

Disambiguating planning from heuristics in rodent spatial navigation

Michael Pereira (michael.pereira@research.fchampalimaud.org)

Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal

Christian K. Machens (christian.machens@research.fchampalimaud.org)

Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal

Rui M. Costa (rui.costa@research.neuro.fchampalimaud.org)

Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal

Department of Neuroscience, Zuckerman Mind Brain Behavior Institute, Columbia University, NY 10027, USA

Thomas Akam (thomas.akam@research.neuro.fchampalimaud.org)

Department of Experimental Psychology, University of Oxford, Oxford, UK

Abstract

A longstanding question in neuroscience is how animals and humans select actions in complex decision trees. Planning, the evaluation of action sequences by anticipating their outcomes, is thought to coexist in the brain with simpler decision-making strategies, such as habit learning and heuristics. Though planning is often required for optimal choice, for many problems simpler strategies yield similar decisions, making them difficult to disambiguate. The scarcity of behavioral tasks that can dissociate planning from other decision mechanisms while generating rich decision data has hindered our understanding of the neural basis of planning. We developed a novel navigation task in which mice navigate to cued goal locations in a complex maze. A targeted search through the large space of possible maze layouts in that environment maximizes the number of decisions that are informative about the use of planning. Over the course of training mice learn shorter paths to goals, and the individual decisions composing these paths are better accounted for by planning than vector navigation. With hundreds of informative decisions per behavioral session, this paradigm opens the door to the study of the neural basis of route planning.

Keywords: planning; navigation; learning

Behavioral paradigm

In the task, mice navigate a tortuous elevated maze to collect rewards at visually cued locations. The apparatus itself consists of a gridlike arrangement of elevated platforms with reward ports connected by removable walkways (Figure 1). By choosing which of the 60 elevated walkways are removed, the experimenter can choose from a large number of possible maze layouts. On each trial, one of the 36 possible reward sites is cued with a stimulus light. The mouse navigates to the cued goal location and upon arrival receives a reward. After a short interval another goal location is randomly selected and cued to start the next trial. The trajectories followed by the animals are stored for analysis.

As reward can be delivered at any one of 36 possible locations and the sequence of reward locations is randomised, the utility of habitual strategies is minimized. Due to the tortuosity of the maze, vector navigation often leads to choices which are not on the shortest path to the goal (see example trial in Figure 2).

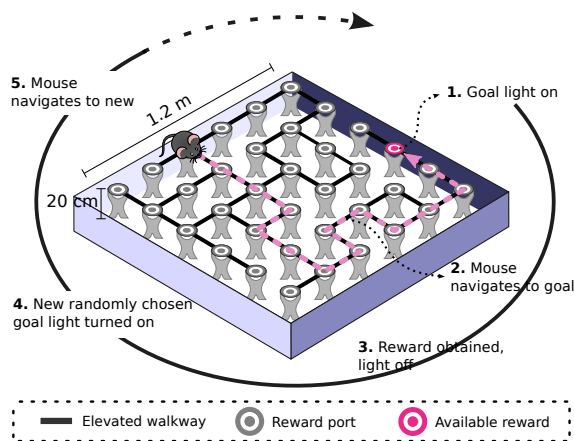


Figure 1: Experimental apparatus and behavioral task flow.

Maze layout optimization

To maximize the discriminability between planning and vector navigation, we performed a guided search through the large space of possible maze layouts. We randomly generated thousands of maze layouts and for each layout we calculated the fraction of decisions for which planning and vector navigation recommended different actions (Figure 3). The higher this value the higher the benefit of planning over vector navigation.

In addition to optimizing for the discriminability of planning we jointly optimised for layouts that have a flat distribution of betweenness centrality across nodes. Betweenness centrality is the fraction of all shortest paths passing through a given location, so nodes with high centrality (e.g. bottlenecks) are on average better choices than nodes with low centrality (e.g. dead ends). Such differences in average value could poten-

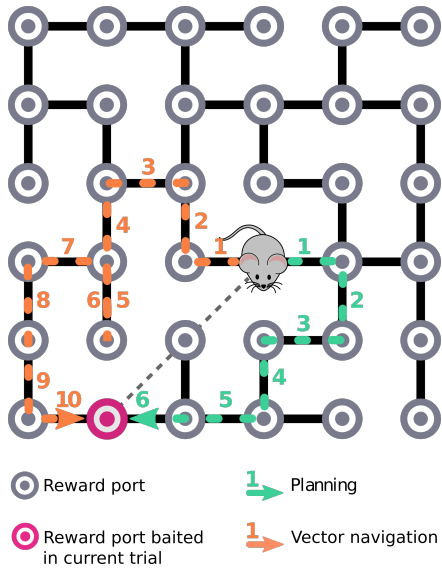


Figure 2: Example trial with trajectories generated by planning and vector navigation. The tortuous maze structure rewards the use of planning with routes shorter than those obtainable by vector navigation.

tially be exploited by habitual learning mechanisms, so we sought to make the centrality distribution as flat as possible while preserving the discriminability of planning from vector navigation. From the optimality frontier (see blue line in Figure 3) different configurations can be selected that represent varying degrees of compromise between the two criteria in this simultaneous optimization.

Analysis of decisions at choice points

We analyzed the routes taken by mice by calculating for each subject an index of how often they chose options favoured by planning and by vector navigation (Figure 4). The {planning, vector navigation} index was defined as the fraction of decisions on which the subject took the option recommended by {planning, vector navigation}, normalized to yield a value of 0 for a random walk and 1 for a deterministic {planning, vector navigation} agent.

After 9 days of training on the same maze layout, the choices observed for all 7 mice tested revealed stronger influence of planning than vector navigation. The obtained results are in the range expected for a planning strategy with some stochasticity.

Acknowledgments

We would like to thank Beatriz Lourenço for her contribution in preparing and performing the next round of behavior experiments, currently under way, aimed at collecting data in different maze layouts.

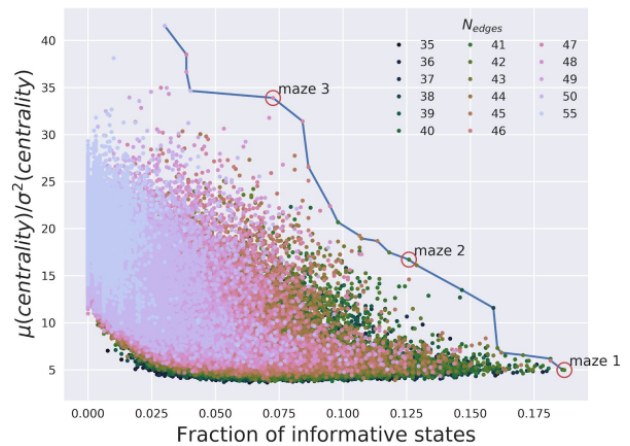


Figure 3: Search for maze configurations that optimally disambiguate planning from vector navigation and habit learning. The fraction of informative states is defined as the fraction of choice points where planning and vector navigation disagree on action selection. The mean/variance centrality ratio is calculated for the distribution of the betweenness centralities across each of the 36 choice points within each maze layout. Blue line represents the Pareto optimal frontier. Different colors represent the number of elevated walkways in the layout (N_{edges}).

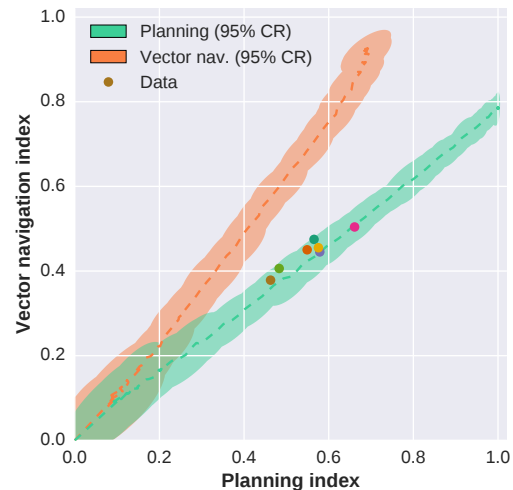


Figure 4: Analysis of decisions at choice points after training. Choices taken by mice (circles, $N=7$) are more compatible with planning (green region) than with vector navigation (orange region). The shaded regions represent the 95% confidence region for simulated behavior at a range of choice stochasticity levels. Note that because planning and vector navigation do not always yield opposing recommendations, the shaded green and orange regions delimiting results expected for planning and vector navigation agents do not lie directly on the respective index axis.