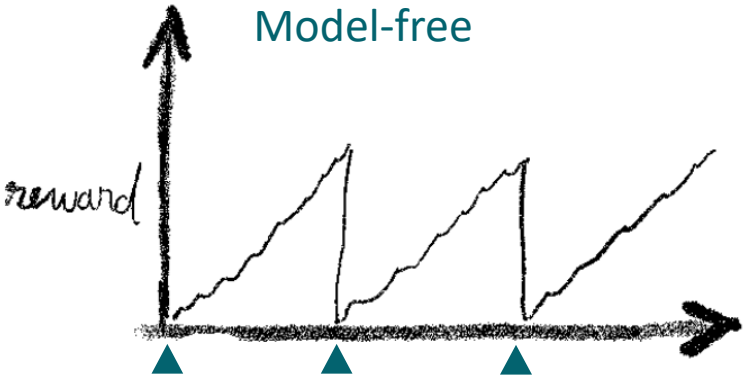


▲: reward function (reward location, motivational state, etc.) changed

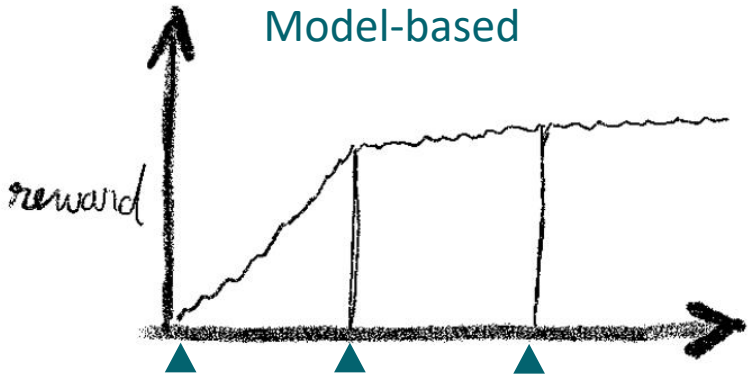
How to compute value for decisions?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$

Model-free



Model-based



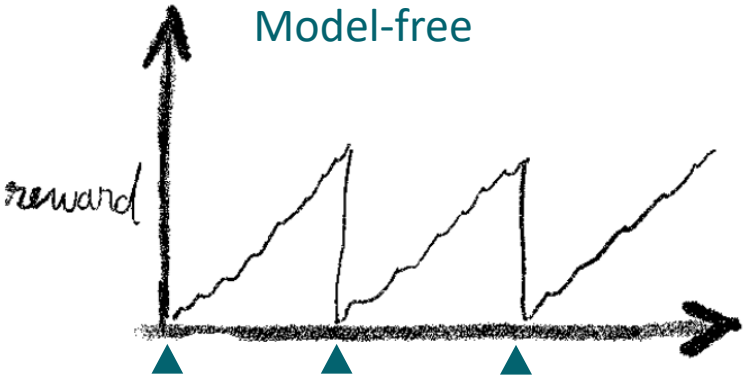
Successor representation (SR)

▲: reward function (reward location, motivational state, etc.) changed

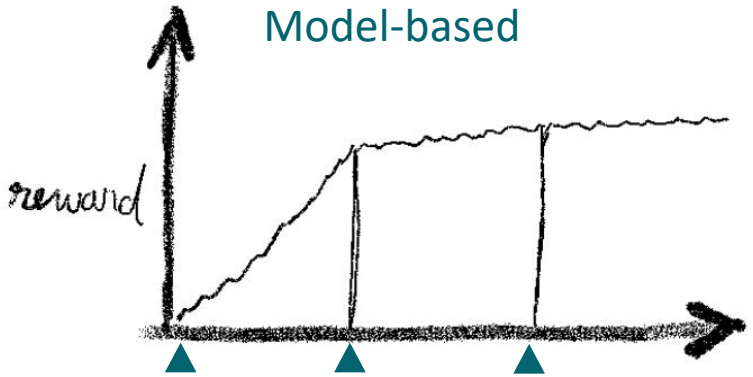
How to compute value for decisions?

Maximize this: $\mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty}\gamma^{t}R(s_t)\,|\,s_0=s\right]$

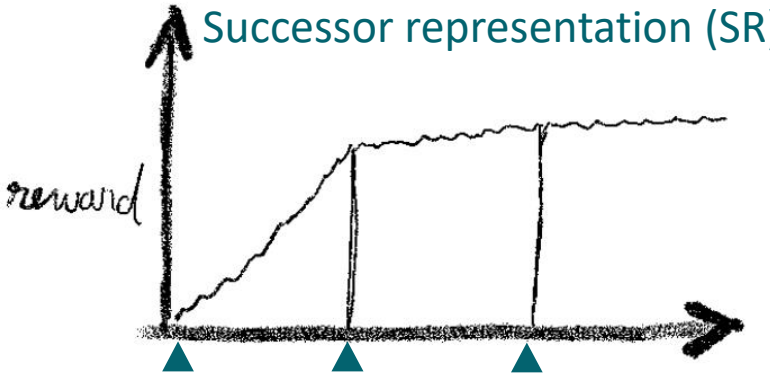
Model-free



Model-based

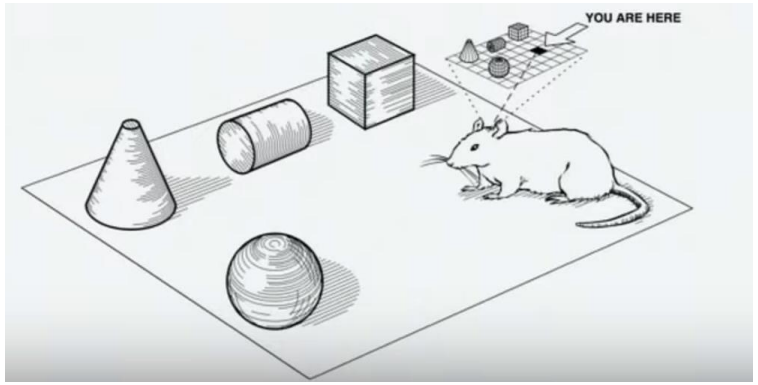


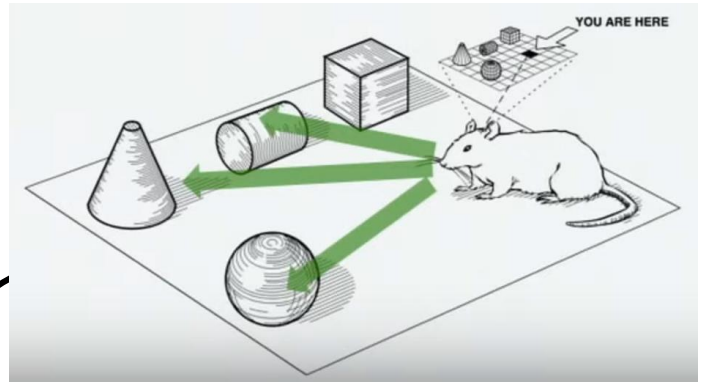
Successor representation (SR)

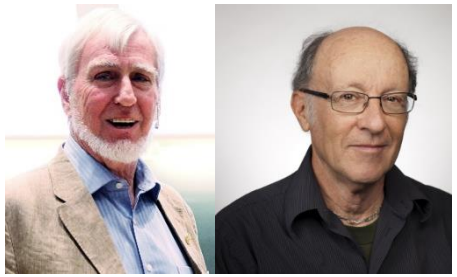


▲: reward function (reward location, motivational state, etc.) changed

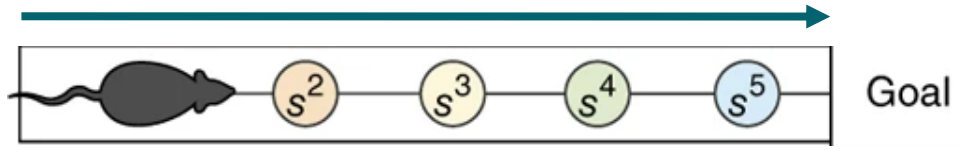
Okay, the SR is interesting but how does the brain encode space?

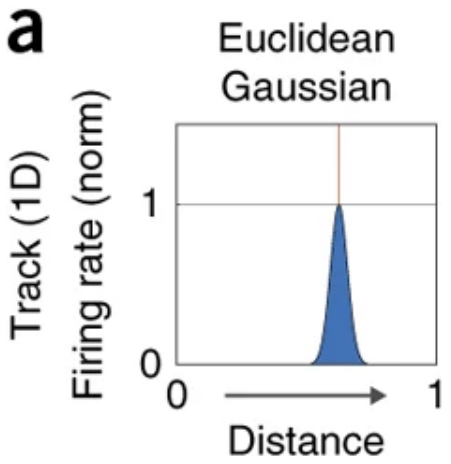

YOU ARE HERE


YOU ARE HERE

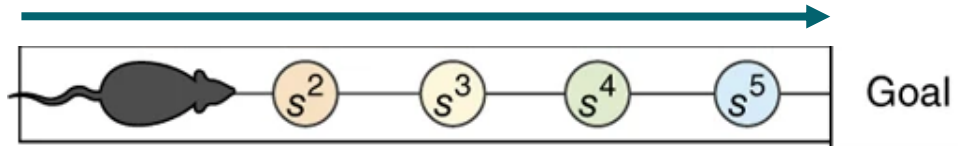


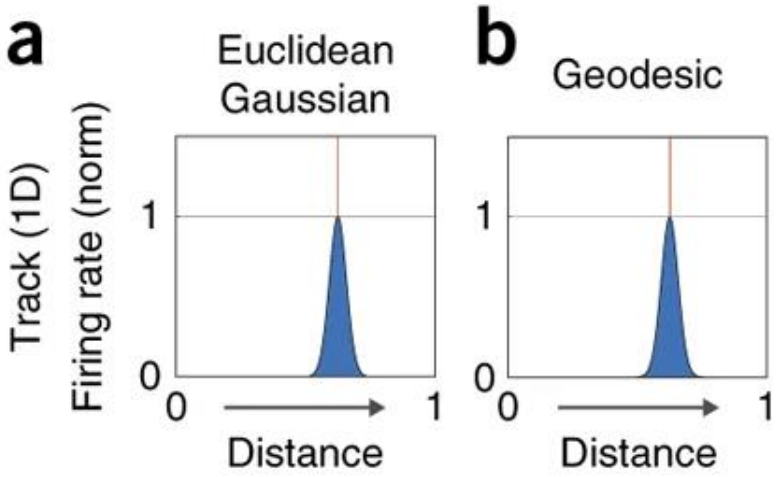
Hippocampal encoding of the SR



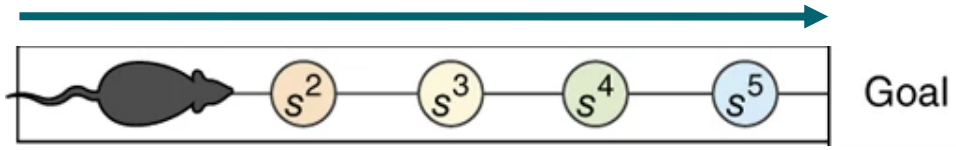


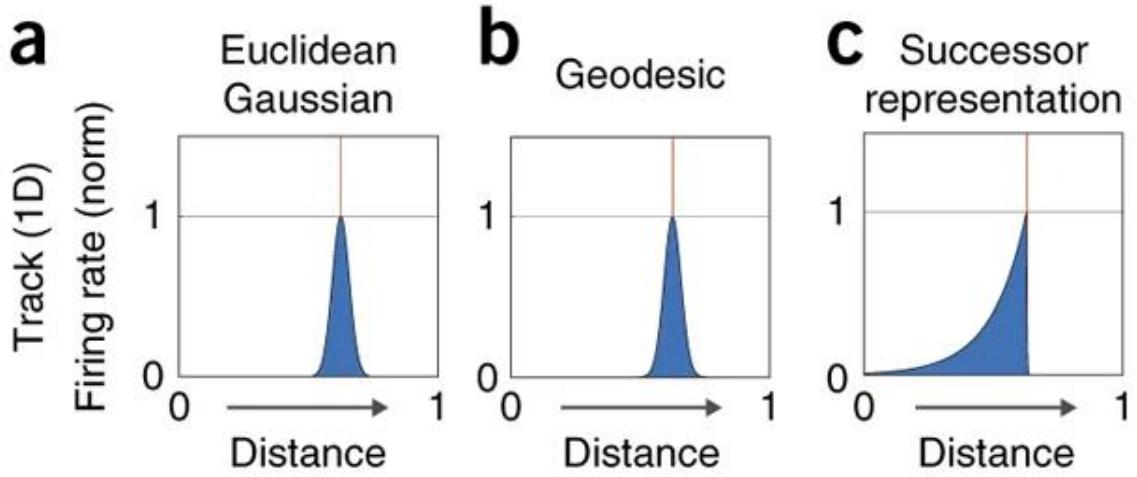
Hippocampal encoding of the SR

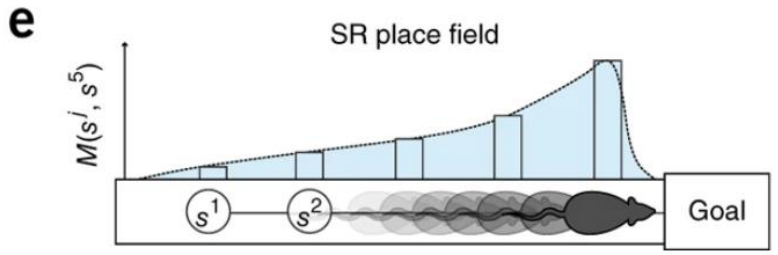




Hippocampal encoding of the SR



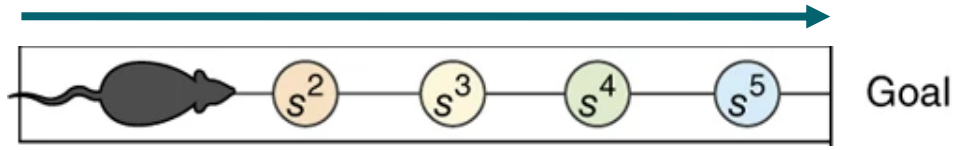


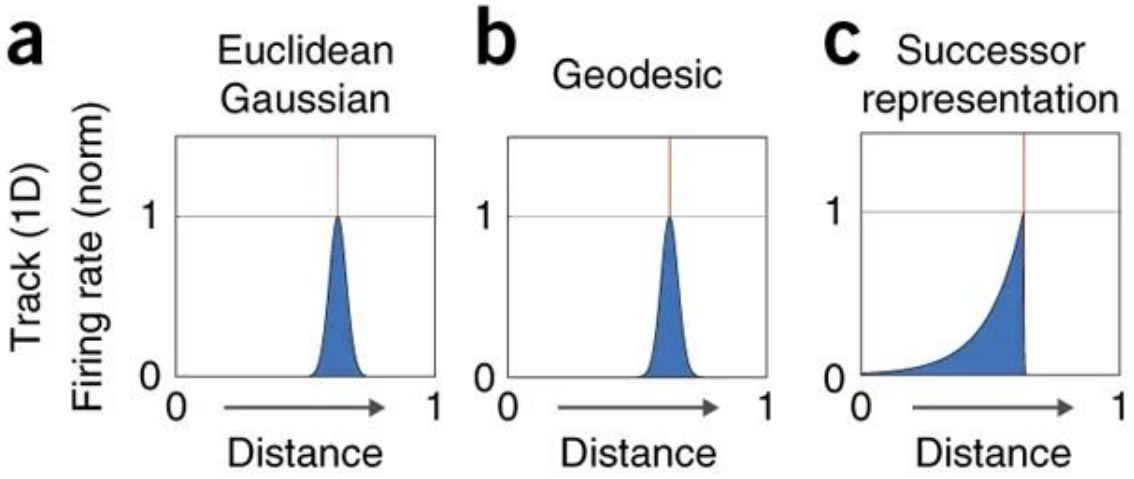


$$M(s,s') = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \mathbb{I}(s_t = s') \mid s_0 = s\right]$$

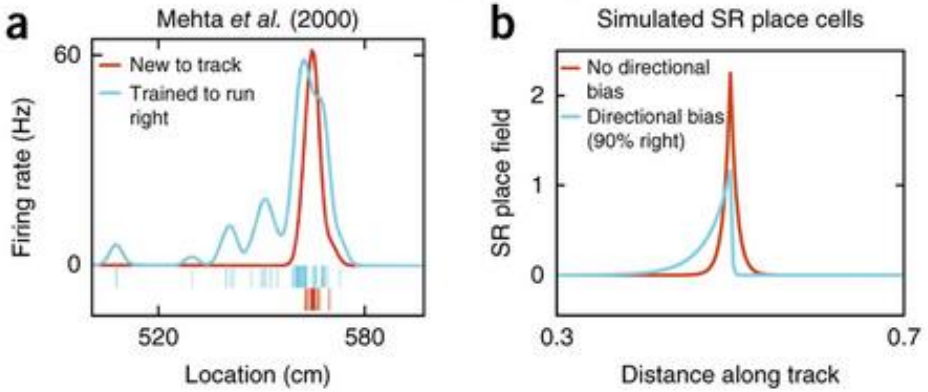
Following the defination: the SR encodes the expected discounted future occupancy of state s' along a trajectory initiated in state s. Does M(s,s') always peak at s' (when s= s' for a specific s'), like in Figure 2e?

Hippocampal encoding of the SR

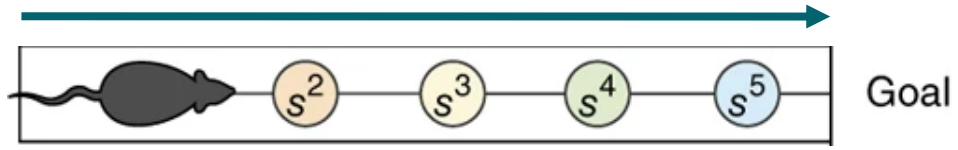


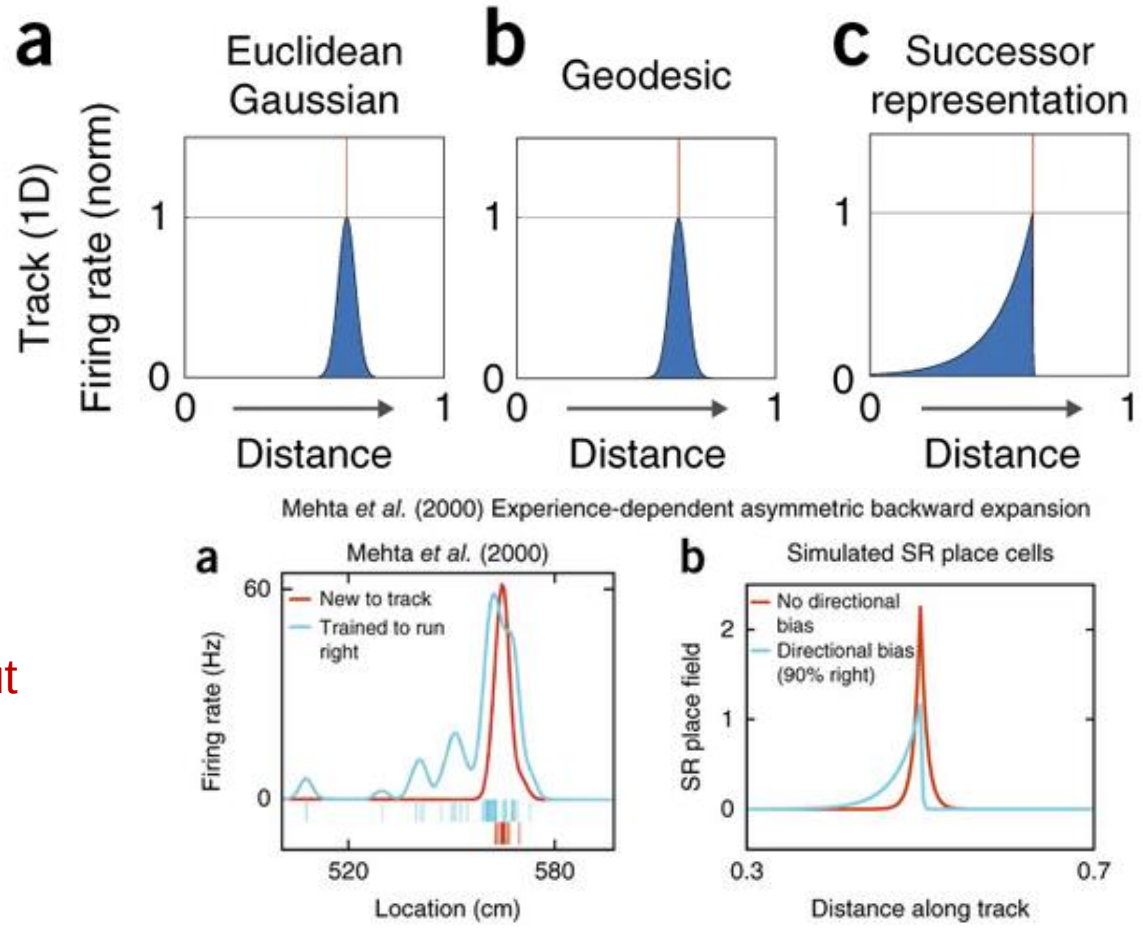


Mehta *et al.* (2000) Experience-dependent asymmetric backward expansion



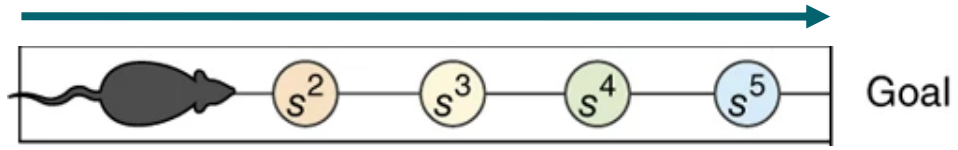
Hippocampal encoding of the SR

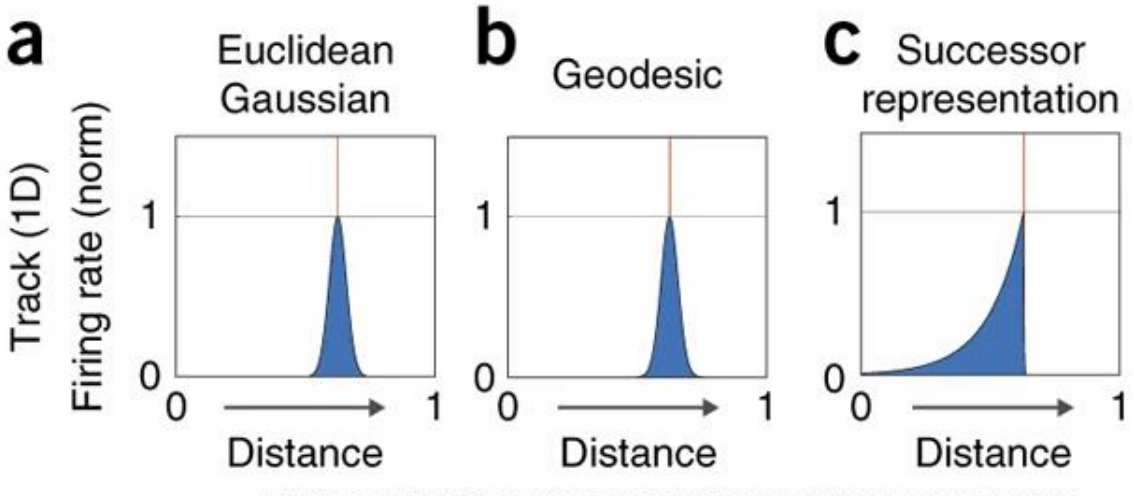




What intrigued me the most was the fact that SR are encoded by population firing rather than single neuron activity. Are there any experiments (or can we figure out any) that would support this hypothesis?

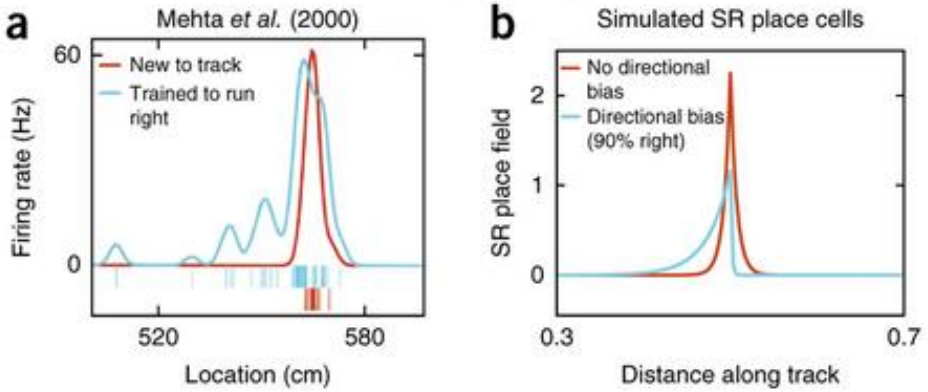
Hippocampal encoding of the SR



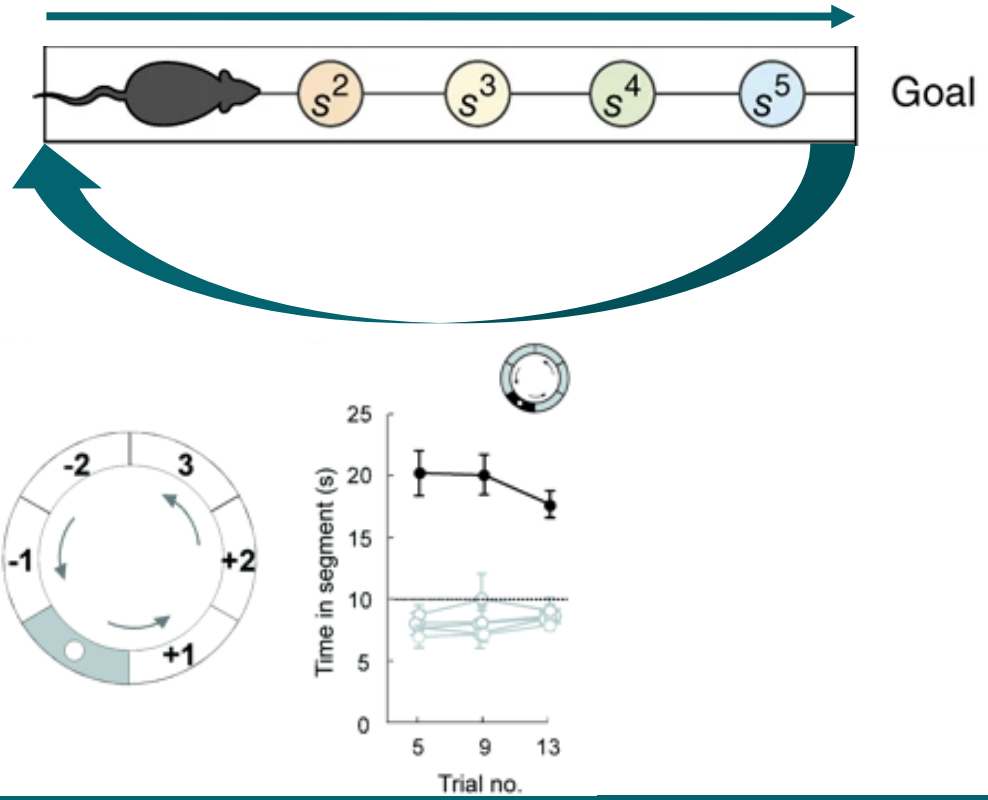


Mehta *et al.* (2000) Experience-dependent asymmetric backward expansion

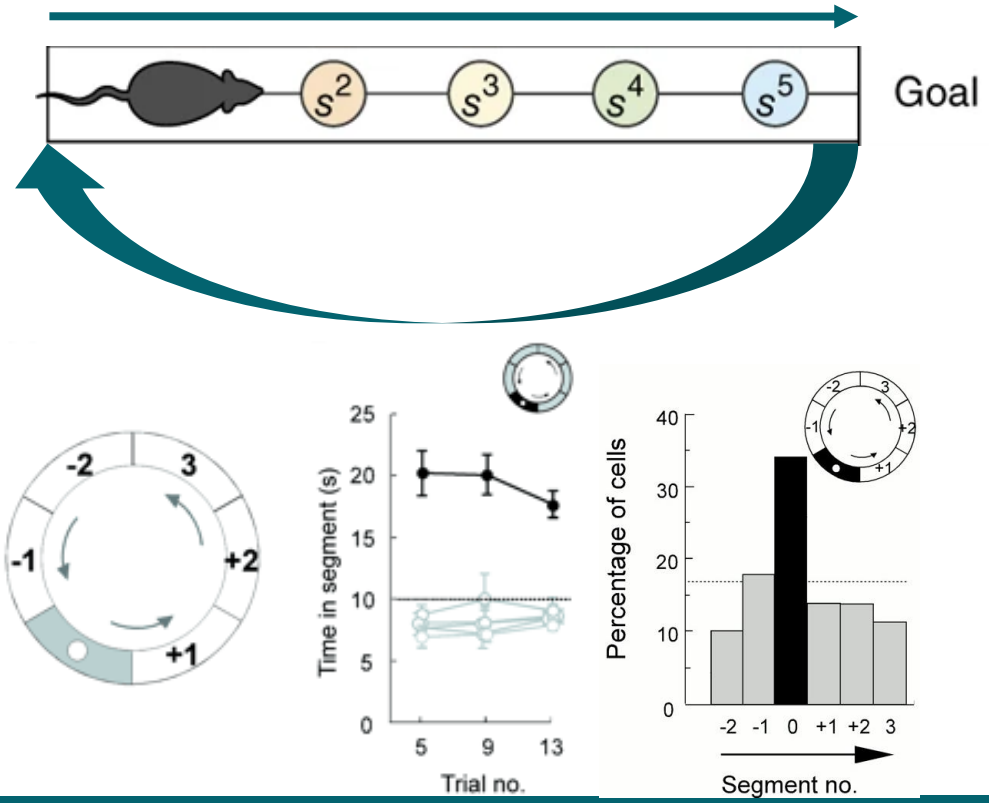
The SR model explains formation of place fields assuming that a fixed optimal policy is being followed (through predictive policy evaluation). How would the model look like if the agent had to concurrently identify the optimal policy (policy iteration/improvement)? Would the same results be obtained in regards to SR place field structure?



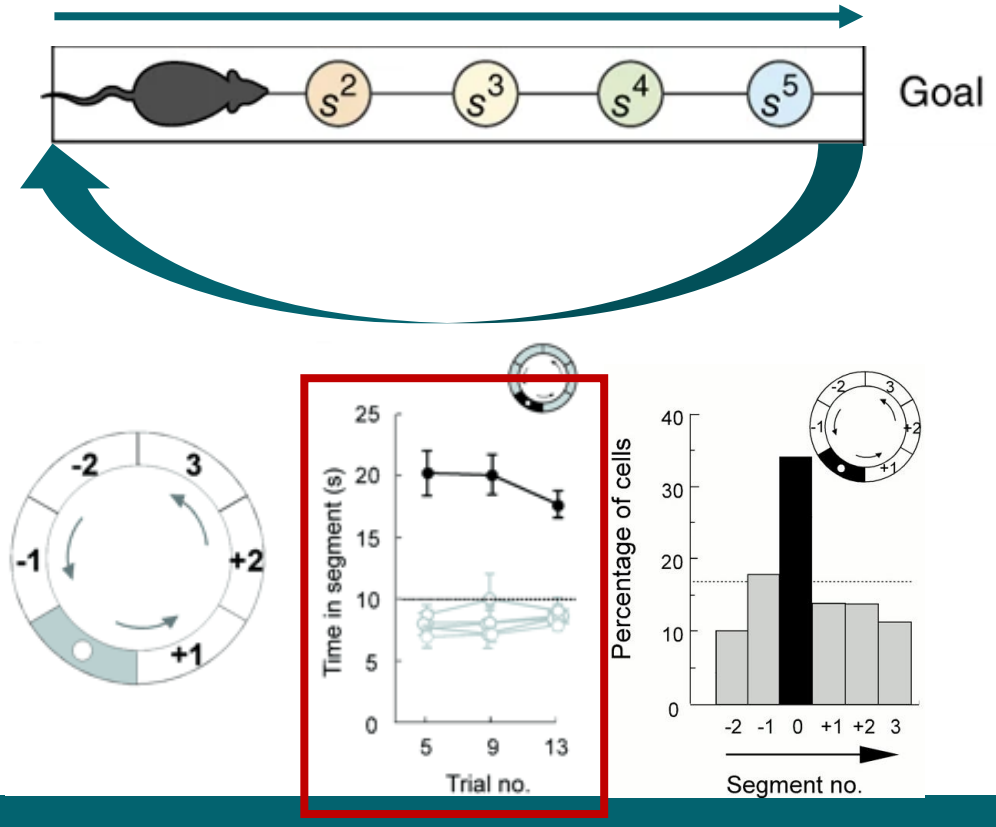
Hippocampal encoding of the SR



Hippocampal encoding of the SR



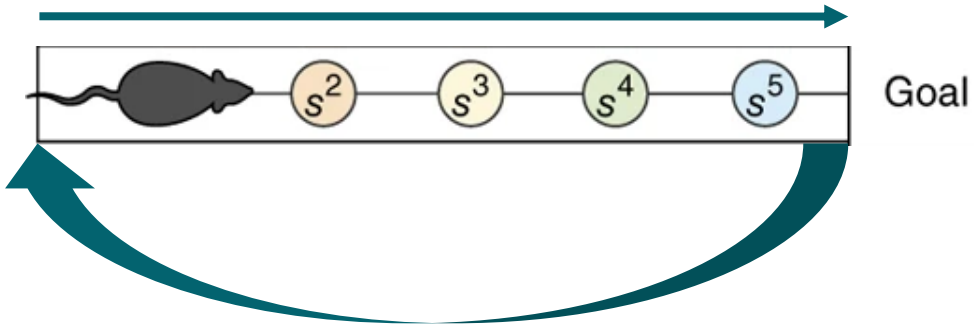
Hippocampal encoding of the SR

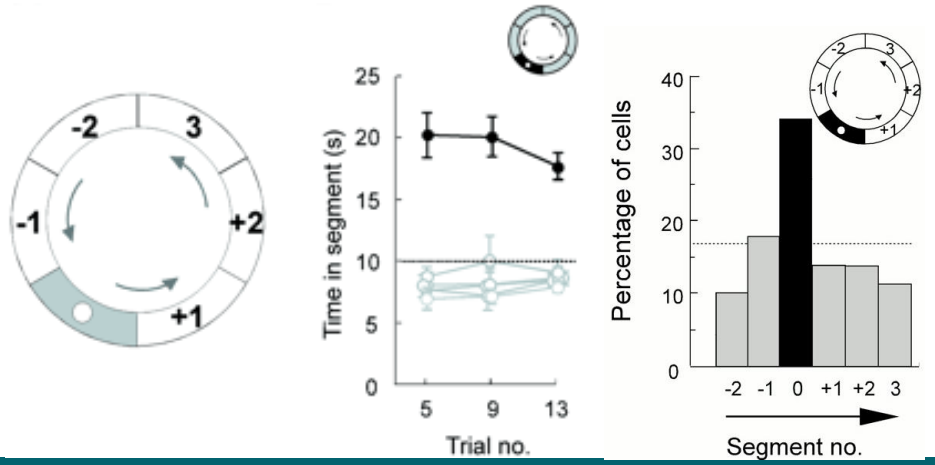


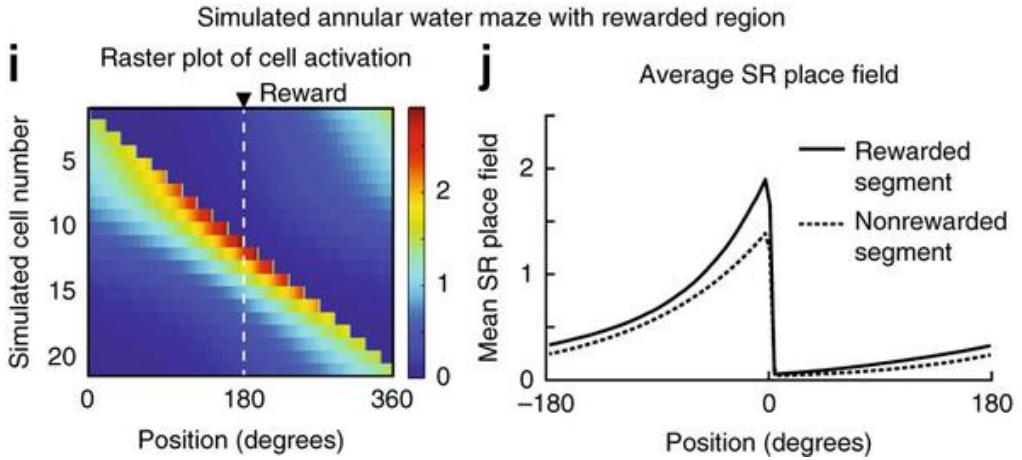
I do not understand why the SR predicts that place fields should be larger at reward locations. From my understanding, a place cell for location s* is large, if even locations s that are far away increase the firing rate. This however should be the case for all place fields s* that are on the policy path, as it is very likely that the place cell is visitied in the future (as opposed to visiting them randomly as is the case for other place cells). And not just for place cells near the reward. This is also reflected in the formulas, as M (which the place cells reflect) is independent of the reward R, and instead dependent on the policy. (Unrelated side note: Could you also expand a bit on how the grid fields reflect a low-dimensional eigendecomposition of the SR?)

"The transition policy was such that the animal **lingered longer near the rewarded location** and had a preferred direction of travel."

Hippocampal encoding of the SR



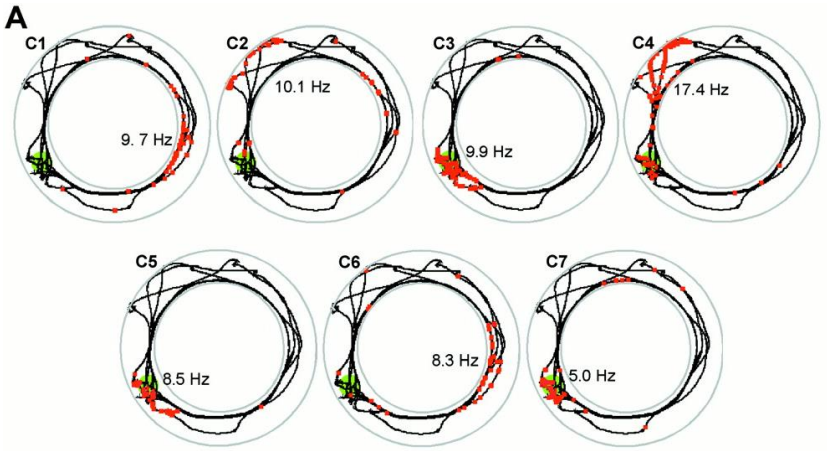


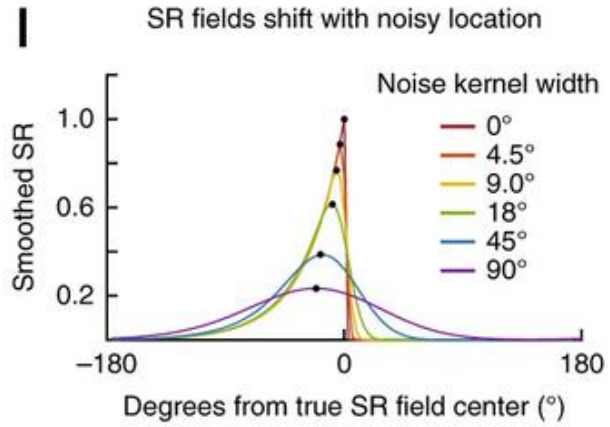


Hippocampal encoding of the SR

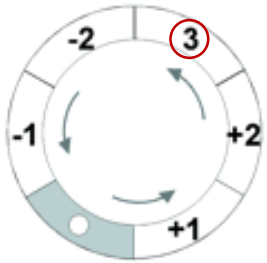
While SR predicts that place fields should be larger near the reward locations, that doesn't match experimental data observed in Hollup et al. (place fields had the same size at rewarded and non rewarded locations). What would be a possible explanation for this in your opinion?

"The size of place fields with peaks within the platform segment was comparable with that of fields in other segments. Place fields in the platform segment covered 18.4% (15.3–22.2%) of the visited area, whereas the fields of the remaining cells covered 18.2% (16.4–20.6%)."





Hippocampal encoding of the SR



In a maze similar to the one used in Figure 3, how would an introduction of two rewards at different locations be represented by place cells? If one reward is preferred over the other, how would this be encoded? By the size of the place fields around the reward?

Hippocampal encoding of the SR

A couple of questions regarding supplementary figure 3

Hippocampal encoding of the SR

A couple of questions regarding supplementary figure 3

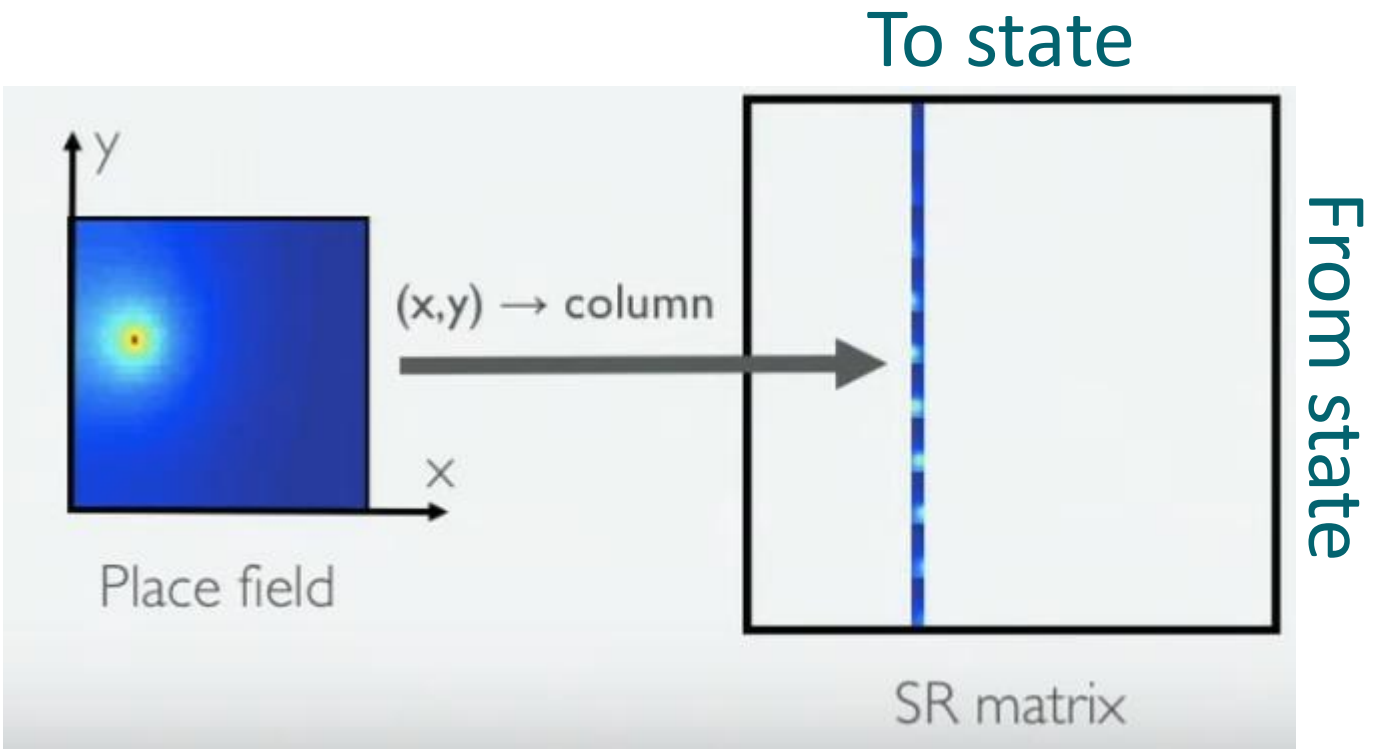
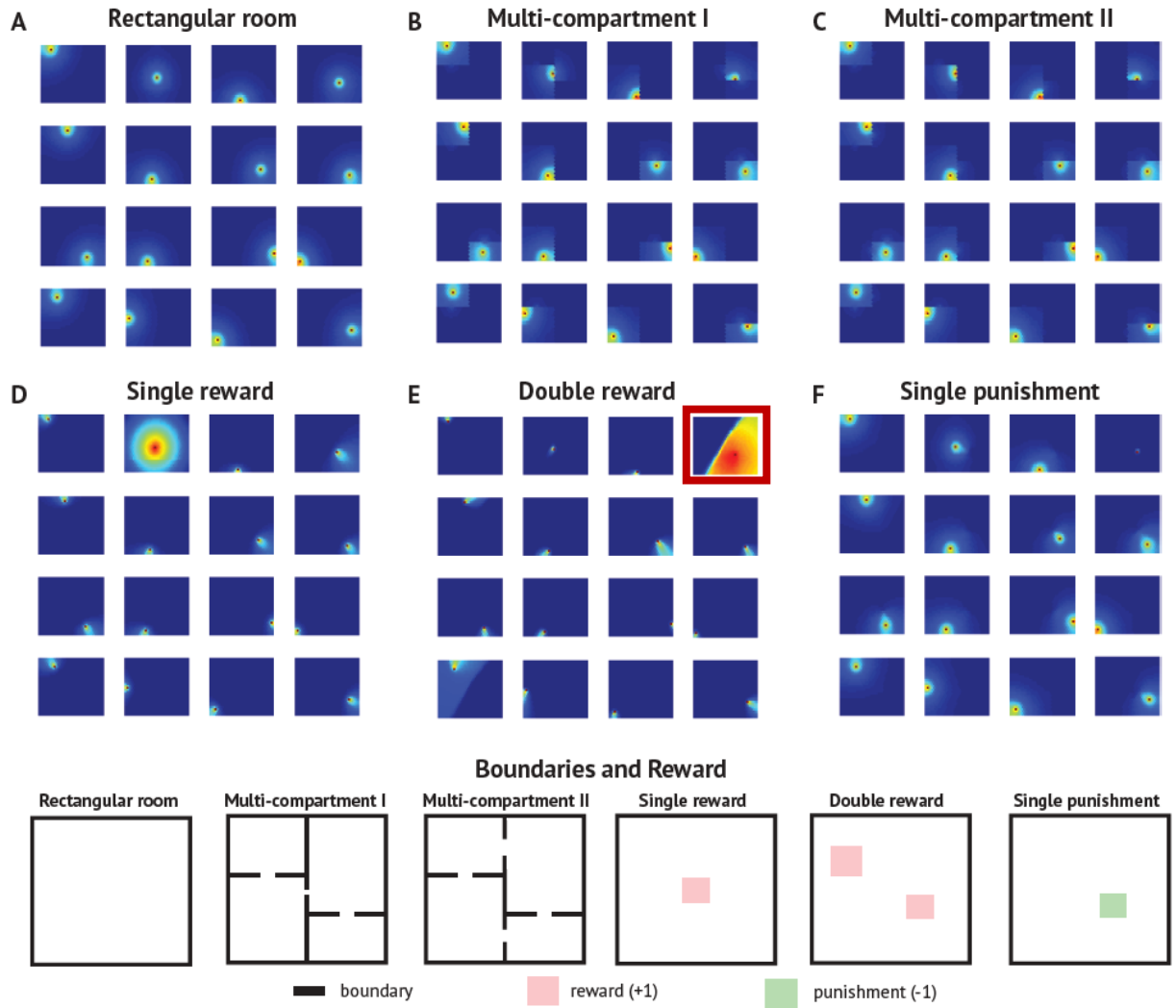


Image from Sam Gershman's talk on youtube

Hippocampal encoding of the SR

A couple of questions regarding supplementary figure 3: For the double reward, half of the room is within the receptive field of one place cell, although only one of the rewards lies within this receptive field – why then from a prediction of successor states would this entire half of the room be preferred?



Hippocampal encoding of the SR

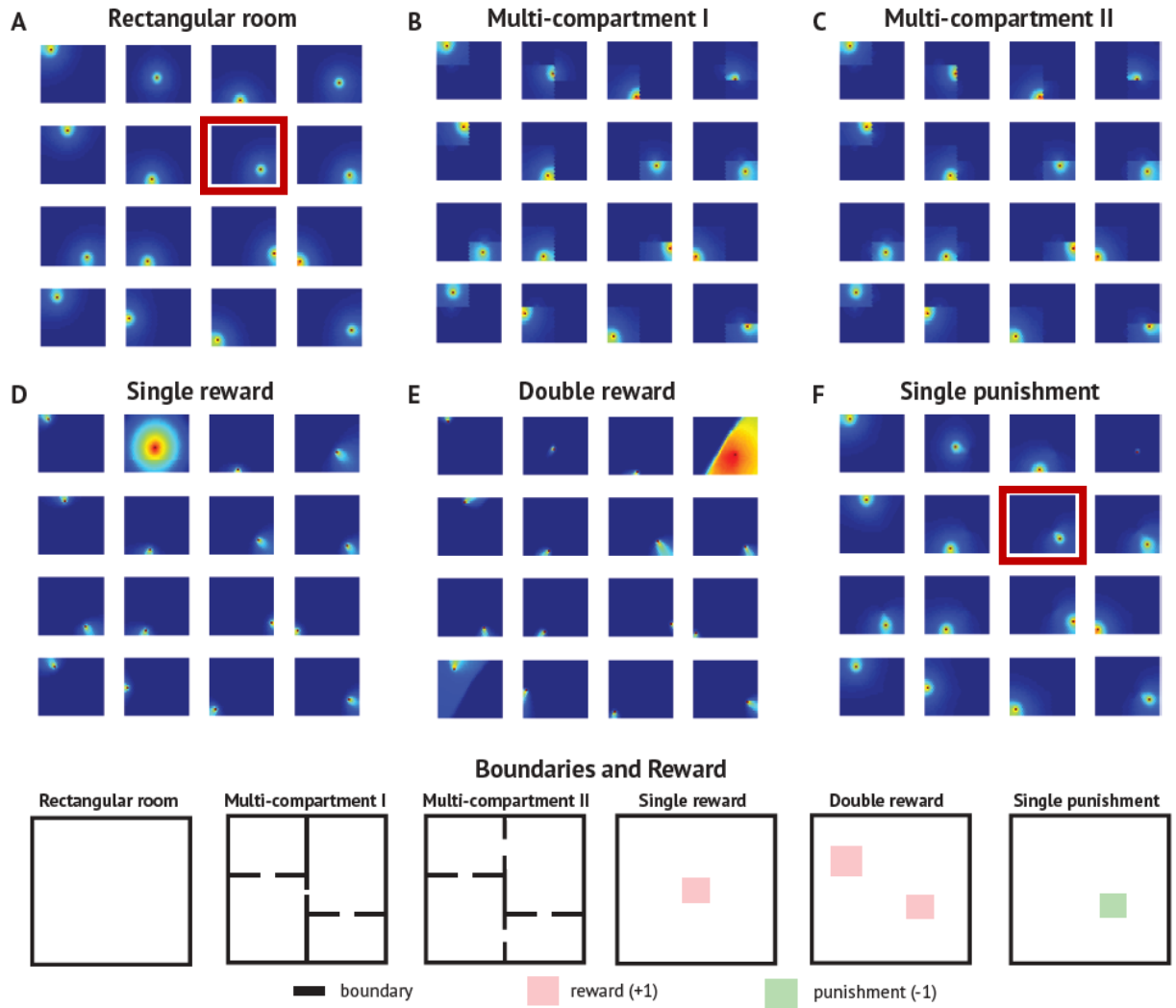
A couple of questions regarding supplementary figure 3: For the double reward, half of the room is within the receptive field of one place cell, although only one of the rewards lies within this receptive field – why then from a prediction of successor states would this entire half of the room be preferred? Secondly, for the punishment context, if only reward is encoded in the receptive fields

$$v_\pi(s) = \sum_{s'} M(s,s')R(s')$$

$$M(s,s') = \mathbb{E}\left[\sum_{t=0} \gamma^t \mathbb{I}\left(s_t = s'\right) | s_0 = s\right]$$

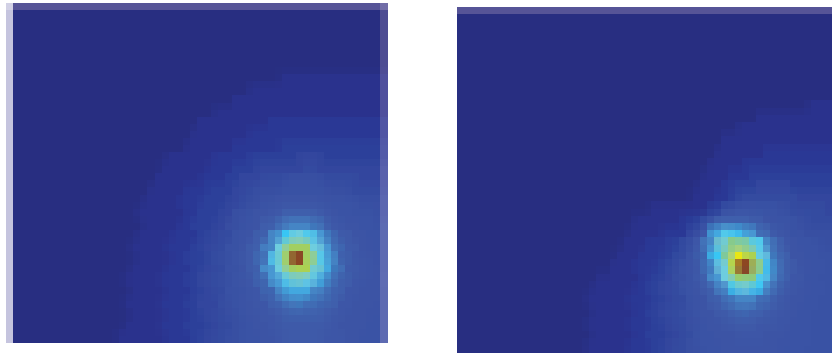
Hippocampal encoding of the SR

A couple of questions regarding supplementary figure 3: For the double reward, half of the room is within the receptive field of one place cell, although only one of the rewards lies within this receptive field – why then from a prediction of successor states would this entire half of the room be preferred? Secondly, for the punishment context, if only reward is encoded in the receptive fields of any of the neurons found, how is predicted punishment represented in the brain? The punished-room looks exactly like the empty room – is this just not represented in the HF?



Hippocampal encoding of the SR

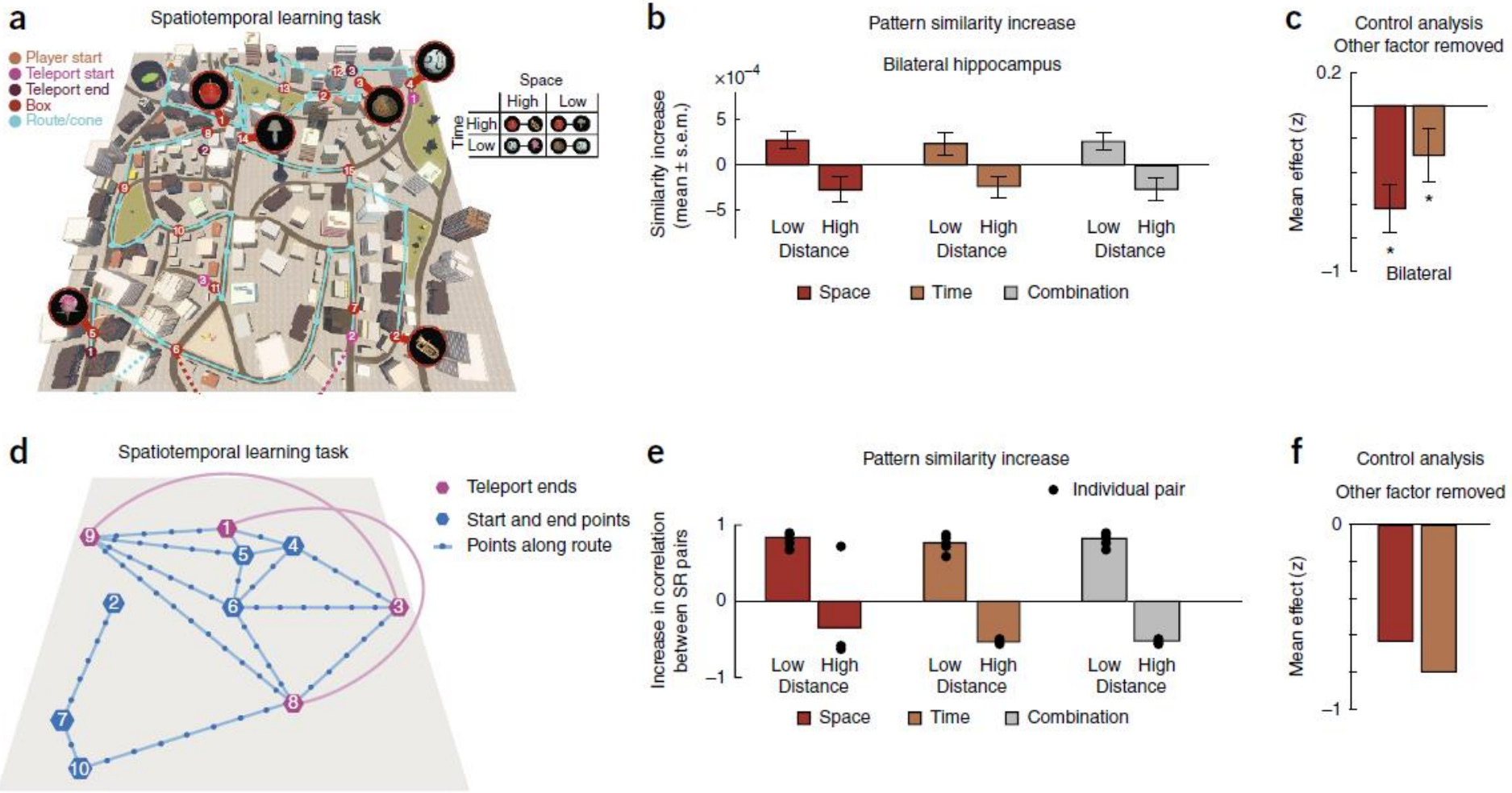
A couple of questions regarding supplementary figure 3: For the double reward, half of the room is within the receptive field of one place cell, although only one of the rewards lies within this receptive field – why then from a prediction of successor states would this entire half of the room be preferred? Secondly, for the punishment context, if only reward is encoded in the receptive fields of any of the neurons found, how is predicted punishment represented in the brain? The punished-room looks exactly like the empty room – is this just not represented in the HF?



Space and time

Humans

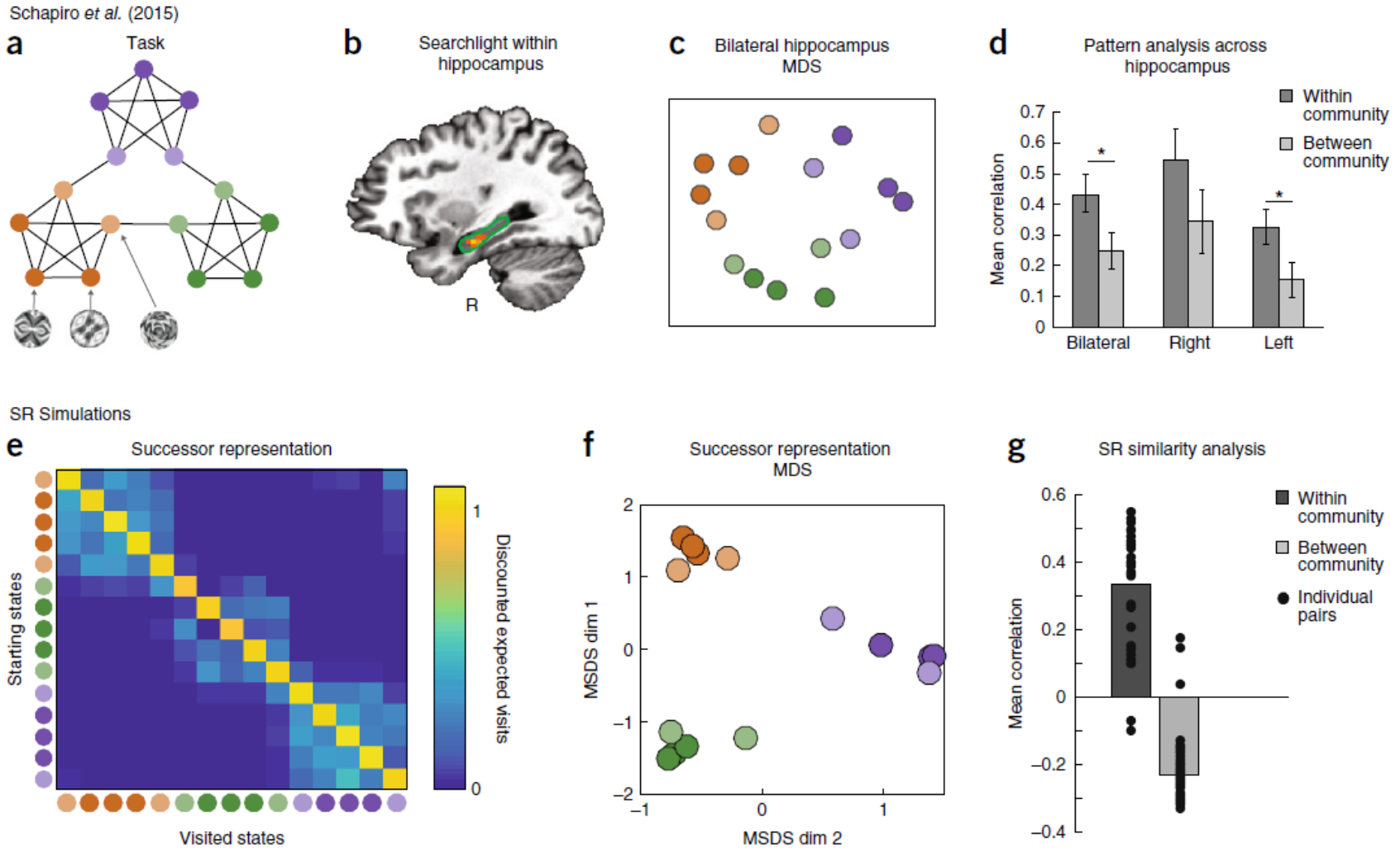
SR



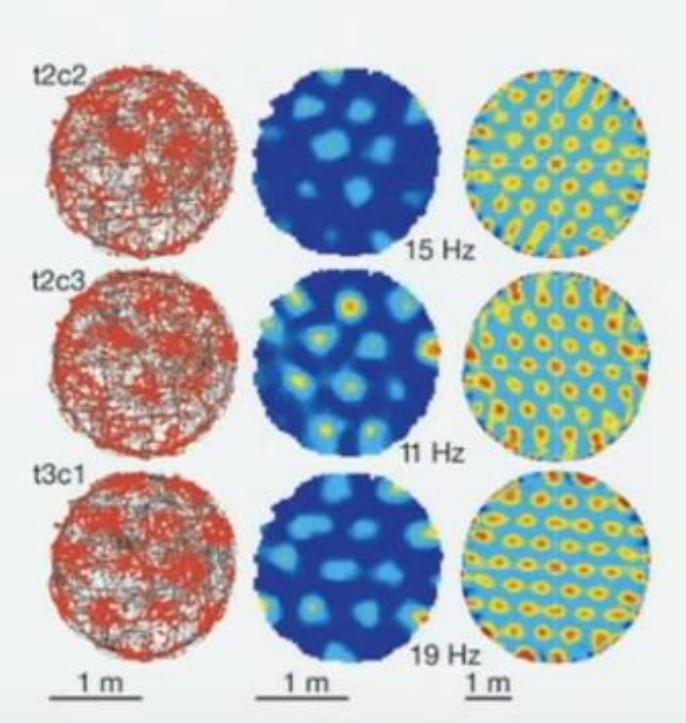
Non-spatial states

Humans

SR



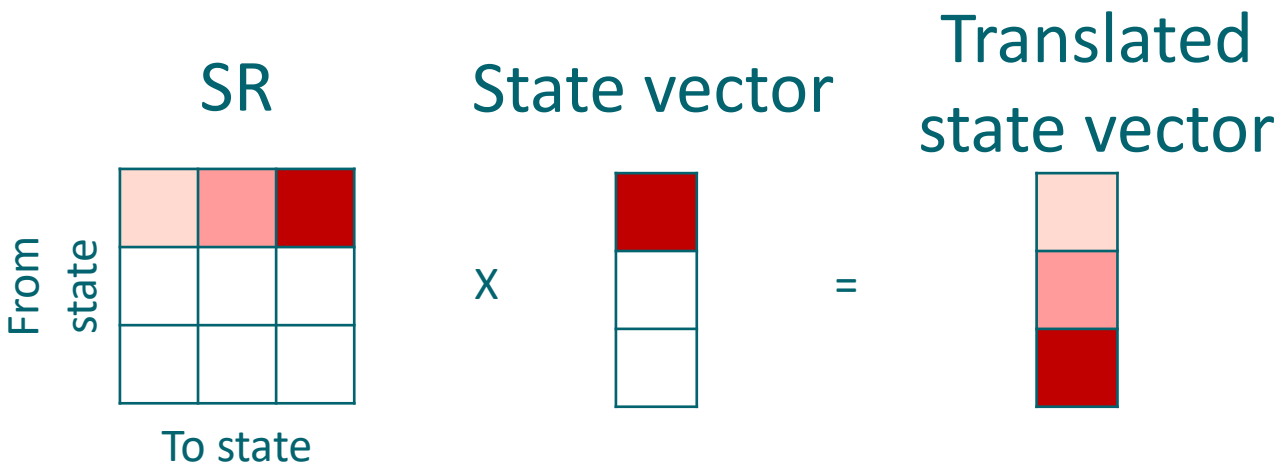
Entorhinal grid cells



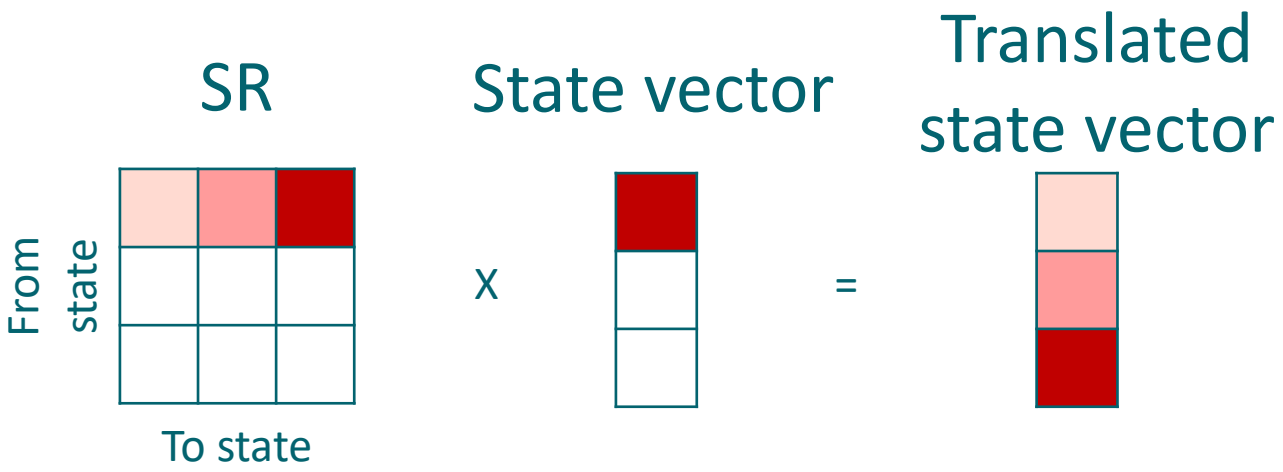
What do they represent? Why are they periodic?

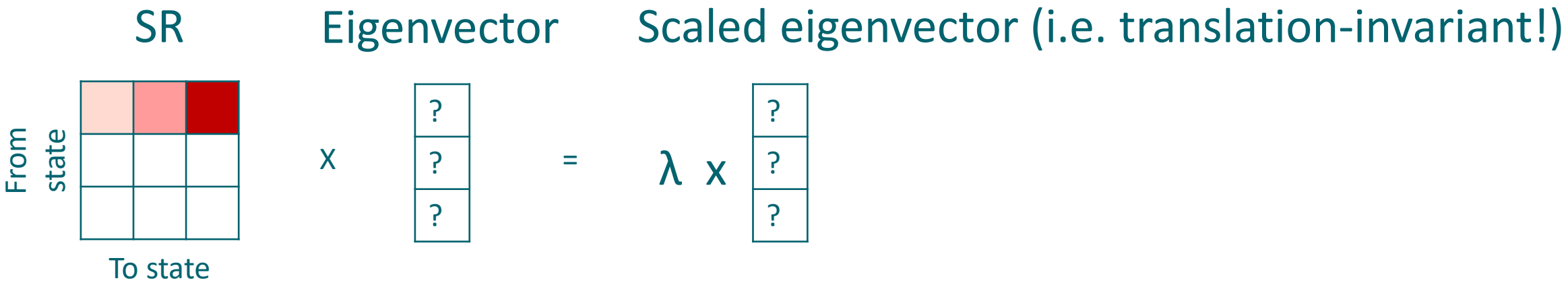
Could you explain/give some intuition about how grid cells being eigenvectors is related to their properties? Are there interesting behavioural properties resulting from this?

Grid cells as a lower-dimensional representation of the SR

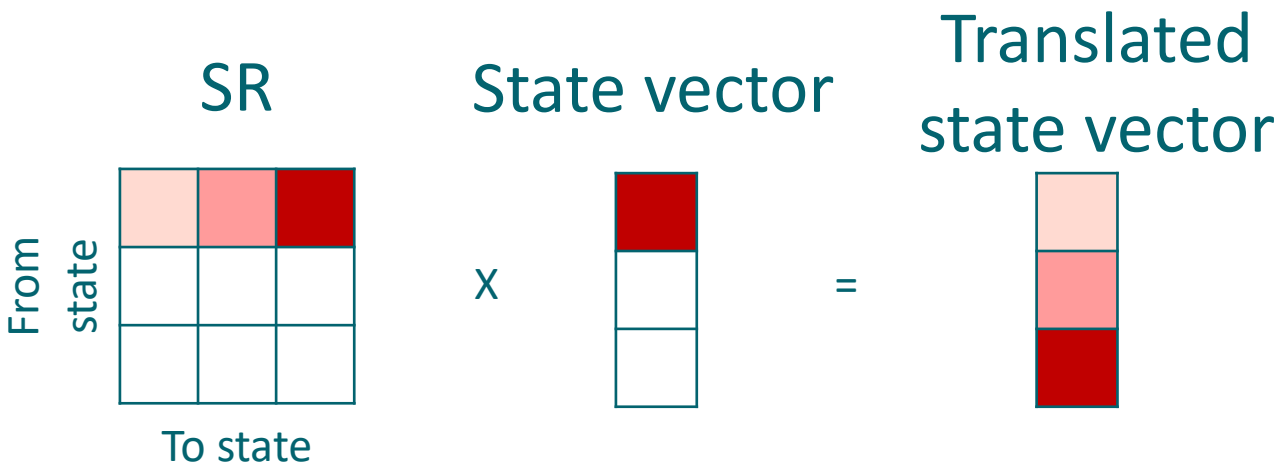


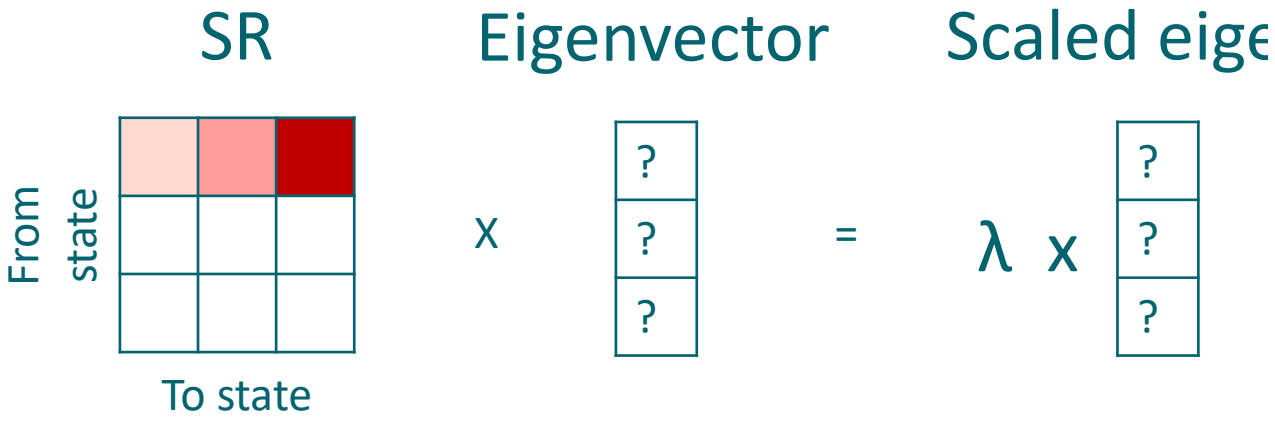
Grid cells as a lower-dimensional representation of the SR

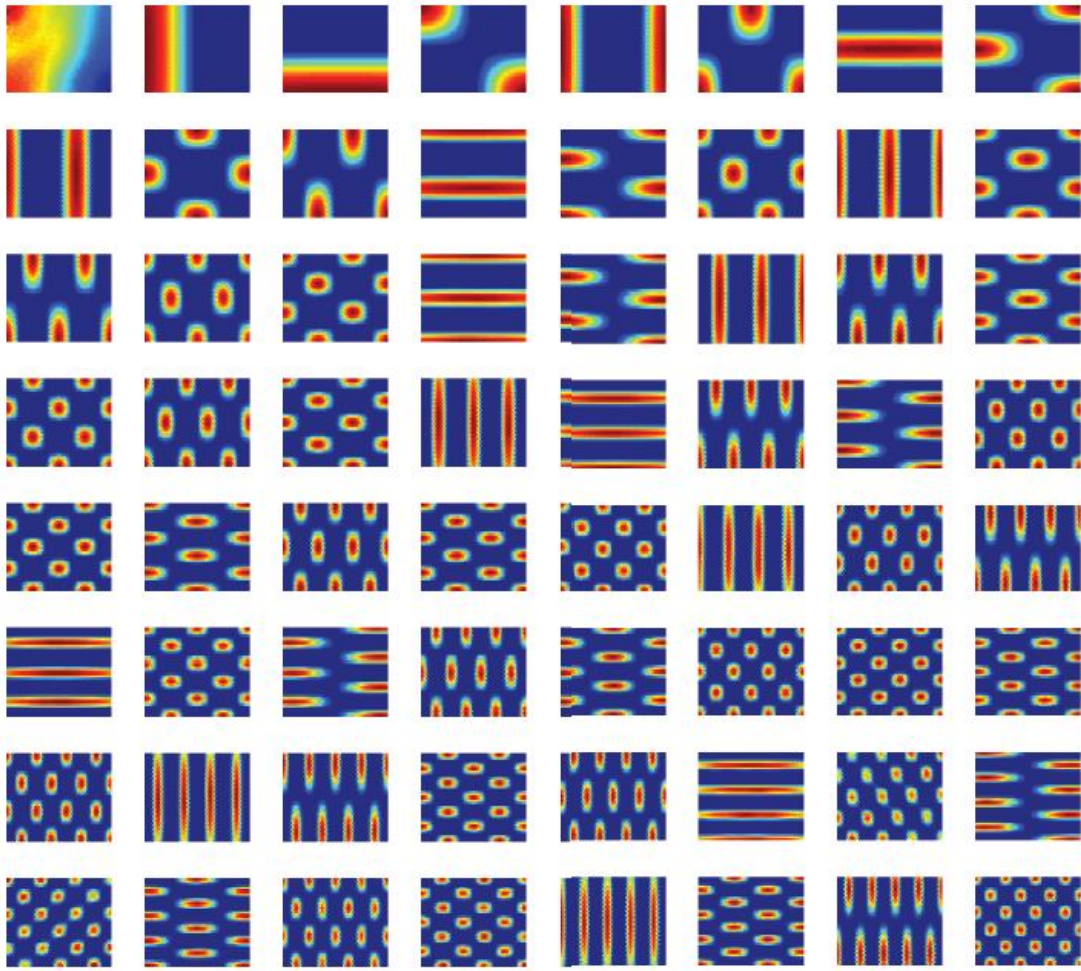




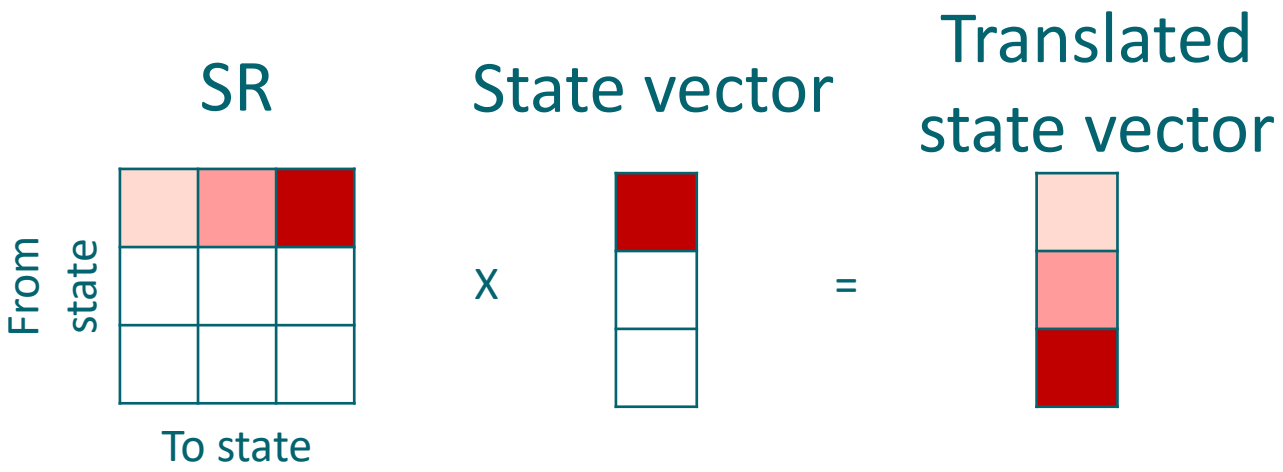
Grid cells as a lower-dimensional representation of the SR

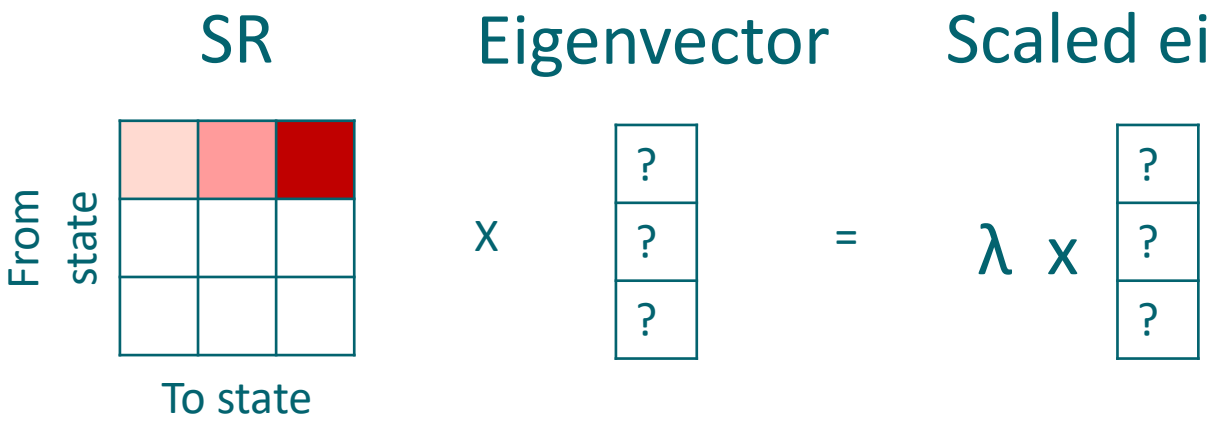






Grid cells as a lower-dimensional representation of the SR





$SR = U\Lambda U^{-1}$

where U is the eigenmatrix (all eigenvectors)

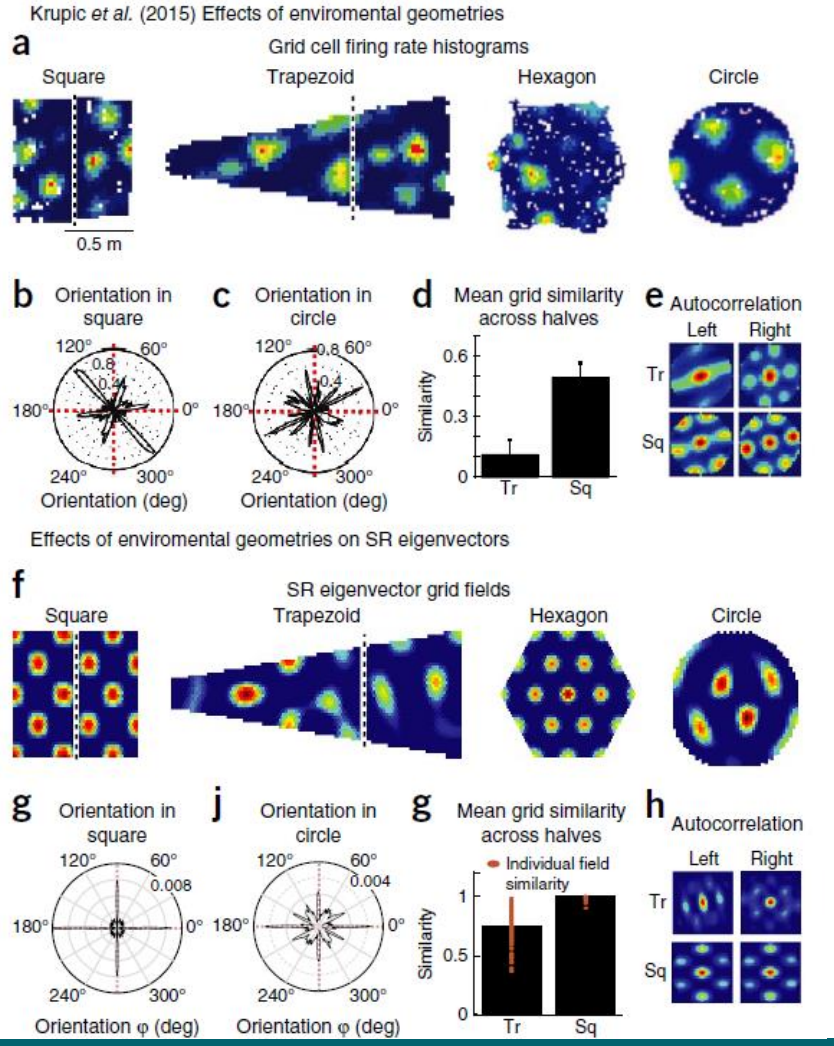
You can

- Reconstruct the SR with the full set of eigenvectors
- Estimate the SR with a just subset of the eigenvectors!

Grid cells as a eigenvectors of the SR

Animals

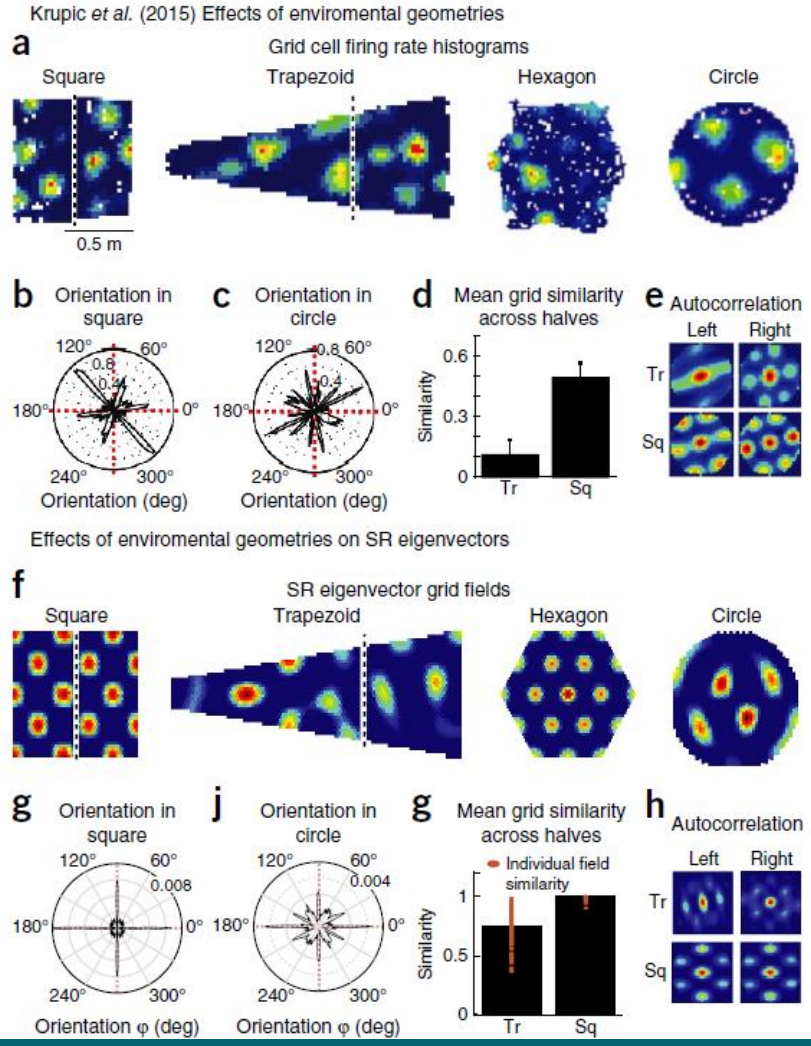
SR

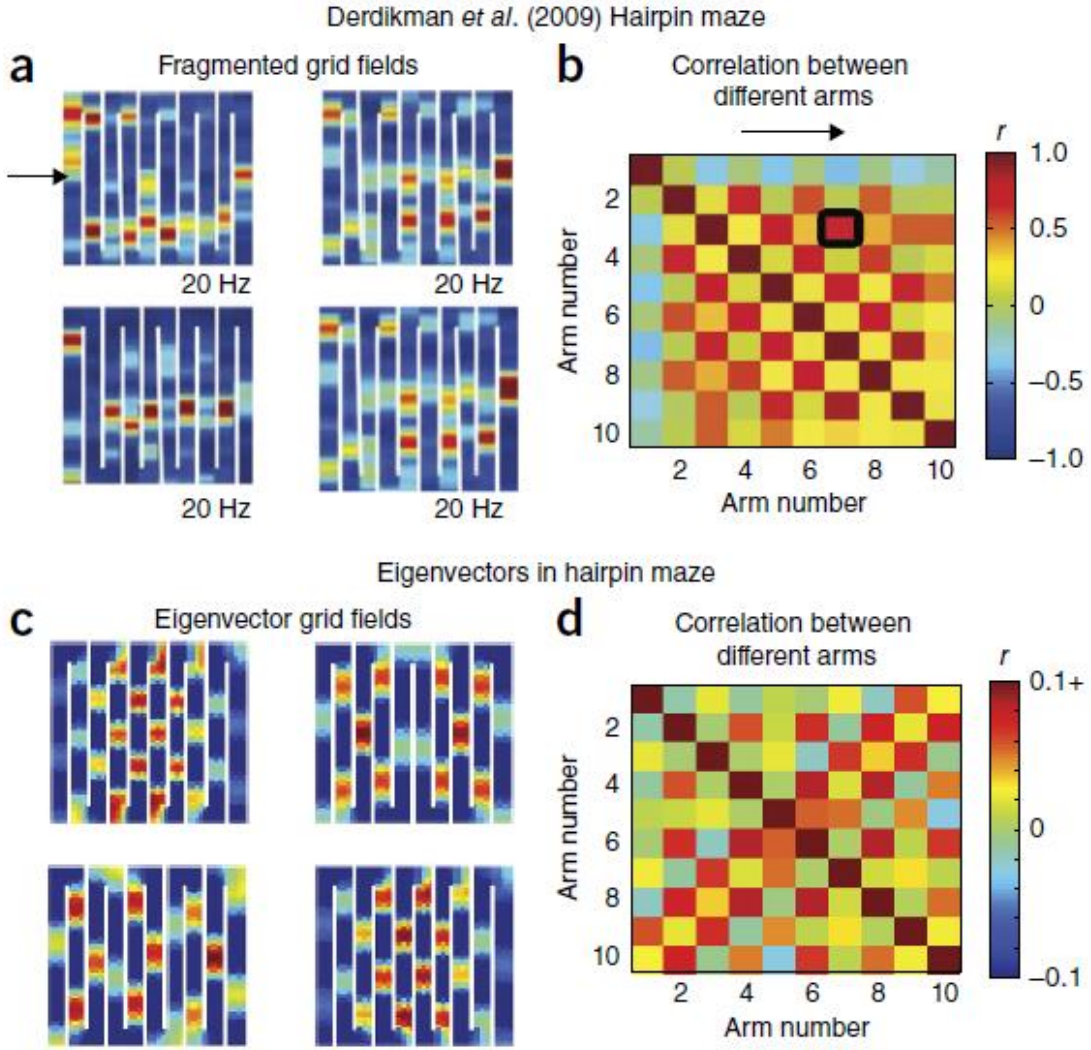


Grid cells as a eigenvectors of the SR

Animals

SR

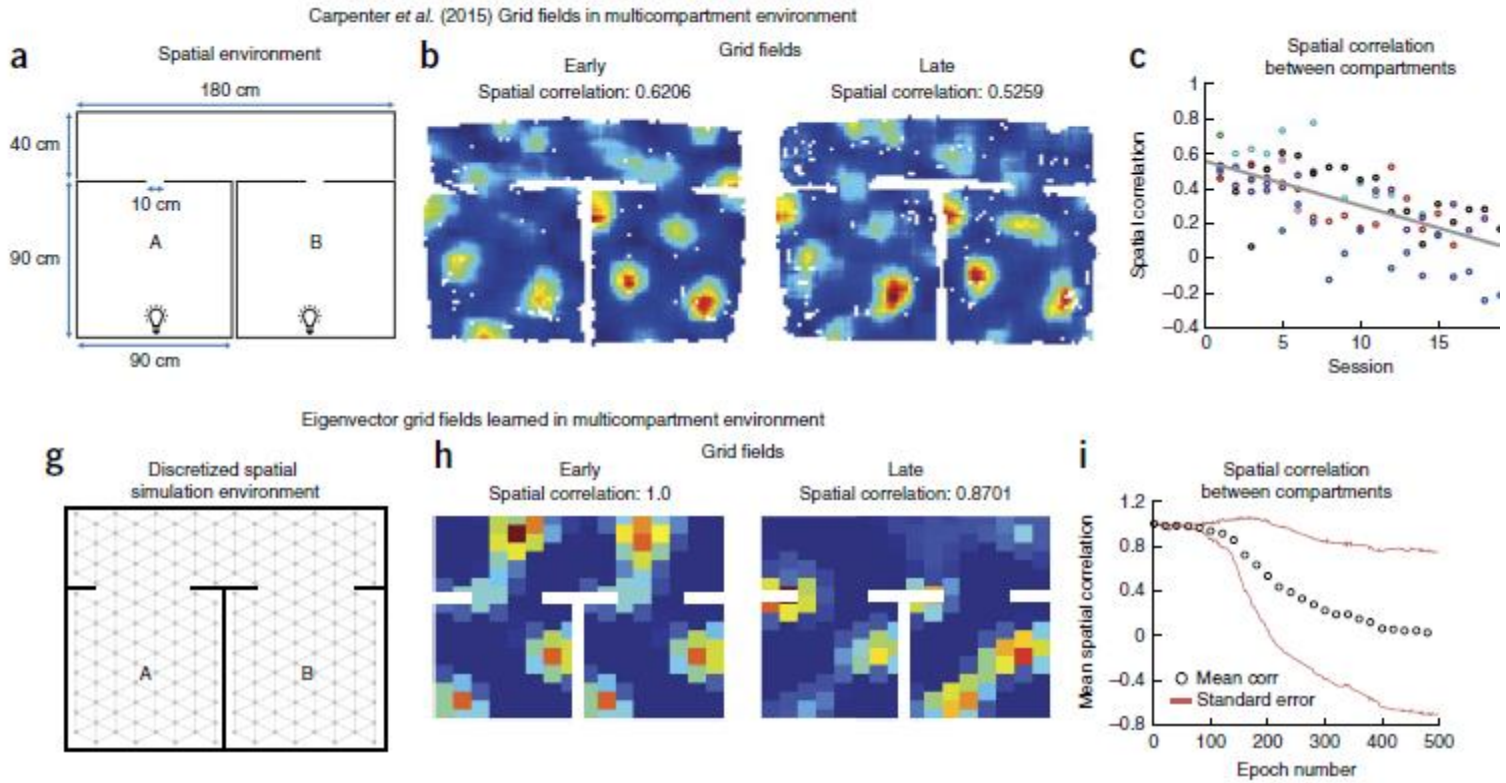




Shift from local grid to global grids

Animals

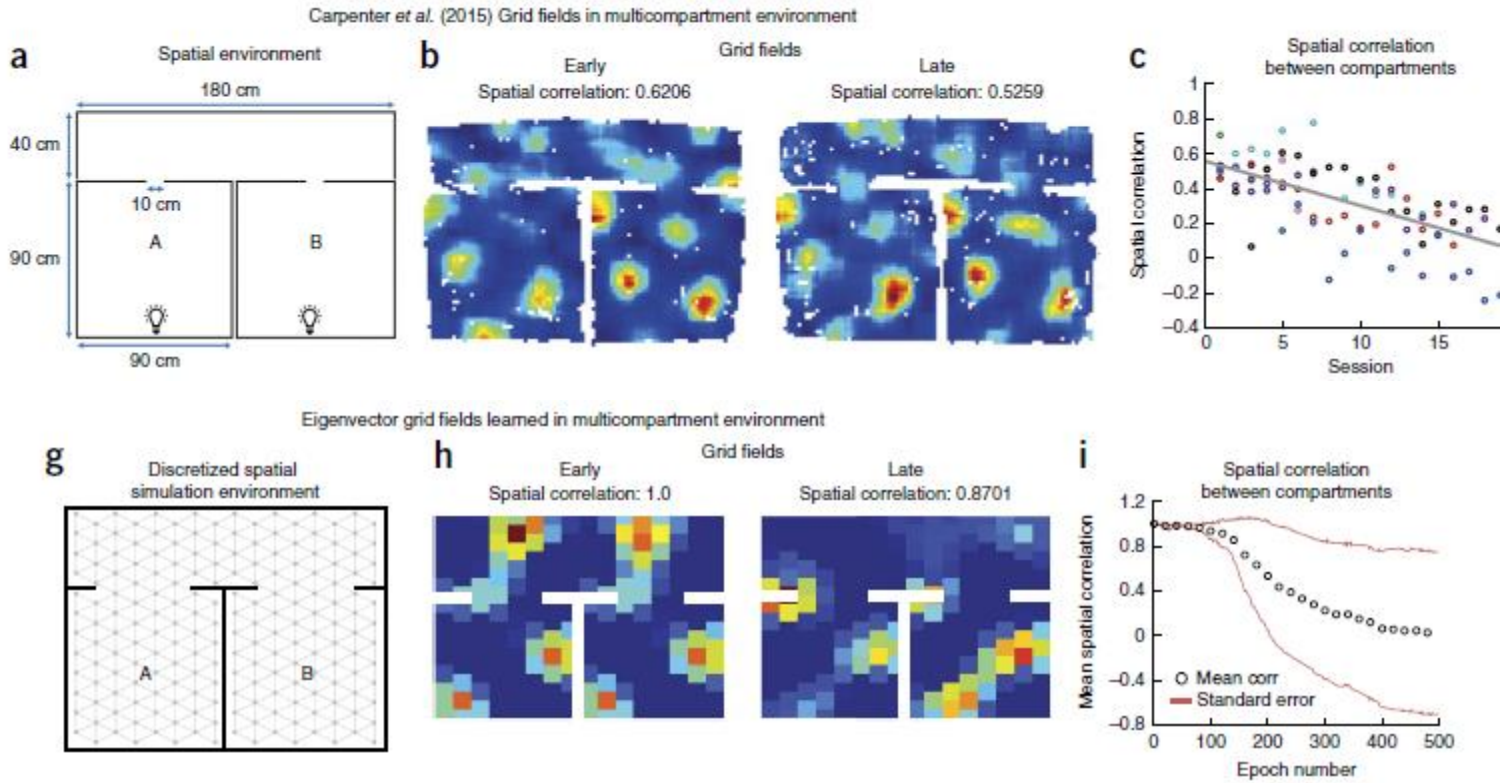
SR



Shift from local grid to global grids

Animals

SR

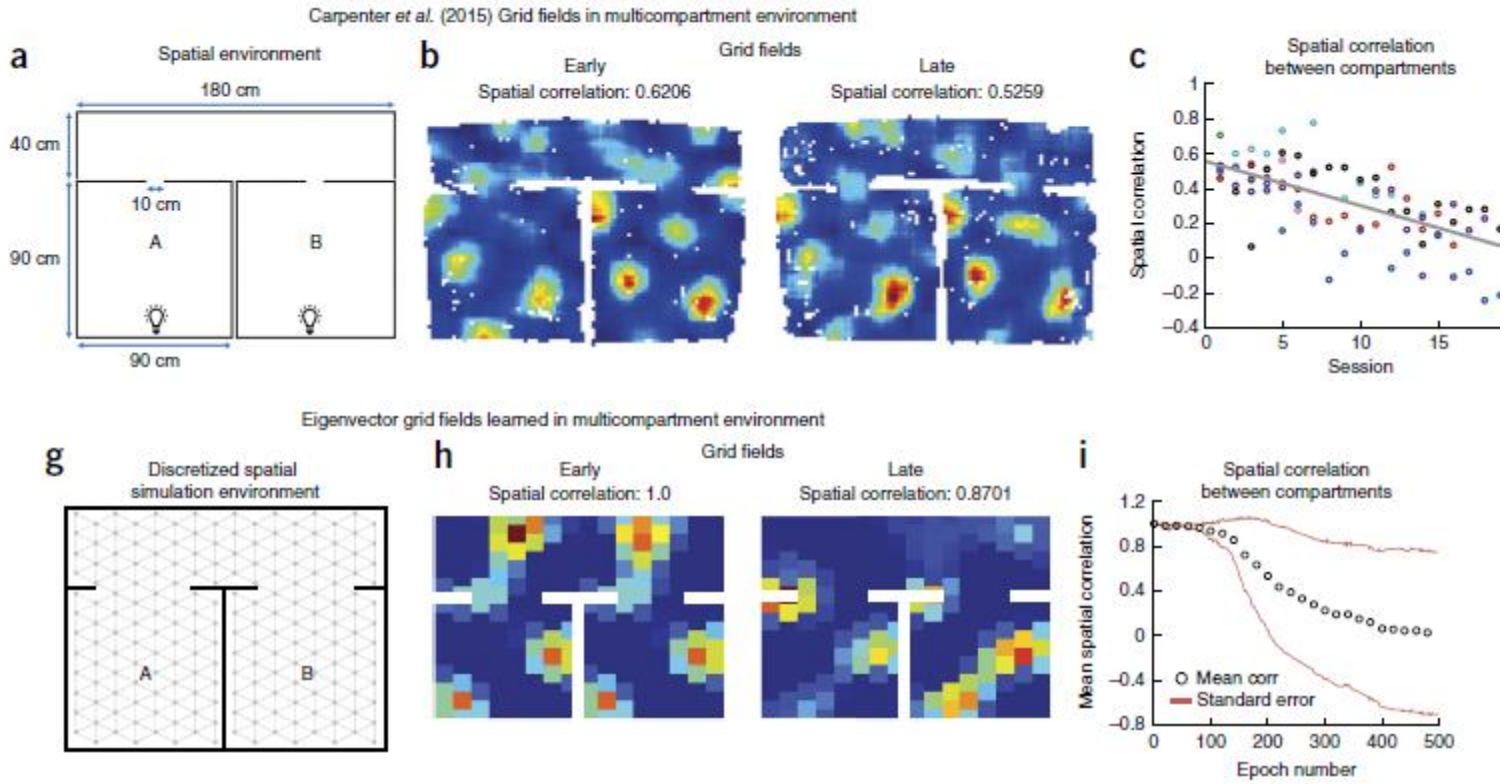


What is an advantage of the global grid as compared to local grids after exploration? How could the underlying changes for this conversion be encoded? Is this just an adaptation or is there more meaning behind?

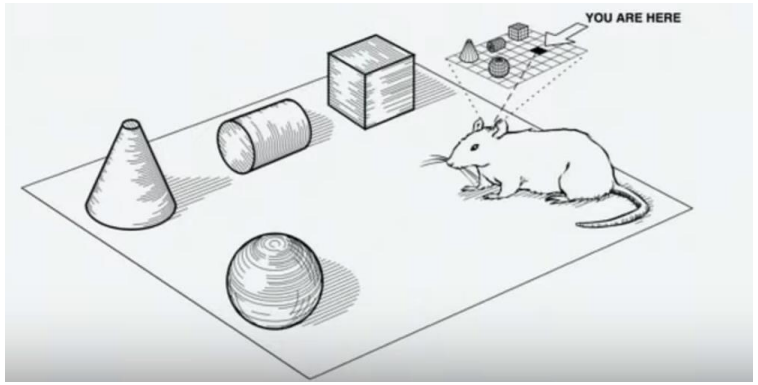
Shift from local grid to global grids

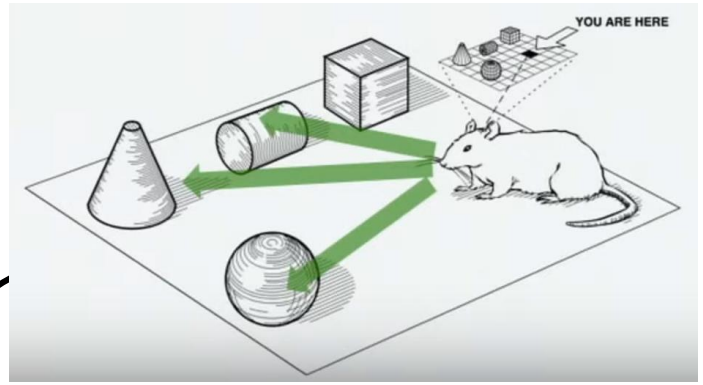
Animals

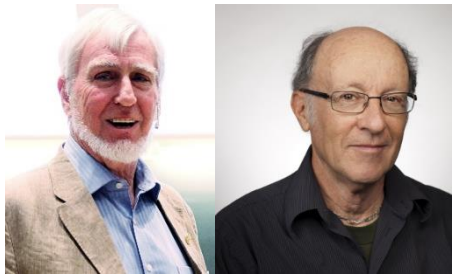
SR



What exactly are the implications for memory formation, if grid cell firing can really be approximated by an SR model? This process has to be connected with a lot more distal projections into other brain areas. How would the implied connection of SR and TCM look on a cellular level?


YOU ARE HERE


YOU ARE HERE



Conclusions

- There are neurons in the hippocampus that mirror predictions of where the animal will be rather than where she is
- This predictive representation is useful because it allows for planning that is easier than model-based but better than model-free (in certain scenarios)

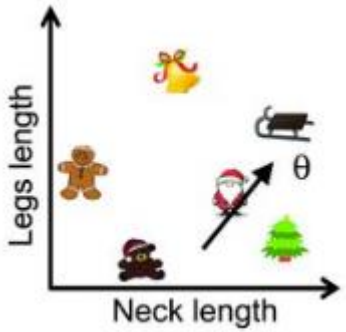
Other questions

- Last Session, time cells as an analogue to place cells were introduced. Are there studies that Show that time cells are also predictive in nature and that they have backward-skewed receptive fields (as shown in this study for place cells)? What would that even mean exactly if time cells were predictive?

- To support spatial information a state value representing a state of mind may be useful. This may guide thought processes and rewards independent and supplementary to spatial information. In form of abstract place cells, what do these cells may encode? Would they represent a model of "how to think successfully"?

- Additional Programming Question to the audience: Are there RL algorithms/frameworks using neuron firing rates. This may be computationally efficient, like spiking neural networks.

I can imagine ‘predictive time cells’ supporting sequential behavior like bird song (e.g. ‘We will repeat this sound 3 more times.’). But in general, it seems paradoxical to predict time (?)



Nir Moneta’s presentation! In particular, Constantinescu et al. (2016)

Other questions

- It is mentioned in the paper that the SR could extend the range of replay forward sweeps in the hippocampus. Could you elaborate on how this should be understood? How do you think the compatibility between the SR model and the theta/sharp wave-ripple activity would look like?
- They affirm that data have been suggesting that "place field stability and organization depends crucially on input from grid cells.". What are the evidences for this correlation? What type of experiments and measures can be made to support this statement?

Thank you for your attention!