Representation learning facilitates different levels of generalization

Fabian M. Renz (renz@mpib-berlin.mpg.de)
Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Shany Grossman (grossman@mpib-berlin.mpg.de)
Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany

Peter Dayan (dayan@tue.mpg.de)
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
University of Tübingen, Tübingen, Germany

Christian F. Doeller (doeller@cbs.mpg.de)
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
Kavli Institute for Systems Neuroscience, NTNU, Trondheim, Norway
Wilhelm Wundt Institute of Psychology, Leipzig University, Leipzig, Germany.

Nicolas W. Schuck (schuck@mpib-berlin.mpg.de)
Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany
Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Berlin, Germany

Abstract:

Cognitive maps represent relational structures and are taken to be important for generalization and optimal decision making in spatial as well as non-spatial domains. While many studies have investigated the benefits of cognitive maps, how these maps are learned from experience has remained less clear. We introduce a new graph-structured sequence task to better understand how cognitive maps are learned. Participants observed sequences of episodes followed by a reward, thereby learning about the underlying transition structure and fluctuating reward contingencies. Importantly, the task structure allowed participants to generalize value from some episode sequences to others, and generalizability was either signaled by episode similarity or had to be inferred more indirectly. Behavioral data demonstrated participants` ability to learn about signaled and unsignaled generalizability with different speed, indicating that the formation of cognitive maps partially relies on exploiting observable similarities across episodes. We hypothesize that a possible neural mechanism involved in learning cognitive maps as described here is experience replay.

Keywords: representation learning; cognitive maps; replay

Cognitive maps for generalization and planning

The ability to generalize is essential for flexible decision making. It is often hypothesized that generalization relies on building a cognitive map (Behrens et al., 2018) that reflects knowledge about both spatial and non-spatial relational structures (O'Keefe & Nadel, 1978; Wu et al., 2020). Cognitive maps support value learning and generalization which in turn allows flexible planning and optimal decision-making (Liu et al., 2021). Together with factorized representations – the characterization of task structure along independent dimensions (e.g., position and context) – cognitive maps might build the basis for flexibly recombining knowledge in novel ways (Behrens et al., 2018). Here, we developed a new graph structured sequence task to gain a better understanding of the emergence of factorized structural knowledge over time.

Experimental Design & Behavioral Results

We designed an episode sequence task (Figure 1), which required participants to infer an underlying structure to make optimal decisions. Our task consisted of 4 sequences of episodes which lead to two slowly drifting probabilistic rewards: three sequences lead to the same reward (common, R1 in Fig 1), whereas one sequence leads to a different reward (rare, R2). Furthermore, two of the three common sequences consisted of perceptually similar items of the same categories (A1 & A1*, B1 & B1* etc., see light orange sequences in Fig 1), while the remaining two sequences (one common, one rare) were composed of unrelated items (E1 through L4). This structure enabled participants to generalize values across the two sequences that share an underlying categorical structure (e.g., A1 and A1*), or across sequences that share a common reward but do not share a categorical structure (e.g., A1 and E1). Participants learned from standard trials in which they experienced different sequences and the associated outcomes (e.g., A1>B2>C3>D4>R1, see Fig 1A). To test whether participants utilized knowledge about the relational structure of the task, they were repeatedly presented with decision trials that allowed them to switch from one episode sequence to a different one, thus potentially reaching a different outcome. In *value generalization* trials (Fig 1B-C), participants would start from a sequence leading to the common reward, and had to decide between two sequences that either led to the common or the rare outcome after one step. In *close generalization* trials (Fig 1B), participants were offered a sequence with shared categories (going from A1>B2*), versus the rare reward option (A1>J2). In *far generalization* trials (Fig 1C), the unrelated sequence of the common reward (A1>F2) had to be compared with the rare reward sequence (A1>J2). In *value* trials, participants started in the unique sequence of the common reward and could switch to either a different sequence of the common reward or to the rare reward sequence. Finally, in *position trials* (Fig 1D) different sequence positions of the same common outcome sequence were offered, such that only ordinal sequence position was relevant for the decision, independently of reward and category membership. We analyzed the proportions of correct choices for the different choice trials, which reflect participants knowledge about different levels of generalizability, i.e., near generalization to episodes from the same category, far generalization to episodes of the different category and position-based generalization. Behavioral data demonstrate the participants learned to leverage all levels of generalizability, thereby quickly adapting to the fluctuating reward contingencies. Knowledge of near generalization was apparent earlier than knowledge of far generalization.

Conclusions

These results shed light on the factors that drive learning of environmental structures, which can be leveraged to support decision making. A main implication is that learning of cognitive maps is enhanced by observable similarities but also includes inferred relationships. We speculate that a possible neural substrate for establishing the structural knowledge as described here is experience replay (Wittkuhn et al., 2021, 2022). According to this idea, replay goes beyond the application of preexisting structural knowledge, and takes an active role in establishing task-state representations by combining experienced sequences in novel ways that shape subsequent neural representations and generalization behavior.

Acknowledgments

FMR is funded by the Max Planck School of Cognition. SG is supported by Zuckerman-CHE program. CFD's research is supported by the Max Planck Society, the European Research Council (ERC-CoG GEOCOG 724836), the Kavli Foundation, the Jebsen Foundation, Helse Midt Norge and The Research Council of Norway (223262/F50; 197467/F50). PD's research is supported by the and the Alexander von Humboldt Foundation and the Max Planck Society. NWS is supported by the Max Planck Society and the European Research Council.

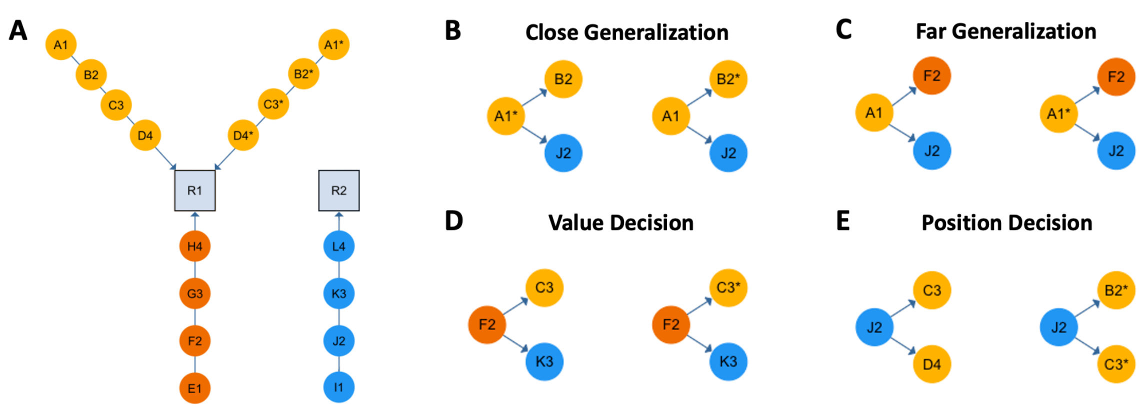

Figure 1. Panel A shows an overview of the graph structured sequence task. A total of 4 sequences result in a common and a rare reward. The common reward is associated with a total of 3 sequences, 2 of which share their category structure (light orange). Different decisions probe for different types of knowledge. Value generalization probes for category knowledge expecting reward generalization to benefit from shared categories (B) compared to unrelated categories (C). Value decisions probe value learning without category information (D). Position decisions probe for representations of the sequence position independent of value and category (E).

References

Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, *100*(2), 490–509.

Liu, Y., Mattar, M. G., Behrens, T. E., Daw, N. D., & Dolan, R. J. (2021). Experience replay is associated with efficient nonlocal learning. *Science*, *372*(6544), eabf1357.

O'keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford university press.

Wittkuhn, L., Chien, S., Hall-McMaster, S., & Schuck, N. W. (2021). Replay in minds and machines. *Neuroscience and Biobehavioral Reviews*, *129*, 367–388. https://doi.org/10.1016/j.neubiorev.2021.08.002

Wittkuhn, L., Krippner, L. M., & Schuck, N. W. (2022). Statistical learning of successor representations is related to on-task replay. *bioRxiv*.

Wu, C. M., Schulz, E., Garvert, M. M., Meder, B., & πSchuck, N. W. (2020). Similarities and differences in spatial and non-spatial cognitive maps. *PLoS Computational Biology*, *16*(9), e1008149.