Determining Allocentric Locations of Sensed Features

Numenta

Marcus Lewis and Jeff Hawkins

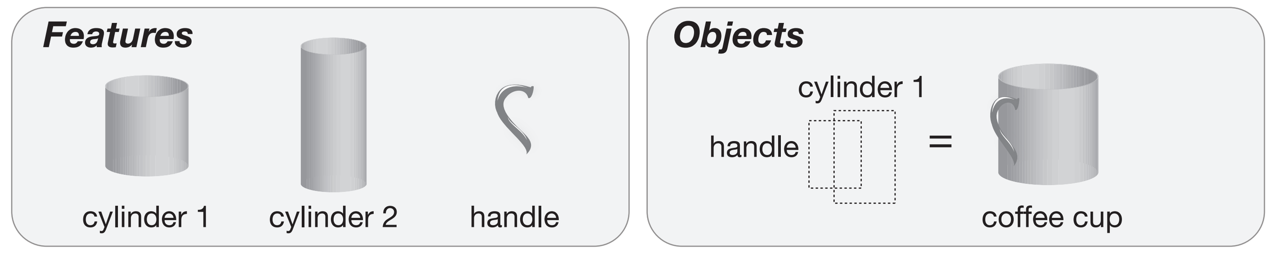
mlewis@numenta.com, jhawkins@numenta.com

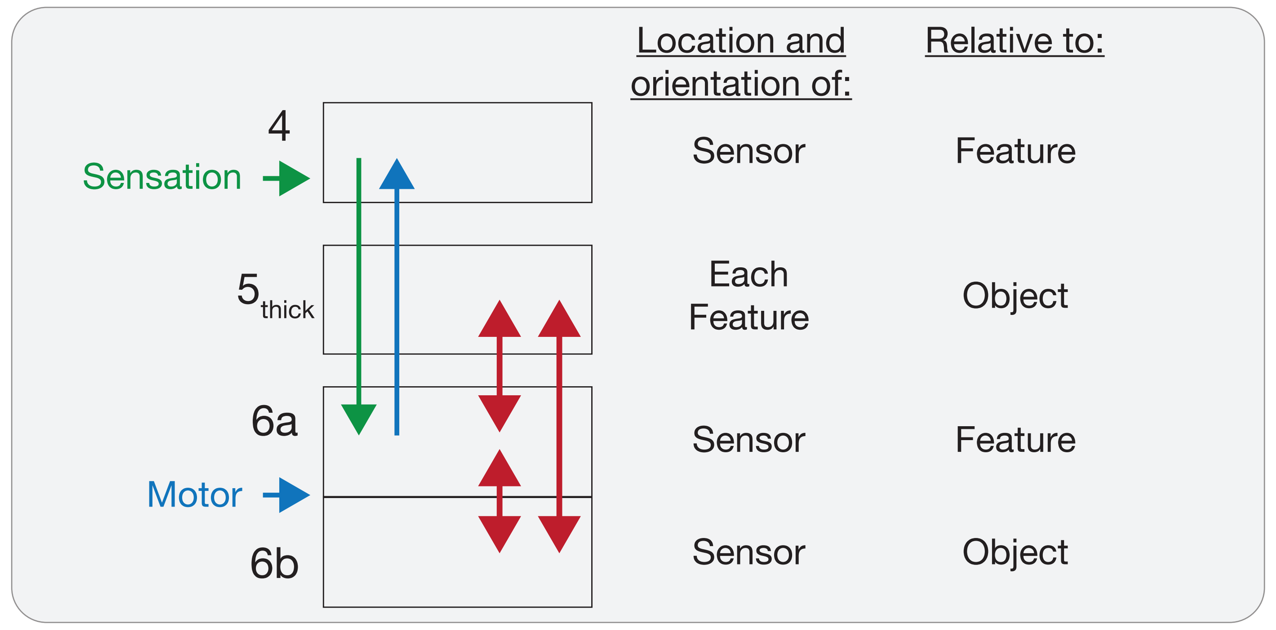
1. Summary

We propose that neocortical columns learn maps of objects, similar to how entorhinal cortex and hippocampus learn maps of environments.

Each column associates sensory inputs with the sense organ's location and orientation relative to the sensed feature.

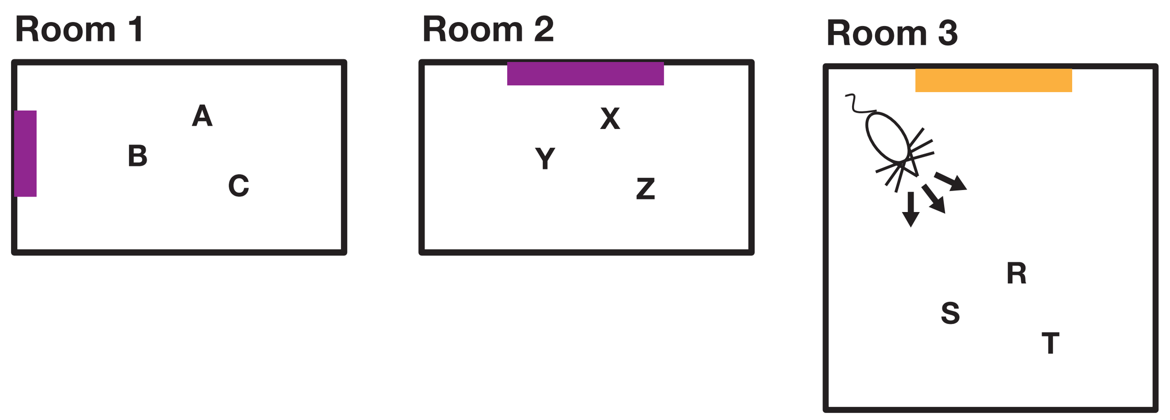
Columns can then arrange these features into novel objects.



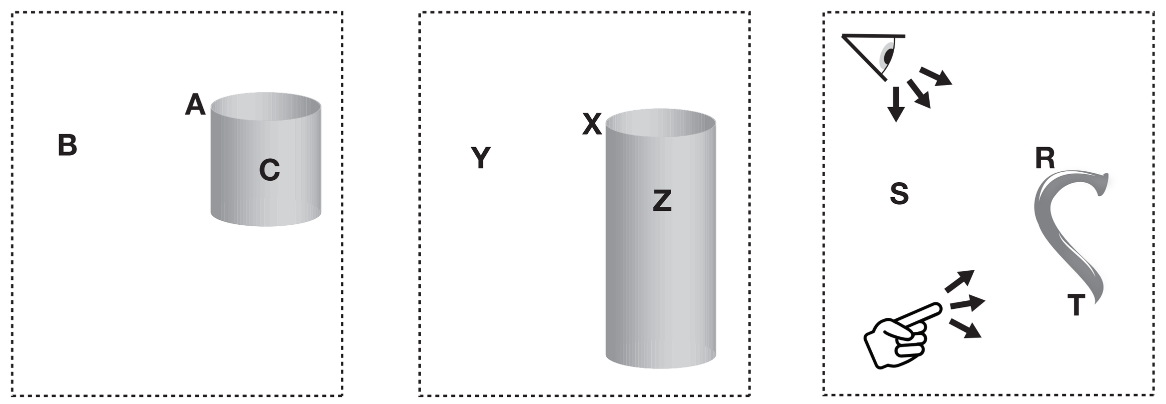


2. Cortical columns model objects as entorhinal cortex models environments.

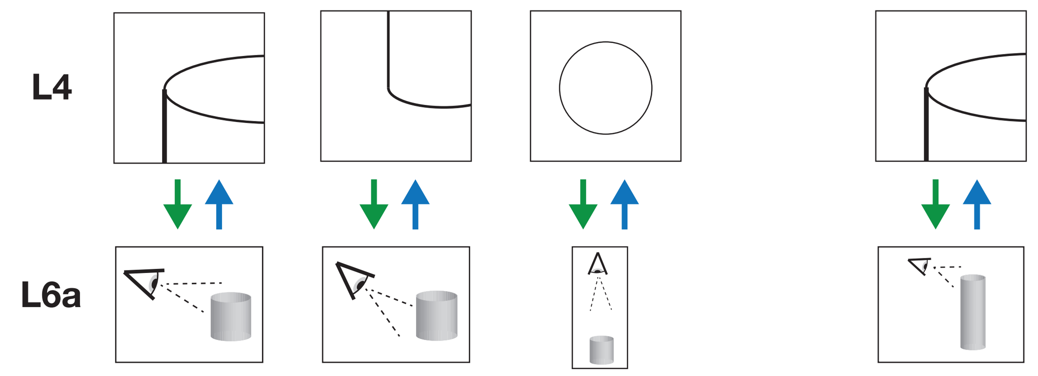
Entorhinal cortex represents locations and orientations relative to specific environments via grid and head direction cells.



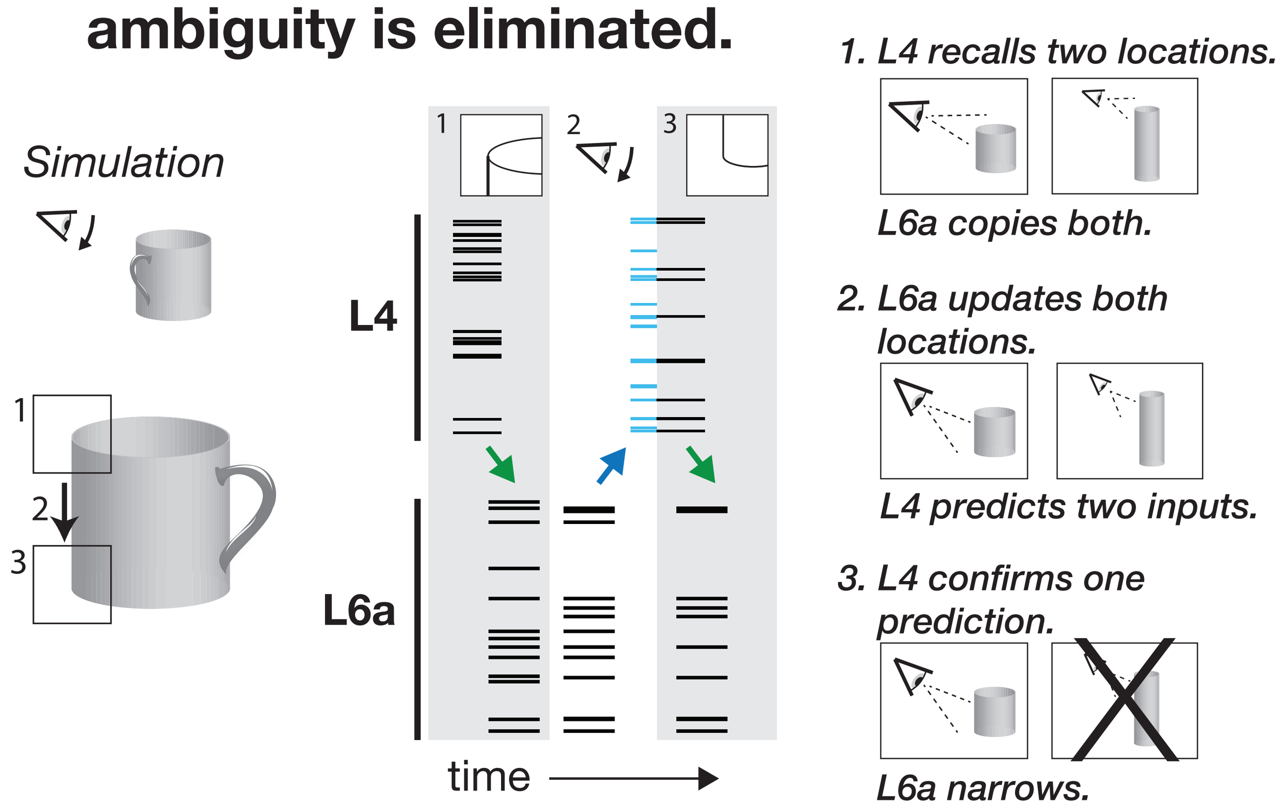
We propose that cortical columns contain analogs of grid and head direction cells. These cells represent the location and orientation of the sensory patch relative to specific objects.



3a. L4 and L6a learn the sensory input at each sensor location.

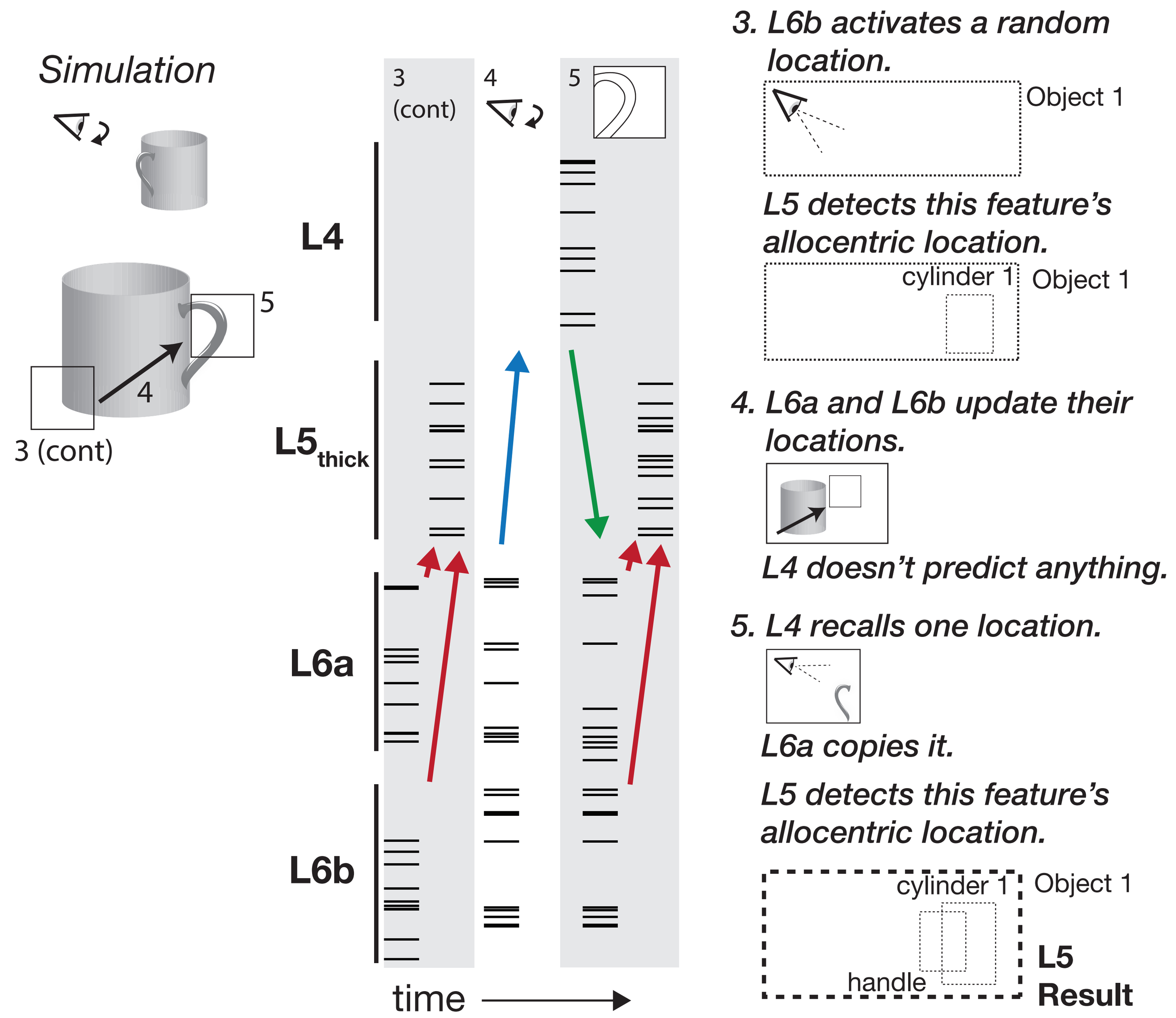


3b. Through movement and sensation, ambiguity is eliminated.



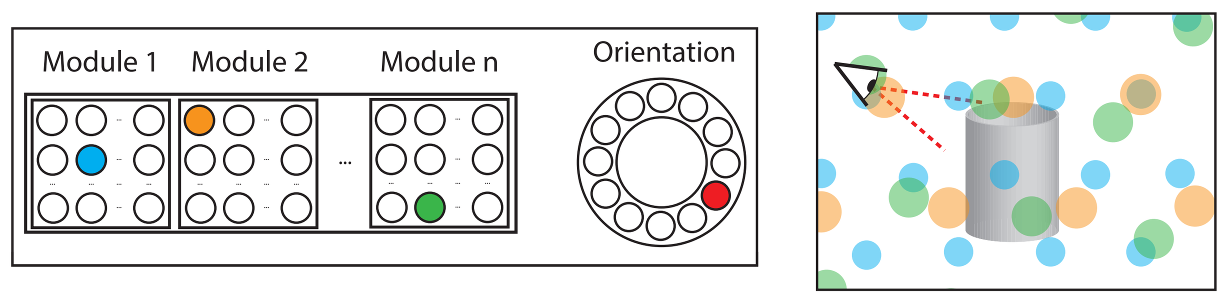
4. $L5_{thick}$ and L6b arrange previously learned features into novel objects.

The cortical column builds a complete model of an object by recognizing each of its component features.



Locations are specific to features/objects.

We propose that $L5_{thick}$, 6a, and 6b contain analogs of grid and head direction cells.



The population represents a conjunctive location + orientation relative to a specific feature or object.

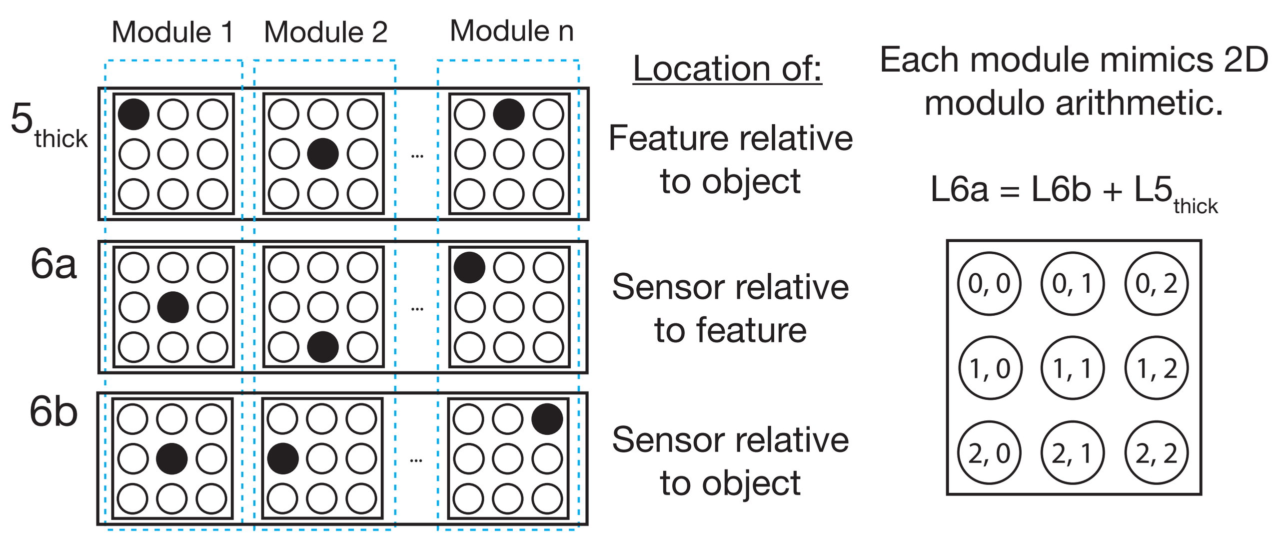
As the sensor moves, the 6a and 6b modules update their location, while the 5_{thick} modules are stable.

The column generates new locations via randomness and path integration. The L4 / L6a algorithm is as follows:

1. Try to recognize the location:
   - Use sensory input to recall ***a set*** of possible locations.
   - Perform path integration ***on the set*** as the sensor moves.
   - Use subsequent sensory input to narrow the set.
2. If recognition fails, begin learning a new feature:
   - Activate a random location,
   - Perform path integration as the sensor moves.
   - Associate the sensory input with these locations.

Grid cells can perform coordinate transforms.

We propose that cortical columns arrange features into objects by performing coordinate transforms with grid cells.



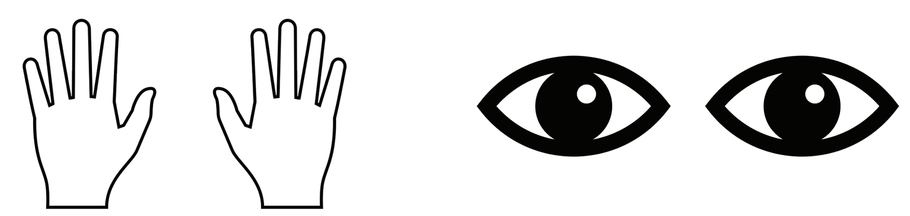
These layers interconnect in "module" compartments. In this simplified model, cells connect to each other in autoassociative triples. If a pair of cells activates, the third cell of the triple will become active.

$L5_{thick}$ is analogous to a transform matrix between 6a and 6b. This "transform" can also be interpreted as "the location of the feature relative to the object".

$L5_{thick}$ represents an object by activating multiple transforms.

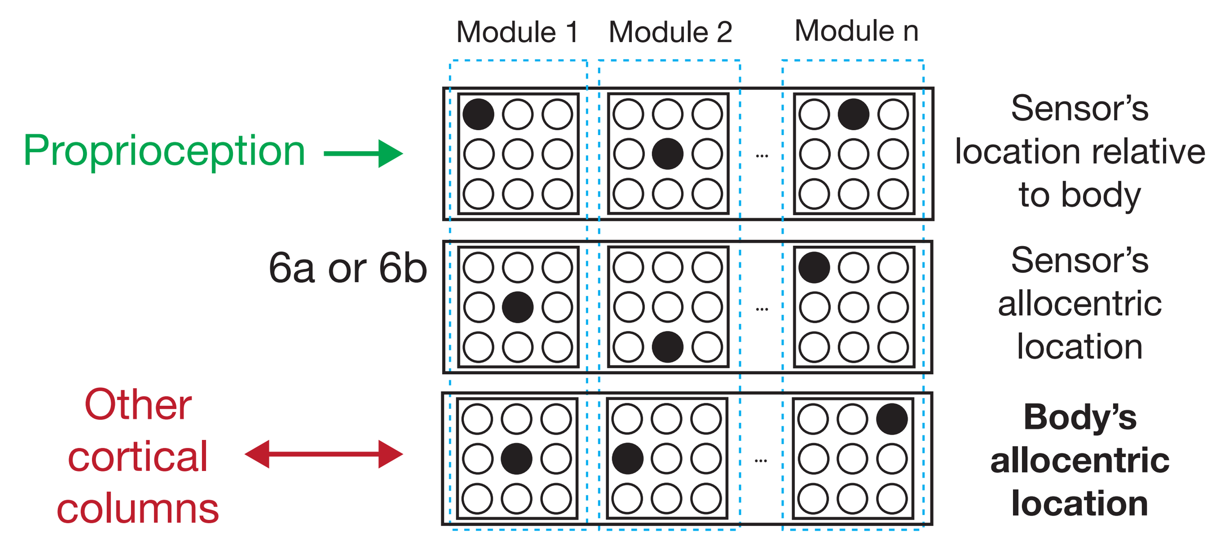
Cortical columns might work together by voting on the body's allocentric location.

Different cortical columns process input from different sensors.



When you grasp an object, you often recognize it with no additional movement, even though no single sensor has received enough information to recognize the object.

Cortical columns might accomplish this by voting on the body's location relative to the feature/object. The column could compute this location with a grid cell transform circuit.



Predictions of the theory

- While an animal attends to an object, cells in the deep layers of sensory cortex have grid cell and head direction cell firing properties. Instead of anchoring to the environment, these cells anchor to the object.
- $L5_{thick}$ builds stable representations of objects that are invariant to the sense organ's location and orientation relative to the object.
- $L5_{thick}$ quickly builds these stable representations for novel objects if the novel objects consist of familiar features.
- L6a cell activity oscillates between being driven by movement cues and by sensory cues.

Current research

- How are objects and locations passed up / down / sideways in the hierarchy? How does this theory change the way we should think about the hierarchy?
- How do neurons do more advanced "transforms" between features and objects? (Not just translation, but rotation.)
- These $L5_{thick}$ representations can be recalled from sensory input via autoassociation in L5 and via the L3->L5 projection. What's the detailed mechanism?

References

- Fiete, I.R., Burak, Y., and Brookings, T. (2008). What grid cells convey about rat location. J. Neurosci. 28, 6858–6871.
- Hawkins, J., Ahmad, S. & Cui, Y. (2017) A theory of how columns in the neocortex enable learning the structure of the world. Front. Neural Circuit., 11, 81.