Hippocampal representations in a complex route planning task



Beatriz Simões Godinho

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024

Acknowledgements

Science is a marathon, not a sprint.

When I first heard this sentence, I had just started my journey into the neuroscience world, as a master's student. And even though I could understand the words, it was not until my years as a DPhil student that I could entirely assimilate and grasp their meaning. This DPhil was perhaps the most challenging project I ever had on my hands, but, if you are surrounded by the right team all the way through, it all becomes more enjoyable, and you'll be reminiscing the process in the end. Likewise, this DPhil was undoubtedly one of the most rewarding and amazing journeys I've ever been through, and all thanks to the people I was privileged to be surrounded by.

Firstly, I would like to thank my wonderful supervisory team: Tim Behrens, Thomas Akam and Mark Walton. Working with the three of them was never difficult, but always a pleasure, as they are all such kind and brilliant scientists who complement each other very well. I am grateful to them for believing in me, when they met me as a young master's student, and for having given me the opportunity to do this DPhil with them.

I thank Mark for all the supporting conversations and guidance during this journey and for always showing interest in the project. His total availability to discuss anything that worried me and to provide valuable advice has always made me feel reassured.

Thomas is the most versatile scientist I have ever met, and so I thank him for all that he has taught me, experimentally and computationally, and for all the days spent together building mazes and assembling probes, or outside in the rain spraying electrical boards - doing it together made it all more fun. I want to thank his hospitality too, for putting an effort into making me feel at home from the moment I arrived at Oxford, and for being such a kind and understanding human all the way through.

And to Tim. I want to thank Tim for always having my back during this project and for his contagious optimism every time we looked at the data - his passion for science is inspiring and I always knew that just by talking to him I would feel refreshed and more motivated than ever to solve the brain (again!). He is an incredibly generous scientist and person, always eager to teach but equally eager to learn - it's been an honour to learn how to do rigorous and elegant science with you, and to know how to always make the journey lighter and more enjoyable.

Additionally, I want to thank Steve Kennerley and Neil Burgess for the genuine interest, great feedback and generosity they brought to my Viva, which largely contributed for it to being a positive and happy memory I will always carry with me.

I feel fortunate for all the amazing people I have met in Mark's and Tim's labs. To Lauren, Inês and Sebastian for making me feel at home in Oxford from day one. To my gals Marta, Raquel and Merima with whom I shared every day of this DPhil, all the frustrations and achievements, and that will forever be a part of my life. A special thanks to Marta, for our immediate friendship and for being the most caring and lovely true friend from day one.

I want to thank Anna, Alon and Jacob for being my core within the lab from before I joined until the very end. A special note to Anna for our friendship built over sushi nights and love advice from the very start. From Oxford to London, the Behrens team is big and mighty, and I want to thank Avital, Diksha, Xiao, Mathias, Sandra, Kris, Francesca, Svenja, Will, Jo, Jiali, Ryan, James, Peter and all the others for all the office days, the fun projects and collaborations, the meaningful chats, for hanging out at conferences and ultimately for making the Behrens' lab such a stimulating and friendly work environment.

I also want to highlight the vibrant atmosphere I encountered at the Sainsbury Wellcome Centre, which provides great support for science to progress, thanks to its people - a special word to Elena, Mattias and Michael.

For all the years I spent in England, I always felt close to my Portuguese friends. For all the visits, the calls, the chats and the friendship, I want to thank Kiki, Madalena, Dario, Mónica, Ananim, Ananão, Lulu, Ana Rita, Miguel, Lucas e Ariana. Wherever we are, whenever we meet it's like time and distance have never come between us.

My family has been an unwavering source of love, comfort, and encouragement, providing me with the strength and support I needed to complete my PhD.

Thanks to my grandma Jaquina for constantly worrying about what I ate or whether I slept enough, for not speaking much but for always caring very deeply.

To my grandpa Piranho for being more than a grandfather, for his sense of humour and for always pushing me to try and exceed myself - I will carry forever the parts of you imprinted in me.

And to my mom. Mima, thank you for the unconditional love and support, for all the memories we've built together over the past 30 years, for always listening and giving the best advice and for being my confidant, the woman I admire the most and my greatest inspiration in life.

Finally, thank you Duarte for choosing this adventure with me, every day. Thank you for your patience, your humour, your intelligence, your kindness and your generosity. Thanks for always brightening up every single English greyish sky. Life is better with you.

Abstract

Behavioural flexibility is a hallmark of human and animal intelligence. It relies on internal models of the world - cognitive maps - which allow inference of the behavioural consequences of new information. Rodent spatial navigation is a powerful model system for studying the neural basis of cognitive maps, due to its ethological validity and our precise knowledge of how neurons represent space. However, most recordings come from simple behaviours and environments. A major open question is how spatial representations support model-based behaviours, such as planning, in complex and dynamic environments.

We recorded dorsal CA1 (dCA1) neurons as mice performed a complex route planning task, navigating to visually cued goal locations in an elevated maze, with different start and goal locations on each trial. We added or removed one link from the maze each day such that its structure gradually evolved over time.

Analysis of choices revealed mice use knowledge of the maze structure, often taking optimal paths even when they require heading away from the goal, showing rapid learning of the structural change within the first session. Further behavioural analysis provides preliminary evidence that mice plan compositionally with com-

monly used routes.

Neural data analysis reveals that dCA1 cells on the maze are not typical place cells. Few cells have a single localised firing field. Most firing fields are directionally selective, and many are spatially extended - entirely (rarely) or partially (commonly) - as if representing behavioural trajectories (routes). Dimensionality reduction of the neural space reveals neural factors mirroring the compositional elements of behaviour. Analysis of sharp-wave ripples (SWRs) in the task setting did not align with the view of online replay serving planning of future trajectories. However, we observe that an increase in the number of ripples during sleep supports mouse navigation performance in the following session, suggesting a role of offline replay in learning and consolidation of experiences.

Keywords: Hippocampus, Cognitive Maps, Navigation, Mouse, Maze

Contents

List of Figures xiii

Acronyms xix

1 Introduction **1**

1.1 Flexible behaviour 1

1.1.1 The Foundations: Darwin and Tolman 1

1.1.1.1 What came before 1

1.1.1.2 Cognitive maps - the origins 4

1.1.2 Relational memory maps 6

1.1.3 Goal-directed vs habits 7

1.1.4 Modern model-based decision-making 9

1.2 Hippocampus for flexible behaviour 15

1.2.1 Lesion studies 15

1.2.2 Anatomical and functional connectivities 19

1.2.2.1 Within the hippocampal formation 19

1.2.2.2 Beyond the hippocampal formation 21

1.2.3 Representations in hippocampus 23

1.2.3.1 Place cells 23

1.2.3.2 Remapping 28
1.2.3.3 Sequences 29
1.2.3.4 Replay 31
1.2.3.5 Other relevant representations 38
1.3 Theories and models of hippocampal flexibility 43
1.3.1 Planning in space 44
1.3.2 Successor representations 51
1.3.3 Models of replay 53
1.3.3.1 Replay as a mechanism for planning in reinforcement learning (RL) 55
1.3.3.2 Online and offline replay contributions for planning . 57
1.3.3.3 Replay as a mechanism of compositionality 58
1.4 Thesis outline 59

2 General Methods **63**

2.1 Subjects 63
2.2 Elevated grid maze 64
2.2.1 7x7 elevated grid maze 65
2.2.2 3x3 elevated grid maze 67
2.3 Goal-directed navigation task 68
2.4 dCA1 electrophysiological recordings experimental schedule 70
2.5 Data acquisition 72
2.6 Surgery 73

2.6.1 General 73
2.6.2 Silicon electrophysiology probe implant 74
2.6.2.1 2x 64-channel silicon probe stack assembly 75
2.7 Data pre-processing 75
2.7.1 Video processing 75
2.7.2 Electrophysiological signal processing 78
2.7.3 Signal alignment 79
2.8 Data analysis 80
2.9 Histology 80

3 A sequence of maze structures optimiser for complex navigation testing 85

3.1 Introduction 85
3.2 Methods 88
3.2.1 Optimisation tools 89
3.2.1.1 Genetic algorithm (GA) 89
3.2.1.2 Pareto Front 91
3.3 Results 92
3.3.1 Single maze optimiser 92
3.3.1.1 First iteration - genetic algorithm and Pareto fronts . 94
3.3.1.2 Second iteration - exhaustive search from random maze 96

3.3.1.3 Third iteration - exhaustive search with a maximum limit of structures per step 98
3.3.1.4 Fourth and final iteration - maximising euclidean-geodesic (EG) discriminability and simple rules . . . 99
3.3.2 Sequence of mazes optimiser 101
3.3.2.1 Building a sequence of maze structures 101
3.3.2.2 The chosen sequence 106
3.4 Discussion 110

4 Mice quickly integrate new spatial information into their navigation strategy under a goal-directed paradigm - and other behavioural signatures 113

4.1 Introduction 114
4.2 Methods 116
4.2.1 3x3 grid maze training schedule 116
4.2.2 Mixture of strategies model 118
4.3 Results 120
4.3.1 Mice training in a 3x3 grid maze 120
4.3.2 General behavioural signatures of mice navigation in a 7x7 grid maze 122
4.3.3 Mice incorporate spatial changes in their navigation strategy within the first session of experience 131
4.4 Discussion 134

5 Hippocampal representations in a complex and dynamic maze environment 141

5.1 Introduction . . . 141

5.2 Methods . . . 144

5.2.1 Linearisation of the maze structure . . . 144

5.2.2 Dimensionality reduction of neural space - PCA and NMF . . 145

5.3 Results . . . 146

5.3.1 Complex hippocampal maps in a complex environment 147

5.3.2 Directional maps . . . 149

5.3.3 Dimensionality reduction of the hippocampal neural space . . 152

5.4 Discussion . . . 154

6 Online and offline replay of an intricate hippocampal cell coding system 161

6.1 Introduction . . . 161

6.2 Methods . . . 164

6.2.1 Ripple detector . . . 164

6.2.1.1 local field potential (LFP)-based ripple detector . . . 164

6.2.1.2 Spike-based ripple detector . . . 165

6.2.2 Spatial decoders . . . 166

6.2.2.1 Linear decoder . . . 166

6.2.2.2 Bayesian decoder . . . 167

6.3 Results . . . 169

6.3.1 Hippocampal representations allow accurate position decoding in a complex environment 169
6.3.2 Online replay analyses 170
6.3.2.1 Hippocampal reactivation at planning times 173
6.3.2.2 Hippocampal reactivation at reward consumption times 176
6.3.3 Offline replay analyses 176
6.4 Discussion 184

7 Conclusion **191**

Bibliography **201**

A Supplements **251**

List of Figures

1.1 Tolman's cognitive maps. 5

1.2 Spatial maps as a type of memory space. 7

1.3 RL model-free and model-based agents. 11

1.4 The two-step task. 13

1.5 Hippocampal lesions in rats. 17

1.6 Hippocampal formation connectivity. 21

1.7 Hippocampal place cells and their features. 26

1.8 Hippocampal replay. 35

1.9 Spatial representations outside of the hippocampus. 42

1.10 RL Dyna-Q planning. 45

1.11 Euclidean vs geodesic navigation models. 50

1.12 The successor representation model. 54

1.13 Replay through memory access prioritisation. 56

2.1 7x7 grid maze apparatus. 67

2.2 Discrete coordinate system in the 7x7 grid maze 68

2.3 3x3 grid maze apparatus. 69

2.4 Goal-directed navigation task. 70

2.5 Hippocampal recordings experiment schedule. 72
2.6 Schematic of the probe ground screw location in the mouse skull. . . 74
2.7 128 channel dual mini-amp single nanodrive implant. 75
2.8 Schematic DLC labelling of mouse parts. 77
2.9 Histological assessment of fibre placement - first recordings cohort. 82
2.10 Histological assessment of fibre placement - second recordings cohort. 83

3.1 Schematic of the genetic algorithm steps used for the individual maze structure optimisation. 91
3.2 Maze example of vector and geodesic navigation strategies on the maze. 93
3.3 Variation in maze structure with its capacity to discriminate between vector navigation and planning 94
3.4 Individual maze structure optimisation - first iteration result. 96
3.5 Individual maze structure optimisation - second iteration result. . . . 97
3.6 Individual maze structure optimisation - first iteration result. 99
3.7 Illustration of the open and closed connected components concept. . 100
3.8 Additional rules for optimising a sequence of maze structures (I) . . 104
3.9 Additional rules for optimising a sequence of maze structures (II). . . 105
3.10 Chosen sequence of maze structures used in this project's experiments (I). 106
3.11 Chosen sequence of maze structures used in this project's experiments (II). 107

3.12 Chosen sequence of maze structures used in this project's experiments (III). 108
3.13 Chosen sequence of maze structures used in this project's experiments (IV). 109

4.1 Initial 3x3 maze training schedule. 117
4.2 Optimised 3x3 maze training schedule. 117
4.3 Vector and planning strategies model cartoon. 119
4.4 Vector, planning and structure change integration strategies model cartoon. 120
4.5 Training performance on the 3x3 grid maze - behaviour and first recordings cohorts. 122
4.6 Vector, planning and structure change integration strategies model cartoon. 123
4.7 Overall performance on the 7x7 grid maze. 124
4.8 Behavioural trials in the 7x7 grid maze. 125
4.9 Variation of mouse velocity across sessions. 127
4.10 Reward distribution across the 7x7 maze structure. 128
4.11 Time spent occupancy distribution across the 7x7 maze. 129
4.12 Number of visits occupancy distribution across the 7x7 maze. 130
4.13 Mice exhibit a mixture of vector and shortest path navigation strategies for solving the task. 131
4.14 Vector vs shortest path navigation across maze structures. 132

4.15 Mice integrate the changed-link information into their navigation strategy within the first session of experience. 134
4.16 Mice integrate the changed-link information differently depending on the nature of the change. 135

5.1 The ”skeleton” maze structure. 145
5.2 Schematic of the neural space dimensionality reduction analysis steps. 146
5.3 Detailed dCA1 neuronal activity session firing maps. 148
5.4 Directional session firing maps of dCA1 neuronal activity. 150
5.5 Neural factors of hippocampal activity in the maze. 153

6.1 Schematic of the linear decoder analysis steps. 167
6.2 Recorded dCA1 neural activity allows decoding of the mouse position in the maze. 170
6.3 Bayesian decoder accuracy. 171
6.4 Trial ripple activity. 172
6.5 Online planning window ripple activity maps. 174
6.6 Decoding replay representations at planning and reward consumption times. 175
6.7 Decoding replay representations at planning and reward consumption times. 177
6.8 Mean SWR firing rate (Hz) at reward consumption increases with the length of the trial’s trajectory. 178
6.9 Post-change sleep session decoded ripple firing maps. 181

6.10 Increase in offline SWRs promotes a better performance of goal-directed navigation in a previously experienced complex maze structure. 184

List of Acronyms

A-O Action-Outcome.

ACC Anterior Cingulate Cortex.

BD Between-Maze Discriminability.

BLA Basolateral Amygdala.

CC Connected Components.

CF Cost Function.

dCA1 Dorsal CA1.

DG Dentate Gyrus.

EC Entorhinal Cortex.

EG Euclidean-Geodesic.

HPC Hippocampus.

LFP Local Field Potential.

MB Model-Based.

MEC Medial Entorhinal Cortex.

MF Model-Free.

mPFC Medial Prefrontal Cortex.

NMF Non-Negative Matrix Factorisation.

OFC Orbitofrontal Cortex.

PaS Parasubiculum.

PCA Principal Component Analysis.

PFC Prefrontal Cortex.

PI Path Integration.

PrS Presubiculum.

PSR Probabilistic Successor Representation.

RL Reinforcement Learning.

RPE Reward Prediction Error.

RTDP Real-Time Dynamic Programming.

S-R Stimulus-Response.

SR Successor Representation.

SWR Sharp-Wave Ripple.

SWS Slow-Wave Sleep.

VR Virtual Reality.

VS Ventral Striatum.

VTE Vicarious Trial and Error.

WD Within-Maze Discriminability.

1

Introduction

Section 1.1

Flexible behaviour

1.1.1 The Foundations: Darwin and Tolman

The concept of cognitive maps has its roots in the early stages of behavioural psychology, and so we should first do a quick dive through its precursors in the study of animal intelligence for context.

1.1.1.1 What came before

In the second half of the 19th century, there was a clear divide in the interpretation of animal behaviour and in how it should be tested and thought of in relation to

the human one. On one side, naturalists like Charles Darwin and his pupils, Romanes and Lubbock, considered human and animal minds to be comparable and explained them in terms of instinct, reason, and intelligence (Darwin, 1859, 1871; Lubbock, 1889; Romanes, 1886).

> *(...) the difference in mind between man and higher animals, great as it is, certainly is one of degree and not of kind.*
>
> Darwin (1871)

> *(...) there is no fundamental difference between man and the higher mammals in their mental faculties.*
>
> Romanes (1886)

On the other side, unconvinced by the anecdotal proof presented by the naturalists, Conwy Lloyd Morgan argued for a less anthropomorphising view of animals, and a more systematic approach to the study of their intelligence (Morgan, 1886). Known as the father of comparative psychology, Morgan interpreted animal behaviour as a product of a ”direct association of ideas” between a sense (stimulus, action) and an experience (another stimulus, action, result), not as a product of reason (Morgan, 1886, 1894).

Opposed to relying on simple observation or anecdotal evidence, Edward Thorndike

emerged as perhaps the first experimentalist to thoroughly describe and design rigorous animal testing paradigms. He systematically tested animals (cats and dogs in the context of food deprivation, and chicken in the context of social deprivation) put in a 'puzzle-room' from where they should escape to get a reward (food or company, for cats and dogs, and chickens, respectively). After freeing themselves for the first time, through the performance of some action discovered by trial-and-error, the animals would get quicker at escaping in subsequent tests.

Regardless, Thorndike concluded that *"(...) these reactions [could] all be explained by the ordinary associative processes without aid from abstract, conceptual, inferential thinking."* and elaborated these ideas into what is known as the Law of Effect (Thorndike, 1898, 1911). Whilst this principle became the foundation of operant conditioning and influenced the work of other behaviourists, such as B.F. Skinner (Skinner, 1953); a distinct approach to behavioural psychology was emerging in Germany - Gestalt psychology - which defended animals' ability to use prior knowledge to simulate possible actions they had never seen and create entirely new solutions (Köhler, 1917).

> *(...) responses that produce a satisfying effect in a particular situation become more and more likely to occur again in that situation, and responses that produce a discomforting effect become less likely to occur again in that situation*

The idea behind the Law of Effect by Thorndike (1898)

1.1.1.2 Cognitive maps - the origins

In 1932, Edward Tolman, inspired by the Gestalt school of thought, carved a new approach to psychology - Purposive Behaviourism - in a book he dedicated to the rats he typically used in his experiments: *"Lastly, in inscribing this book to M.N.A. ('Mus norvegius albinus') I have sought to indicate where perhaps, most of all, the final credit or discredit belongs."* (Tolman, 1932).

In 1948, Tolman published what would become his most famous and influential paper, proposing the existence of a complex mental representation of the environment - a cognitive map (Tolman, 1948a). His theoretical perspective defended

> *Firstly,* ***that learning consists*** *not in stimulus-response (S-R) connections but* ***in the building up*** *in the nervous system of sets which function like* ***cognitive maps****, and second, that such cognitive maps may be usefully characterized as* ***varying from a narrow strip variety to a broader comprehensive variety****.*
>
> Tolman (1948a)

To support his theory, Tolman reported on five different experiments performed by rats in distinct mazes. In one of these, a *latent learning* experiment (Blodgett, 1929), three groups of rats were put on a maze to navigate. While the control group found food in the goal location at the end of the trials from day one and presented a

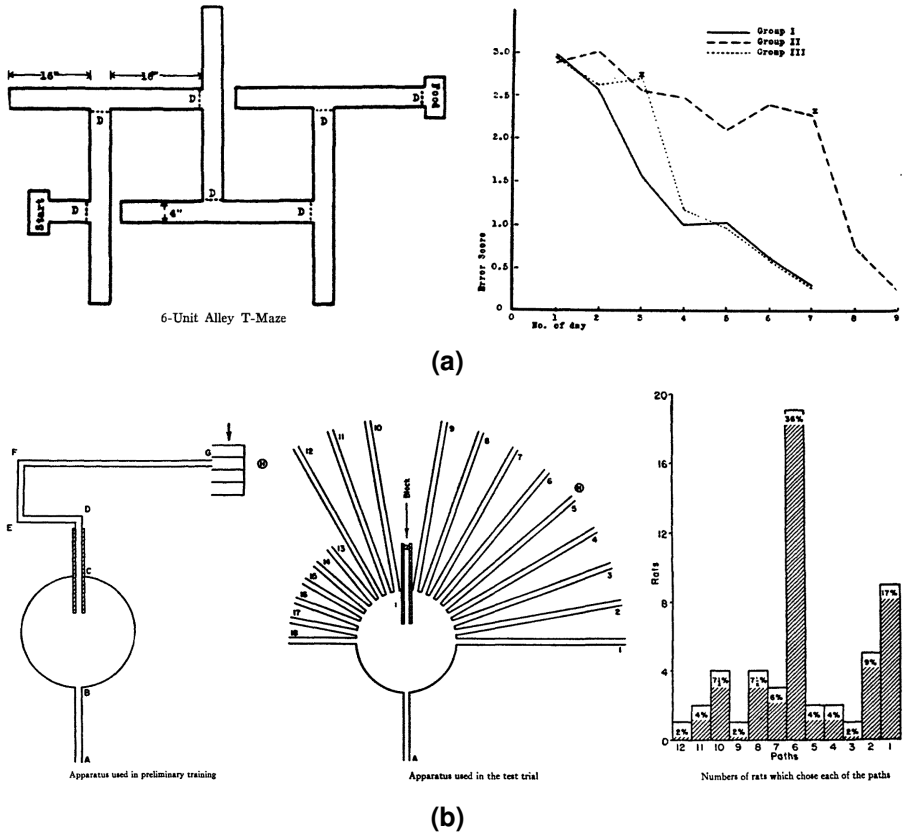


Figure 1.1: Tolman's cognitive maps. (a) Alley T-maze (left) used by Blodgett (1929) to demonstrate latent learning, where groups I, II and III of rats run the maze with reward available at the end from day 1, from day 3 and from day 8, respectively. (right) After the first rewarded day for groups II and III the number of errors dropped massively, suggesting learning of the space in absence of rewards. (b) Spatial orientation task, where rats demonstrated shortcut behaviour when transitioning between environments (from left to centre maze), exhibiting preference for the novel arm that would take them in the same Euclidean direction of the previous maze goal. Adapted from Tolman (1948a).

slow and steady improvement in performance; the other two groups did not find any sort of reward in the maze until days three and seven of nine, respectively. In the days without reward, both test groups presented a consistently low performance (i.e. high error rate, high number of entries in alleys before reaching the end of the trial). However, on the day immediately after the introduction of food at the goal-box, both groups massively improved their performance, and at a higher pace than that observed in the control group, showing that rats could learn a map of space in

the absence of reinforcers.

In another experiment, of *spatial orientation*, Tolman and colleagues showed that rats would quickly find new routes to experienced goal locations after major changes in the environment (Tolman et al., 1946). In sum, these and other results supported the existence of cognitive maps in animals and, furthermore, built the intuition that these maps are comprehensive and not narrow, which allows for greater behavioural flexibility and adaptation, just like in humans (Tolman, 1948a).

1.1.2 Relational memory maps

The concept of cognitive maps has emerged in association with animal navigation and spatial tasks, as these are ethological platforms for the study of internal model representations in animals. Nonetheless, the idea of cognitive maps is not specific to space but rather generalises to non-spatial, more abstract scenarios.

For Cohen and Eichenbaum (1994), this system should be thought of as providing a general relational processing mechanism that supports declarative memory. The hippocampus (HPC) forms a ”memory space” of relational representations that bind the elements of experiences in memory and associate different memories based on their shared components (Eichenbaum, 1999). This is achieved by this system’s capacity to rapidly establish associations between any subset of its inputs and between these inputs and reactivated relational memories. The hippocampal formation thus plays a key role in the generation, recombination and flexible use of knowledge across species and behavioural domains beyond the

spatial one (Eichenbaum and Cohen, 2014).

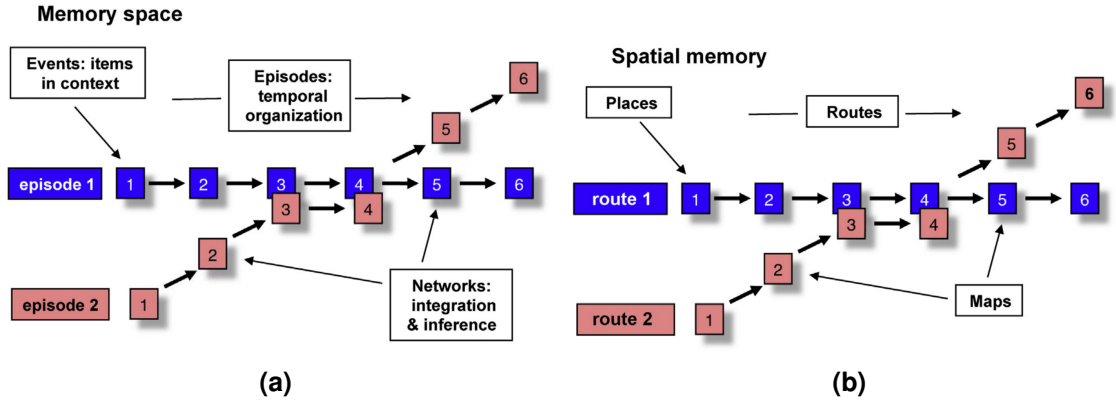


Figure 1.2: Spatial maps as a type of memory space. (a) While integration and inference are supported by mechanisms of 'memory space' where events are experienced across different episodes that are organised into networks, (b) spatial maps are a special case where places are spatial memories we traverse along distinct routes. These routes then connect at overlapping locations thus creating an internal map that relates places encountered at different times. Adapted from Eichenbaum and Cohen (2014)

1.1.3 Goal-directed vs habits

Flexible behaviours are often goal-directed. Goals, rewards or positive reinforcers, seem to be everywhere in our daily lives and to direct almost our every action. For us humans, they can take many forms, whether it is having dinner after having skipped lunch, a glass of water after exercising, meeting friends for a chat, getting a compliment from someone we like, or achieving something of great desire (like submitting your PhD thesis!). For other animals, to our knowledge, rewards tend to be less complex. For instance, in Tolman's maze experiments, sometimes rats had to find their way through the corridors towards a food reward that was waiting for them in the reward box, and this was enough to motivate them to perform efficiently. Moreover, in cases where it was not available (the latent learning experiment described in Section 1.1.1.2), when a reward was introduced in the environment, it

caused drastic improvements in performance (Tolman, 1948a).

With the purpose of understanding the learning processes guiding flexible behaviours, such as goal-directed ones, experimenters designed *reinforcer devaluation* experiments. With the aim of testing if the animals knew the consequences of their actions and had, thus, learned the instrumental contingency (e.g. that pressing a lever leads to reward delivery), they were initially trained to perform an action to get a reward (i.e. pressing a lever for water). After the behaviour was established, the outcome was devalued (e.g. by providing free access to the reward), and subjects were then tested through extinction - animals were put in the same instrumental context but in the absence of rewards. The cognitive map enthusiasts predicted that the animals would decrease their response to the instrumental condition in the absence of an outcome after devaluation, whereas, if the response level was not affected, the conclusion would be that the animals had learnt/reinforced a simple S-R association instead of building a model of the environment. Inconclusively, both results were true for different sets of experiments (found devaluation effects: (Tolman and Gleitman, 1949; Krieckhaus and Wolf, 1968); did not find devaluation effects: (Tolman, 1933; Rozeboom, 1957; Dickinson et al., 1983)).

In 1983, Dickinson and colleagues proposed to get to the bottom of the question and conducted another series of instrumental learning experiments in rats. Interestingly, they observed that the sensitivity of instrumental behaviour to outcome devaluation depends on the amount of training the animals receive. After a moderate amount of training, rats' lever pressing was sensitive to outcome devaluation, indicating goal-directed behaviour. However, after extensive training, the rats' be-

haviour became habitual and insensitive to outcome devaluation (Dickinson et al., 1983). This discovery led to the proposal of a dual-process theory of instrumental learning.

According to this theory, *goal-directed behaviour* is driven by the encoding of action-outcome (A-O) contingencies, i.e. by actions that are performed with the intention of achieving a specific goal or outcome. Such behaviour is flexible and can be quickly adapted when the value of the outcome changes or when the contingency between the action and the outcome is altered (Dickinson, 1985; Balleine and Dickinson, 1998). Conversely, *habitual behaviour* is driven by the formation of S-R associations that render it not flexible but automatic and triggered by specific stimuli in the environment rather than by the anticipated consequences of the action (Dickinson, 1985). Habits are acquired through the repeated performance of an action in a stable context, leading to the formation of S-R associations that can be executed without conscious deliberation (Yin et al., 2006).

1.1.4 Modern model-based decision-making

The distinction between goal-directed and habitual behaviour laid the foundation for the modern study of animal learning and decision-making. Inspired by it, researchers have begun to explore a related but distinct dichotomy: the distinction between *model-free (MF)* and *model-based (MB) reinforcement learning (RL)* (Sutton and Barto, 1998). This distinction has its roots in the field of *RL*, a framework through which individuals learn how to act in order to maximise reward. Here, solu-

tions should be discovered by trial-and-error, using feedback (positive or negative reinforcers) to guide action selection. This has proven to be a successful framework for describing many features of decision-making and learning and identifying the substrates of such computations in the brain (Sutton and Barto, 1998; Schultz et al., 1997). However, while the distinction between goal-directed and habitual behaviour focuses on the sensitivity of actions to changes in outcome value, the distinction between MF and MB learning focuses on the use of an internal model of the environment to guide decision-making (Balleine and Dickinson, 1998; Daw et al., 2005).

In 2005, Daw, Niv and Dayan mapped the distinction between goal-directed and habitual behaviours in psychology onto the RL distinction between MB and MF algorithms (Daw et al., 2005). A MF RL agent does not have an explicit map of the environment's transition structure and so it cannot predict which states will different actions lead to. It does, however, maintain a representation of the environment by receiving at each timestep information about the current state, and it learns about the consequences of actions strictly in the sense of them leading (or not) to reward Figure 1.3(a). While computationally simple, this type of learning cannot rapidly adapt to changes in the environment or the value of rewards. In contrast, a MB RL agent, as the name suggests, implies the construction of an internal model of the transition structure of the environment, which represents the relationships between different states, actions, and outcomes, and thus allows the prediction of future states based on the current state and chosen action Figure 1.3(b). Although more flexible and allowing quicker adaptation to change, this kind of learning can be

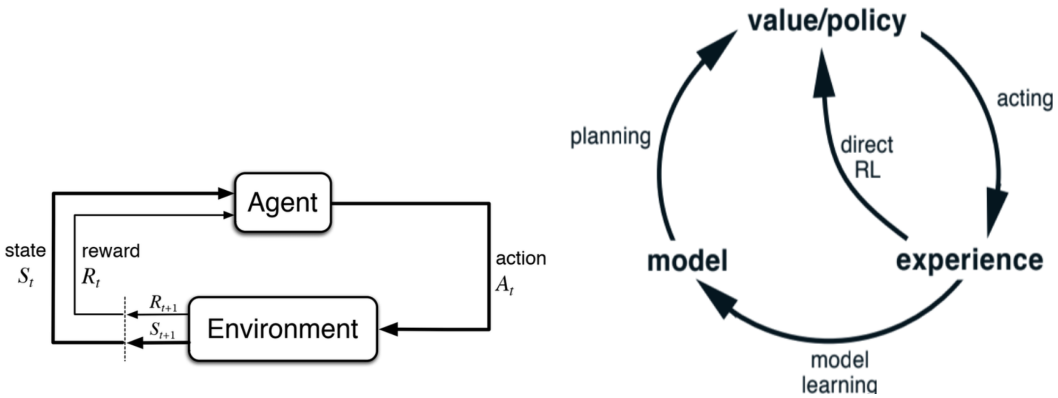


Figure 1.3: RL model-free and model-based agents. (a) agent-environment interaction in a Markov decision process, i.e. a process where all future states only ever depend on the current state, and not on the previous history (b) interactions in a Dyna-Q model. In an interaction between MF and MB, a model of the environment is learned through experience. This model simulates additional experience for planning, which is used to update a value function and policy that are also updated through direct experience in a MF manner (adapted from Sutton and Barto (1998)).

highly computationally demanding, especially in large state-space environments.

Therefore, neither MF nor MB learning are entirely satisfying on their own. Thus, the authors proposed that, although the two learning modes can coexist in the brain, an uncertainty signal in the brain would arbitrate between the two and select, at each moment, the dominating system to guide behavioural choices. This uncertainty mediation, initially hypothesised to rely on the locus coeruleus (LC) noradrenergic system (Daw et al., 2005; Aston-Jones and Cohen, 2005; Yu and Dayan, 2005; Jepma and Nieuwenhuis, 2011), remains itself not entirely understood, with its role being more recently associated with the dopaminergic and cholinergic systems (Doya, 2008; Pezzulo et al., 2013), as well as the prefrontal cortex (PFC) (Lee et al., 2014). MF and MB behaviour is also thought to be mediated by different areas. MF learning is primarily associated with the dorsolateral striatum, also involved in habit formation (Yin et al., 2006), while MB learning re-

lies on the PFC (prelimbic (PL) region) and dorsomedial striatum, responsible for goal-directed behaviour and representations of A-O contingencies (Balleine and O'Doherty, 2010; Yin et al., 2005).

Several experimental paradigms have since then emerged in the hope of disambiguating the contributions of these learning systems to decision-making and unravelling the computational mechanisms behind them. The two-step task is one of them (Daw et al., 2011). In this task (Figure 1.4(a)), participants make a series of choices between two options, which probabilistically lead to one of two second-stage states, that will themselves probabilistically lead to a rewarded or non-rewarded state. Importantly, the transition probabilities between the first-stage choices and second-stage states remain fixed, while the reward probabilities at the second stage change over time. The extent to which subjects are employing a MB or MF strategy can be extracted by their choice patterns in this task - while a MB strategy promotes staying or switching behaviour based not solely on the outcome experienced in the previous trial, but also on the transition that occurred; a MF strategy will lead to repeating or switching actions only based on the outcome received - reward or omission. Evidence advocated for a hybrid system in humans (Daw et al., 2011) (Figure 1.4(c)), with subsequent studies identifying neural correlates of these two forms of learning in the striatum and prefrontal cortex (e.g., Daw et al. (2011); Smittenaar et al. (2013)).

The success of this task translated into several adaptations to rodent models, permitting a deeper investigation of these RL algorithms via the use of neural recordings and manipulation techniques. In a rat adaptation of the two-step task,

Miller et al. (2017) probed a causal role of the orbitofrontal cortex (OFC) and dorsal CA1 (dCA1) in MB behaviour. In another simplified version, this time for mice and with additional reversals in the transition probabilities (Figure 1.4(b)), activity in the anterior cingulate cortex (ACC) is shown to be modulated by future states, and its inactivation impairs MB choices (Akam et al., 2021).

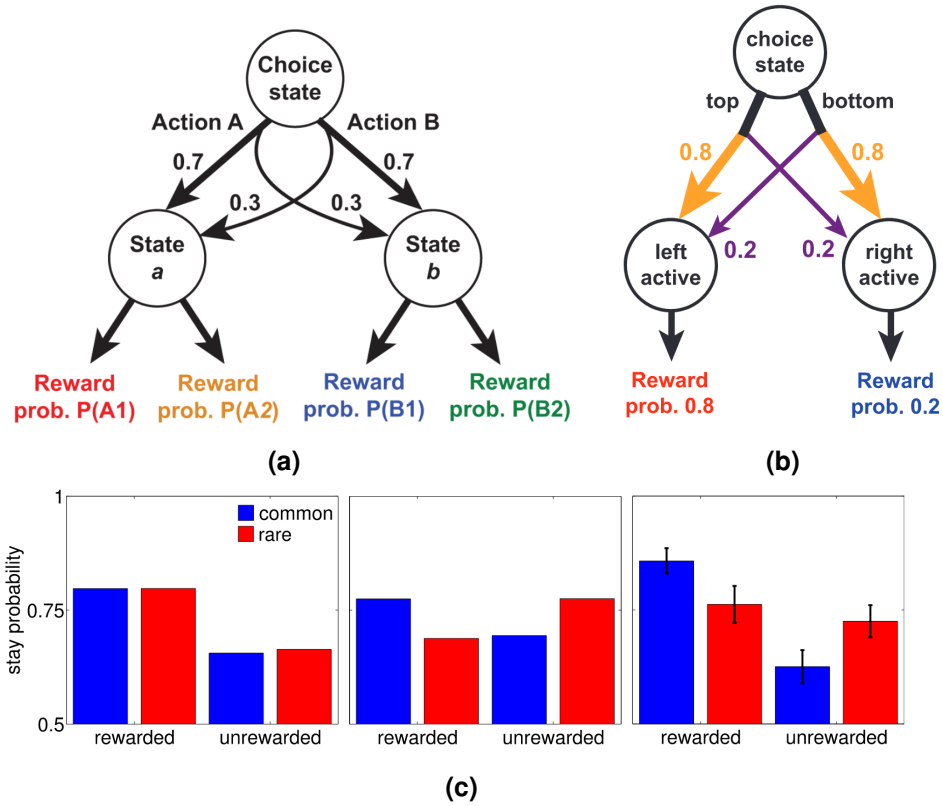


Figure 1.4: The two-step task. (a) schematic by (Akam et al., 2015) of the original version of the two-step task, and (b) of the adapted version of it for rodent testing (adapted from Akam et al. (2021)). (c) probability of repeating a choice in the first step as a function of the previous trial outcome and transition probability - MF subjects repeat previously rewarded action independent of the nature of the transition (left), while MB subjects modulate their choice based not solely on the outcome but also the transition type of the previous trial (centre). Human subjects responses (right) (adapted from Daw et al. (2011).)

These same decision-making processes and more complex sequences of decisions can also be assessed using a spatial context and even in physical space. An example of that is Hasz and Redish (2018) adaptation of the two-step task to

a spatial maze for rats through which they concluded that, like humans, rats use a hybrid strategy for solving this task, with OFC activity and vicarious trial and error (VTE) behaviour being correlated with MB choices. Moreover, Simon and Daw (2011) found evidence for MB learning in the HPC and PFC of humans while they ”navigated” a series of choices to find a rewarded location. Virtual reality (VR) environments have also been popularised to investigate the relationship between MB learning and spatial memory, with studies concluding that subjects with better spatial memory are more likely to use a MB strategy during navigation (Penny et al., 2013) and that this is modulated by activity in the OFC and HPC (Vikbladh et al., 2019).

Alternatively, another well-known strategy that can analogously explain the flexibility inherent to MB behaviours is Bayesian inference. Defined by the computation of a belief state that combines relational knowledge between stimuli or events (Brogden, 1939), inference-based tasks are widely employed in the study of the neural mechanisms behind MB decision-making, and have so far highlighted the crucial role of the OFC, ACC, medial prefrontal cortex (mPFC) and dCA1 in the formation, updating and use of internal models to guide behaviour (Hampton et al., 2006; Behrens et al., 2007; Costa et al., 2016; Bartolo and Averbeck, 2020; Vertechi et al., 2020).

Section 1.2

Hippocampus for flexible behaviour

As mentioned in Section 1.1.4, HPC has been implicated in MB decision-making in humans and animals, suggesting an important role for this brain region in shaping behaviour. In this section, we review the literature investigating the role of HPC in flexible behaviour.

1.2.1 Lesion studies

Lesion studies in humans were initially led by the desire to fix certain conditions, such as epilepsy, yet ended up being a gateway to our understanding of the function of various brain regions, like the HPC.

In the 1950s, lesions in humans and rodents concluded that memory should be located in the brain, albeit not agreeing on the extent of its distribution within it (Penfield and Rasmussen, 1950; Lashley, 1950). Soon after, Scoville and Milner (1957) used the same technique to attribute its dependence to a specific brain region—the HPC. In the famous H.M. patient study, bilateral resection of the medial temporal lobe (MTL) caused *"(...) one striking and totally unexpected behavioural result: a grave loss of recent memory (...)"*. The patient's mother reported that he did *"(...) not know where objects in continual use [were] kept (...) [and that] he will do the same jigsaw puzzles day after day without showing any practice effect (...)*

[while] his understanding and reasoning [were] undiminished.". Such observations motivated the recognition of memory as a separate function in the brain and suggested the hippocampal formation to be crucial for declarative memory, a subtype of memory that can be recalled and declared (Cohen and Squire, 1980).

Around the same time, lesion studies in animals further elucidated the role of the HPC in memory, but also in spatial cognition. Unlike the common behavioural response inhibition observed in rats with hippocampal damage at that time (Isaacson, 1982; Kimble, 1968), HPC started to be associated with behavioural flexibility. While hippocampal lesioned rats successfully learned to go to one arm of a T-maze, they needed many more trials to learn a reversal of this preference (Douglas, 1967; Greene and Stauff, 1974; Kimble, 1963), and also failed to follow spontaneous or rewarded alternation patterns (Johnson et al., 1977; Sinnamon et al., 1978). Morris et al. (1982) had hippocampal lesioned rats perform a spatial navigation and memory task in a water maze. Lesioned rats struggled to locate a hidden platform (just below the water surface) compared to control animals (Figure 1.5(a)). These impairments were observed across trials, including those after the animals had seen where the platform was located, and were proportional to the extent of the lesion. This indicated an essential role of HPC in spatial learning and memory. Moreover, Olton and Papas (1979) confirmed it to be crucially involved when flexible behaviour is required - hippocampal lesioned rats were significantly impaired in a working memory radial arm maze task, while their performance was not affected in a reference memory task in the same apparatus (Figure 1.5(b)).

In mice, cytotoxic lesions of the HPC lead to the impairment of spatial work-

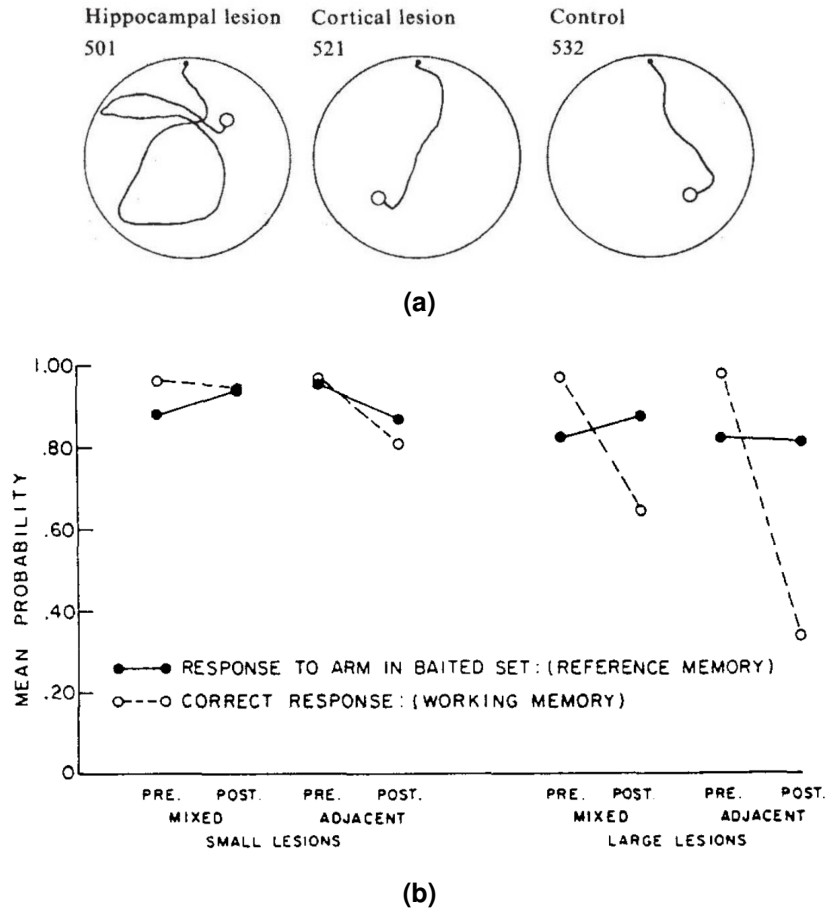


Figure 1.5: Hippocampal lesions in rats. (a) Rats in a water maze struggle to find a hidden platform just below surface level even after having seen it in previous trials (adapted from Morris et al. (1982)), while (b) in a radial maze lesioned rats show impaired working memory but not reference memory, with the impairment being proportional to the lesion size (adapted from Olton and Papas (1979)).

ing memory in a T-maze delayed alternation task (Bannerman et al., 1999) and a radial arm maze (Schmitt et al., 2003). Furthermore, hippocampal lesions disrupt the ability of mice to form integrative memories that combine spatial and nonspatial knowledge (Deacon et al., 2002). The development of genetic tools allowed for more localised and temporally precise inactivation of the mouse HPC. For example, mice with a conditional knockout of the NMDA receptor in the CA1 region of the HPC exhibited impairments in spatial memory (Tsien et al., 1996), while opto-

genetic silencing of hippocampal neurons disrupted the retrieval of contextual fear memories (Goshen et al., 2011). While these results provide further evidence for the role of the HPC in spatial learning and memory consolidation in mice, other experiments with selective inactivation of its subregions (Hock and Bunsey, 1998; McHugh et al., 2007) conclude spatial working memory to especially rely on dorsal but not ventral HPC.

Hippocampal lesions in monkeys have also been shown to lead to deficits in delayed non-matching-to-sample tasks (Zola-Morgan et al., 1989) and object-in-place memory (Gaffan, 1994), thus impairing associations between spatial and non-spatial components of space.

In the context of habitual vs goal-directed behaviour (Section 1.1.3), the HPC has been suggested to mediate the arbitration between these two modes of behaviour, favouring a MB strategy and preventing over-generalisation across contexts (Packard and McGaugh, 1996; McDonald and White, 1993; Redish, 2016). It has further been implicated in the rapid updating of A-O contingencies when environmental conditions change (Bradfield et al., 2020), and in the use of prospective memory to guide goal-directed behaviour, which allows animals to flexibly plan and execute actions based on anticipated future outcomes (Doll et al., 2015; Miller et al., 2017; Vikbladh et al., 2019). Another example that well illustrates the HPC as a flexible memory hub is the inability to combine memories into novel scenarios, i.e. to imagine new experiences, in hippocampal-lesioned patients (Hassabis et al., 2007). Additionally, the recombination of information critically depends on this network in humans (Preston et al., 2004; Zeithamova et al., 2012) and in rats

(Bunsey and Eichenbaum, 1996; Dusek and Eichenbaum, 1997) when performing transitive inference or paired-associate learning tasks; as does the ability to generalise knowledge through the abstraction of commonalities between experiences (Kumaran et al., 2009; Kumaran and McClelland, 2012; Koster et al., 2018).

1.2.2 Anatomical and functional connectivities

To better understand the various roles of the HPC in the generation and updating of cognitive maps, it is vital to consider the anatomical disposition of the different subregions of the HPC, and its connections to other brain regions within and outside of the hippocampal formation.

1.2.2.1 Within the hippocampal formation

Located in the medial temporal lobe, the hippocampal formation encompasses four principal regions (Figure 1.6) - the entorhinal cortex (EC), the subicular complex (divided into subiculum, presubiculum (PrS) and parasubiculum (PaS)), the dentate gyrus (DG) and the hippocampus proper (Witter et al., 2000). The hippocampus proper comprises the Cornus Ammonis (CA) regions: CA1, CA2 and CA3 (Amaral and Witter, 1989; Andersen et al., 2007). Anatomically, the hippocampal formation retains a largely conserved structure across species, with some size and shape/orientation differences (Amaral et al., 2007; Clark and Squire, 2013).

Concurrently, the connectivity patterns within the hippocampal formation and between the HPC and other brain regions are also highly conserved across species. An example of that is the trisynaptic pathway, originally described by Santiago Ra-

mon y Cajal, and considered the main route for hippocampal information processing in mammals (Andersen et al., 2007). It is composed of three main synaptic connections: from the EC to the DG, from the DG to the CA3 and from the CA3 to the CA1(Amaral and Witter, 1989). It starts with the passage of information from the EC layer II neurons to the DG via the perforant path (Witter et al., 2000), conveying multimodal sensory information from various cortical areas (Amaral et al., 2007). On the second synapse of the pathway, the DG granule cells axons - mossy fibres - transmit the information to CA3 pyramidal neurons, to facilitate the rapid encoding of new information (Amaral et al., 2007; Rollenhagen and Lübke, 2006; Rolls, 2013). Lastly, the third synapse of the trisynaptic circuit will connect the CA3 pyramidal cell axons - Schaffer collaterals - and the CA1 pyramidal neurons.

Importantly, the CA1 then projects back to EC via the subiculum, forming a loop mechanism for memory formation and consolidation (Knierim et al., 2006). The EC serves as the main interface between the HPC and the various cortical areas, receiving input from the perirhinal and parahippocampal cortices, which are believed to process object and spatial information, respectively (Witter et al., 1989; Burwell, 2000; Eichenbaum et al., 2007).

The DG will then transform similar input patterns into distinct neural representations through a process called pattern separation (Leutgeb et al., 2007). Here, the extensive recurrent connections within the CA3 region allow the retrieval of stored representations from partial or degraded cues - pattern completion (Knierim et al., 2006).

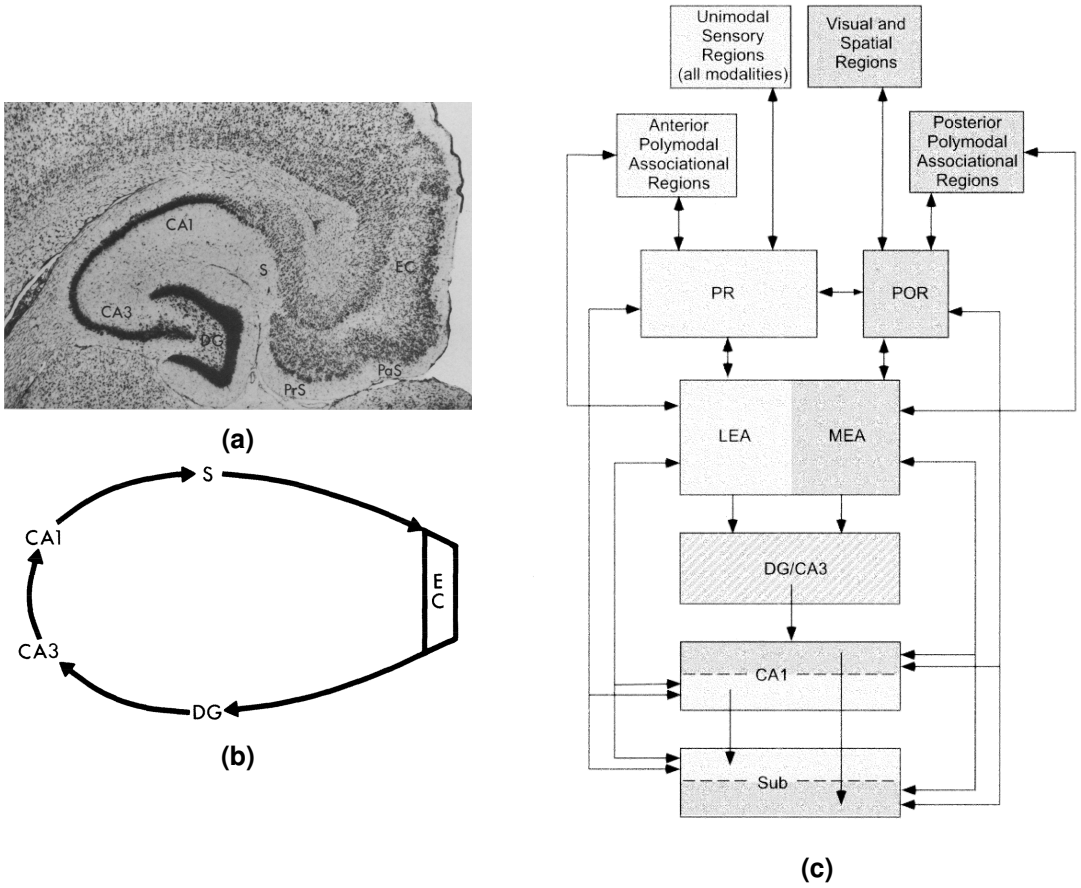


Figure 1.6: Hippocampal formation connectivity. (a) rodent hippocampal subregions - hippocampus proper (subdivided into CA1, CA2 and CA3), DG, EC and subicular complex (subdivided into subiculum (S), PrS and PaS); (b) summary diagram of the trisynaptic pathway loop ((a) and (b) adapted from Amaral and Witter (1989)); (c) schematic of the corticohippocampal circuitry (adapted from Burwell (2000))

1.2.2.2 Beyond the hippocampal formation

Thus far, we have looked at some of the main connections within the hippocampal formation supporting memory formation and consolidation; here, we will discuss some of the key connections between the hippocampal formation and various other brain regions that are essential for information integration.

One of the most relevant connections is between the HPC and the mPFC region, which has been implicated in memory consolidation and retrieval, as well as

integration of new spatial and contextual information (Preston and Eichenbaum, 2013). On the one hand, the CA1 and subiculum regions send direct projections to the mPFC (Hoover and Vertes, 2007), thought to support the transfer of information for long-term storage and the integration of spatial and contextual coding during decision-making behaviours (Euston et al., 2012). On the other hand, the mPFC projects back to the HPC via the thalamus, facilitating the update of stored representations and top-down control of memory retrieval (Dolleman-van Der Weel et al., 2019; Eichenbaum et al., 2007).

Besides, the hippocampal-amygdala connection is viewed as a means for emotional and contextual information integration (Phelps, 2004) - the basolateral amygdala (BLA) projects to the subiculum, CA1 and CA3 regions of the HPC, and, in return, it projects back to the BLA (Pitkänen et al., 2000). This interconnection is particularly significant in the context of fear conditioning and consolidation of emotionally salient memories (Richardson et al., 2004; Ritchey et al., 2008).

Allegedly important for spatial memory and navigation are the projections from CA1 and subiculum to the anterior thalamic nuclei via the mammillary bodies, forming the hippocampal-thalamic-mammillary system (Aggleton et al., 2010; Jankowski et al., 2013). Moreover, the medium septum, also a subcortical structure, communicates to the HPC via cholinergic and GABAergic projections, modulating its activity and contributing to the generation of theta oscillations (Vertes and Kocsis, 1997). Finally, as mentioned previously, the EC is the principal input to the HPC and the main intermediary (together with other parahippocampal structures) between this structure and various cortical areas responsible for sending sensory,

motor and associative inputs to the HPC (Suzuki and Amaral, 1994; Lavenex and Amaral, 2000).

1.2.3 Representations in hippocampus

Earlier lesion studies in the HPC have helped establish the importance of this region in the performance of flexible behaviour. However, it was not until the advent of intracortical recordings that we could begin to elucidate the underlying mechanisms behind it.

1.2.3.1 Place cells

The first neurophysiological evidence for map-like representations in the rat's brain emerged roughly thirty years after Tolman's famous claims. In 1971, O'Keefe and his student Dostrovsky implanted rats with microelectrodes in dorsal HPC and recorded single-neuron activity while manually moving the animals around a rectangular arena. To their surprise, the activity of some of the recorded neurons was modulated by specific locations in the environment - each cell fired only when the animal was at a particular point in the arena (O'Keefe and Dostrovsky, 1971). A population of these place cells in HPC (O'Keefe, 1976) would tile the entire space to allow navigation, thus providing the substrate of the cognitive map structure in the brain (O'Keefe and Nadel, 1978).

Each place cell fired in a unique location - its place field (O'Keefe, 1979) - with different cells covering the entire explored space with different fields (Wilson and McNaughton, 1993); and the shape and size of the fields were modulated by the

size and shape of the environment. While normally place cells exhibit elongated firing fields along the animal's direction of movement (Muller et al., 1994; Mehta et al., 1997); these fields stretch and become more elliptical in larger environments of the same type (O'Keefe and Burgess, 1996), and place cells can even change to a multifield representation system as reported by Fenton and colleagues (Fenton et al., 2008; Park et al., 2011). Furthermore, field locations can change when the shape of the environment changes (Muller and Kubie, 1987), reflecting the dynamic nature of hippocampal representations in a phenomenon called spatial remapping (discussed in more detail in Section 1.2.3.2).

More often than not, navigating an environment is done with a goal in mind, for example finding food or other types of rewards. Identified in the HPC (Hölscher, 2003; Sarel and Finkelstein, 2017; Gauthier and Tank, 2018) (but also in EC (Gauthier and Tank, 2018) and PFC (Passecker et al., 2019)), reward cells, or goal cells, can encode the spatial location of such rewards, the animal's proximity to them, and even their relative value. Hence, in coordination with the many cells already mentioned, goal cells are thought to provide crucial information for efficient goal-directed navigation and decision-making (Zhu et al., 2021).

Although extensively studied in rats, hippocampal place cells have been identified in multiple other species - mice (Kentros et al., 2004), monkeys (Rolls, 1999), humans (Ekstrom et al., 2003) and even bats (Ulanovsky and Moss, 2007). Place cell representations in bats exhibit spatially selective fields like those in other mammals, but with significantly larger and three-dimensional fields (Ulanovsky and Moss, 2007; Sarel and Finkelstein, 2017; Omer et al., 2018). They are strongly

modulated by echolocation behaviour (Ulanovsky and Moss, 2011), and indeed bats can create spatial representations based on echolocation cues alone (Geva-Sagiv et al., 2015). The similarities and differences in place cell representations between the bat and other mammals highlight the conserved spatial role of the HPC across species and its wonderful adaptability to the navigational challenges that each species faces.

Perhaps the most surprising features of hippocampal place cells is their capacity to reactivate previously exhibited firing patterns in a totally different context. These sweep events happen either during theta cycles (Buzsáki, 2002) - sequence events -, or sharp-wave ripples (SWRs) (Pavlides and Winson, 1989; Wilson and McNaughton, 1993) - replay events. Having been hypothesised as a key mechanism in the creation and consolidation of cognitive maps with great contribution to planning and decision-making behaviours, we explore these topics in more detail in Section 1.2.3.3 and Section 1.2.3.4, respectively.

'Place cell' non-spatial representations

However, cells in the HPC can also be involved in the representation of non-spatial features of the world. An example of this has been found in rats performing a 1D non-spatial operant task - rats were required to manipulate a joystick to navigate a sound frequency space and release at the rewarded frequency (Aronov et al., 2017). Place-like cells in the HPC were active at specific frequencies within the frequency space, similar to the representation of location in physical space. Further evidence of these abstract representations has been observed in human and non-

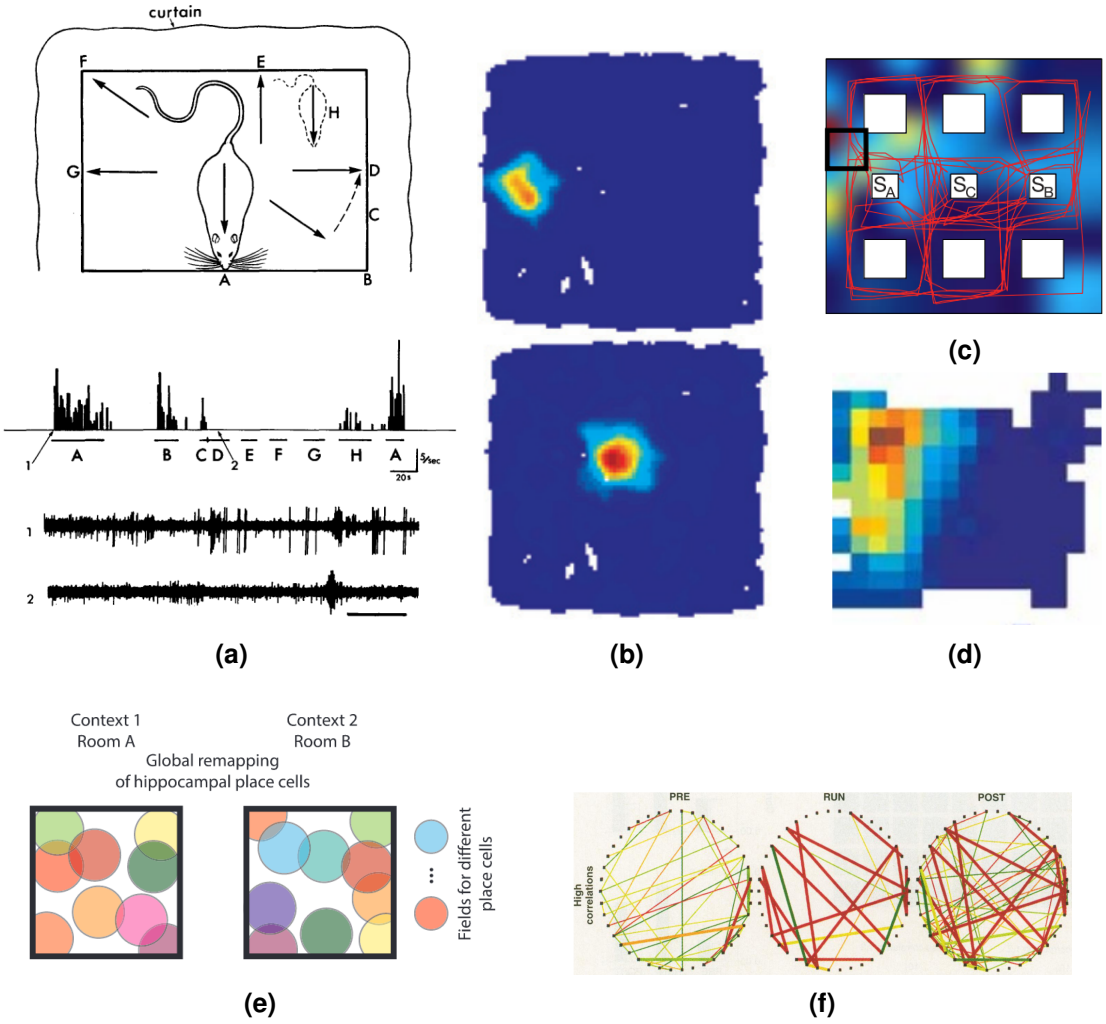


Figure 1.7: Hippocampal place cells and their features. (a) first recorded place cell in rats that fires for locations A and B only (adapted from O'Keefe and Dostrovsky (1971); (b) firing fields from two example place cells as they are often plotted in recent work (adapted from Hafting et al. (2005); (c) fields from human place cell (adapted from Ekstrom et al. (2003); (d) place cell field from a bat (adapted from Ulanovsky and Moss (2007); (e) schematic of place cell remapping phenomena (adapted from Behrens et al. (2018); (f) first representations of remote place cell reactivation (replay) during sleep (adapted from Wilson and McNaughton (1994).

human primates. Hippocampal place cells that would fire at specific locations in physical space fire in association with particular stimuli within the space of available stimuli, as is the case of cells that prefer certain faces or monuments (Quiroga et al., 2005), or construct a map of subjective value space orthogonal to other

sensory information spaces in a decision-making task (Knudsen and Wallis, 2021), thus highlighting a role of HPC beyond navigation and episodic memory.

Another particular case of abstract representation of 1D spaces that can be found in hippocampal cells is time. Accomplished by place-like cells that track time - time cells -, the encoding of the temporal dimension is complementary to the encoding of spatial dimensions in the formation of episodic memory (Eichenbaum and Cohen, 2014; Pastalkova et al., 2008). Hence, when time is a relevant feature of task space, while space is tiled by place cells, other cells in the HPC will independently tile the time interval for task completion (Kraus et al., 2013). These are also present in non-spatial tasks where they have been shown to tile the delay period in an odour memory task (MacDonald et al., 2013). These time-dependent representations have also been studied in humans (Umbach et al., 2020). Hippocampal neurons are also known for their ability to encode states (Schuck et al., 2016), relations (Deuker et al., 2016), concepts (Theves et al., 2019) and task-specific events. In a task requiring animals to move around a circular track multiple times before receiving a reward, Sun et al. (2020) found mouse CA1 cell activity to be modulated not just by spatial location on the track but also by lap number. This lap-specific firing was independent of the lap length and generalised across different spatial structures, reinforcing the idea that hippocampal cells code for an abstract organisation of experience that is discrete and can, therefore, be applied to different contexts.

The HPC has also been implicated in the representation of social relationships and hierarchies in humans and other animals. Tavares et al. (2015) observed

neural responses modulated by social rank in humans, with distinct neural patterns being associated with different levels of social hierarchy and the representation of social distance being encoded in HPC. Analogously, the same has been found in bats - studies have shown that bats' hippocampal place cells can form individual representations for other bats independent of their location (Omer et al., 2018), and that this same representation is modulated by the distance to the other bat and the social relationship between them, with stronger activity for closer distances and more closely affiliated bats (Eliav et al., 2021). Such representations support the navigation of a social space rather than a physical one.

1.2.3.2 Remapping

Spatial remapping of place maps - i.e. changes of place cell firing patterns - happens when an animal changes to a differently shaped environment but not when transitioning between two of the same environments (Lever et al., 2002). The extent of remapping can vary with how different the spaces are and the animal's recent experience in the space (Quirk et al., 1990), and it can happen at a slower or faster pace (Leutgeb et al., 2005; Frank et al., 2004). The remapping of spatial representations can be partial or global if only a subset of the firing fields has changed and another remains unaltered or if the resulting firing map is considered orthogonal to the previous one (Cai et al., 2016; Keinath et al., 2022; Mankin et al., 2012; Ziv et al., 2013; Sanders et al., 2020). Usually, partial remapping is associated with smaller changes in the environment - altering a portion of the entire space or introducing new elements in it (e.g. barriers) -, thus maintaining a stable repre-

sentation of space while allowing the update in the representation of the modified areas (Skaggs and McNaughton, 1998; Bostock et al., 1991). Global remapping, through changes in the active cell population, the appearance of new fields or reorganisation of pre-existing ones; is interpreted as a response to significant changes in the environment, to a new space or to aversive stimuli, reflecting the formation of a completely independent map (Dupret et al., 2010; Moita et al., 2004; Ormond et al., 2023). Alternatively, the concept of rate remapping, which refers to a change in the firing rates of hippocampal place cells without considerable differences in their spatial allocation, is normally associated with small environmental changes, such as colour or texture, allowing for the discrimination of different contexts or experiences within the same map (Leutgeb et al., 2005).

These remapping phenomena elicit an important characteristic of hippocampal place cells. Contrary to the initial beliefs about place cells (O'Keefe, 1976), these do not encode a simple representation of space. Instead, they are a conjunctive code of spatial and non-spatial information (sensory cues, task-relevant features, context) (Wiener et al., 1989; Markus et al., 1995; Eichenbaum, 1999; Anderson et al., 2000; Nieh et al., 2021).

1.2.3.3 Sequences

Another type of hippocampal representation can be done in the form of sequences. These temporally compressed groups of place cell firing can be detected in the so-called theta cycles. Each theta cycle is a single period of the hippocampal theta rhythm, a sinusoidal oscillation in the local field potential (LFP) of this region

(Dragoi and Buzsáki, 2006; Foster, 2017) with a frequency of approximately 4-10 Hz in rodents and 1-4 Hz in humans (Jacobs, 2014). Theta sequences, or sweeps, originate primarily in the CA1 region of the HPC, but have also been detected in the CA3 and the DG (Buzsáki, 2002).

During spatial navigation, place cells in the HPC fire in sequences that correspond to the animal's trajectory through space. Simultaneously, repeated theta cycles can generate a series of forward sequences of place representations that sometimes alternate between multiple potential future paths to goal (Johnson and Redish, 2007; Kay et al., 2020). Furthermore, these sequences have been shown to coincide with goal representation in the PFC (Zielinski, 2020), and are often associated with the presence of VTE behaviour at choice points in rodents (Redish, 2016; Tolman, 1948a), which is locked to the hippocampal theta oscillation (Johnson and Redish, 2007). For these reasons, sequential hippocampal firing has been suggested to mirror processes of mental simulation of future behavioural solutions, and of being a supporting mechanism of online planning (Redish, 2016; Mattar and Daw, 2018).

Considering the expression of VTE, this occurs when animals are faced with difficult decisions within a world for which they have an internal model. As previously hypothesised by Tolman (1948a), this behaviour goes away in cases of excessive training in stable environments (Van Der Meer et al., 2020; Gardner et al., 2013), and reappears if task contingencies are changed (Schmidt et al., 2013), matching the normative conditions for MF and MB RL, respectively.

However, theta sequences have also been implicated in the consolidation and

retrieval of spatial information (Burgess et al., 2002, 1994). Borrowing from the theta phase precession phenomenon, according to which place cells fire in progressively earlier phases of the theta cycle as the animal moves through their place fields (O'Keefe and Burgess, 2005), the compression of spatial sequences within each theta cycle will thus enable a rapid association between locations and related events, facilitating the encoding of episodic memory. Models of place and grid cells (Burgess et al., 1994; Bush et al., 2015) have suggested that theta sweeps arise from an interaction between hippocampal place cells and inhibitory neurons, along with the hexagonal grid input from the EC. Furthermore, theta phase precession can undergo remapping itself, which refers to theta precession pattern alterations caused by changes in the environment or the animal's experience, and that result in temporal remapping or place cell activity (Skaggs et al., 1996; Ekstrom et al., 2001). This type of remapping is thought to play a role in the temporal coding of spatial information and the formation of memory sequences.

1.2.3.4 Replay

The discovery

Following the report of place cells as cells whose firing could not occur outside of their place field, a study reported that place cells that were active in an environment would be more active in subsequent sleep than the ones that had not been active in the environment (Pavlides and Winson, 1989). This supported the idea that place cells could reactivate away from the initial stimulus location, which was

soon after confirmed by Wilson and McNaughton (1994) in what is established as the first paper reporting hippocampal replay. They had rats run a circular track for food rewards while performing multi-unit recordings in the HPC and found that pairs of cells with overlapping place fields during awake behaviour spontaneously reactivated during subsequent sleep periods. These reactivation patterns replicated the firing patterns observed during exploration in a highly temporally compressed manner, typically lasting between 30 and 300 ms, and primarily occurred during SWR events (Buzsáki, 1989; Nádasdy et al., 1999). With several other studies supporting these findings (Kudrimoti et al., 1999; Nádasdy et al., 1999; Skaggs et al., 1996; Lee and Wilson, 2002; Louie and Wilson, 2001), replay was initially seen as a recapitulation of past experience and hypothesised to play a role in memory consolidation and information transfer to the neocortex.

Types and functions

Since then, much more has been discovered about these replay events, but very little has been consensually agreed upon amongst the scientific community as to what the real function of this remarkable mechanism in our brains stands for.

One fundamental distinction in the study of replay is between offline and online replay.

Soon after discovering place cell self-reactivation during sleep - or offline replay, which extends to rest periods without behavioural engagement -, scientists wondered about the hippocampal ability to represent these prior experiences in an awake behavioural context (Kudrimoti et al., 1999). In fact, years later, several

researchers reported the occurrence of hippocampal replay during ongoing behaviour in rats (Foster and Wilson, 2006; Csicsvari et al., 2007; Diba and Buzsáki, 2007; Karlsson and Frank, 2009; Davidson et al., 2009) - online replay. Although this type of replay shares some of the main features with offline replay - both mostly happen during SWRs and are temporally compressed in a similar timescale -, online replay has its unique characteristics. Observed during periods of awake immobility, reactivation of place cell activity was recorded in the rat's brain within seconds after experience and even after the first lap in the environment (Foster and Wilson, 2006). Moreover, awake replay did not simply recapitulate sequences of hippocampal activity in the same order as experienced before, as mainly reported for sleep replay (Lee and Wilson, 2002) - known as forward replay -, but often replayed prior experience in the reversed order - to which we call reverse replay (Csicsvari et al., 2007; Diba and Buzsáki, 2007; Karlsson and Frank, 2009; Davidson et al., 2009; Gupta et al., 2010).

Furthermore, studies have shown activation of place cells during ripple events representing sequences of future trajectories, or forward sweeps (Pfeiffer and Foster, 2013; Widloski and Foster, 2022), a specific subtype of replay called preplay.

Replay was initially believed to play a crucial role in memory consolidation, supported by the self-reactivation during sleep of place cells that fired together in previous awake experience (Wilson and McNaughton, 1994) in the same order observed during preceding runs (Lee and Wilson, 2002). Additionally, disruption of rats' SWRs and of, consequently, replay events during rest/sleep periods both between learning trials (Ego-Stengel and Wilson, 2010) and post-learning (Girardeau

et al., 2009) of an eight-arm radial maze impaired the animals' performance based on their ability to remember the location of the three baited arms, whereas the disruption of SWRs during task performance did not impact their behaviour. These findings align with the pre-existing idea that offline SWR replay is essential for strengthening and stabilising newly formed memories, facilitating the transfer of information from the HPC to the neocortex and supporting its integration into long-term memory.

The potential role of online replay diverges depending on the direction of sequence reactivation. In the case of forward replay, it is viewed as a mechanism for consolidation of recent experience and reinstatement of existing neural representations (Foster and Wilson, 2006; Diba and Buzsáki, 2007; Davidson et al., 2009). This is corroborated by the increase in replay at choice points during VTE behaviour of rats, and subsequent performance impairment when disrupting hippocampal activity at those exact times (Papale et al., 2016).

Alternatively, reverse replay has been proposed as a substrate for evaluating past experiences (Foster and Wilson, 2006; Diba and Buzsáki, 2007) as it can recapitulate, albeit in the opposite direction, the most recent experience. In addition, reverse replay has been shown to be modulated by reward magnitude, increasing in frequency and extending in content when rewards were present compared to when they were not, and having a direct correlation with reward size (Ambrose et al., 2016). This suggests a role of replay in credit assignment, prioritising updates of more valuable experiences. We will later discuss potential theoretical models for this fascinating hippocampal phenomenon Section 1.3.3.

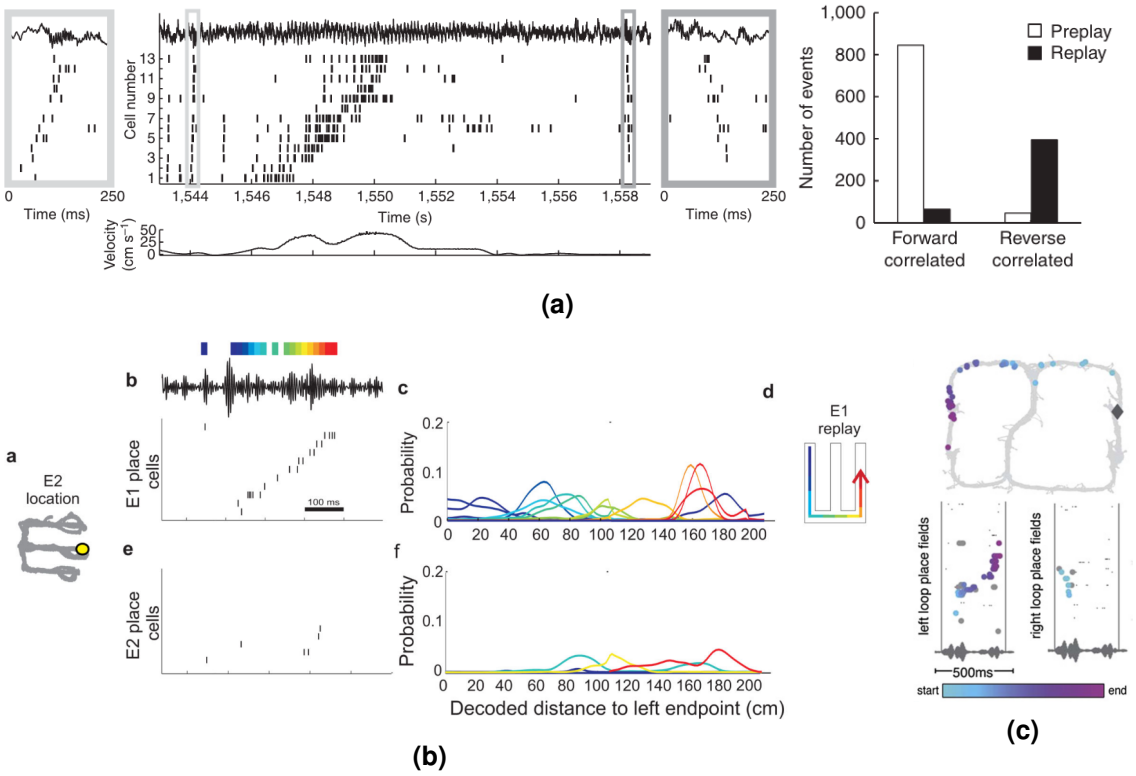


Figure 1.8: Hippocampal replay. (a) identification of forward preplay and reverse replay events out of CA1 neural spike trains before, during, and after behaviour (left) and correlation of forward events with periods before the run and of reverse events with periods after the run (right) (adapted from Diba and Buzsáki (2007)); (b) Online replay of a previous environment E1 when the animal is stopped in environment E2 (adapted from Karlsson and Frank (2009)); (c) example of novel trajectories constructed in replay (adapted from Gupta et al. (2010)).

Replay beyond the hippocampus

Over the past 30 years, hippocampal replay has been widely studied across species, both because it is thought to be a fundamental piece in understanding the brains' cognitive function, but equally for the lack of consensus about its role in the most various forms and even about what should be considered replay in the first place. Nonetheless, this fascinating mechanism is not unique to the HPC, and it has been found in multiple other regions across the brain, inside and outside of the hippocampal formation.

One of the most interesting brain regions to observe together with the HPC in the study of cognitive maps is the EC. Due to its input-output loop into the HPC (Section 1.2.2), it is perhaps not entirely surprising that researchers have recorded grid cell replay in this region. Although evidence suggests that, during sleep, the medial entorhinal cortex (MEC) can generate replay activity independently of the HPC (O'Neill et al., 2017), its capacity to represent recent spatial experience extends to online contexts, where it exhibits strong temporal coordination with hippocampal replay events, and where its spatial coherence is thought to depend on the hippocampal signal (Ólafsdóttir et al., 2016). From these, we can extrapolate MEC replay to contribute to integrating and consolidating new spatial memories, which is supported by models where the HPC and EC work together in the construction and maintenance of spatial representations (Kubie and Fenton, 2012; Erdem and Hasselmo, 2014; Bush et al., 2015), and where MEC replay is pointed to provide a stable spatial framework for the reactivation and consolidation of hippocampal memories.

The belief that replay is a substrate for spatial memory consolidation expands beyond the hippocampal formation all the way to the prefrontal cortex. Observed offline in the rat mPFC, the reactivation of neural activity patterns experienced during a preceding spatial task (Euston et al., 2007) shows evidence for the memory consolidation hypothesis. Additionally, this role in facilitating information consolidation applies to more abstract rules and strategies other than spatial (although still in a spatial context) - e.g. rats in a Y-maze learned various rules that signalled the rewarded arm at different times, and the neural patterns reflecting previously

acquired information were replayed in mPFC during sleep and in close synchronisation to hippocampal SWRs (Peyrache et al., 2009). More recently, Jadhav et al. (2016) performed simultaneous recordings of HPC and mPFC activity in rats during a spatial alternation task and observed coordinated replay activity in both regions during SWR events that took place when the animals stopped at the rewarded locations or choice points. These reactivations mainly reflected the preceding trajectories, and informed the temporal coordination between mPFC replays and SWRs in the HPC (lags of typically $\pm$100ms). With the replay coordination being stronger at early learning stages (Jadhav et al., 2016), mPFC replay is considered more relevant for integrating new information transmitted by the HPC, but it could also be involved in evaluating past experiences and planning future actions due to its presence at choice point locations.

Coordinated reactivation of recently experienced neural patterns can also be found between HPC and the ventral striatum (VS). The temporal correlation between the two signals seems to be more robust during periods of slow-wave sleep (SWS) when compared to rapid-eye movement (REM) sleep or quiet wakefulness, and the replay of neural ensembles in VS after a reward-seeking behaviour could be an auxiliary mechanism for reward-related memory consolidation (Pennartz et al., 2004). This study was later deepened by investigating the impact of reward magnitude on offline VS replay (Lansink et al., 2008) in rats. Again, reactivation of neural patterns in the VS was observed during SWS, and it was proved to be modulated by reward size in the different locations - neurons associated with high-reward locations showed more prominent reactivation compared to those associated with

low-reward locations. Complementarily, coordination between VS and HPC replays was higher for high-rewarded locations than low-rewarded ones, corroborating the idea that both regions interact to process and consolidate reward-related information, which in turn can be essential for reinforcing specific actions over others and, thus, guiding future behaviour.

1.2.3.5 Other relevant representations

Grid cells

Another fundamental piece of the cognitive map machinery is the entorhinal grid code, or grid cells. In 2005, Hafting et al. (2005) recorded the activity of MEC grid cells in freely moving rats in open arenas. These cells fired in multiple locations along the environment with fields, organised in a periodic triangular grid that covered the entire space. The spacing and orientation of the grid pattern vary from cell to cell with an offset from each other in a way that the firing fields of neighbouring cells are shifted in space, thus providing a universal spatial coordinate system as a metric for path integration (Moser et al., 2008). This grid map representation is thought to represent vector relationships and distances between different spatial locations (Bush et al., 2015; Stemmler et al., 2015), exhibiting different levels of coarseness in its representation, with the distance between the firing fields increasing from the dorsal to the ventral MEC (Hafting et al., 2005).

More recently, studies have emphasised the development and plasticity of grid code. Emerging gradually during post-natal development, grid patterns stabilised

around 3-4 weeks in rats (Langston et al., 2010; Wills et al., 2010). Throughout adulthood, these patterns are flexible to the specifics of the environment, with the grid scale and orientation being modulated by the size and shape of the space (Barry et al., 2007, 2012; Stensola et al., 2012; Stemmler et al., 2015) and/or being modulated by the presence of boundaries or landmarks in the environment (Krupic et al., 2016; Hardcastle et al., 2017). This plasticity provides evidence against a solely intrinsic origin of grid cells (Fyhn et al., 2007) and indicates an intricate cooperation between self-motion and external sensory information in their shaping.

Grid-like activity can be recorded in humans while they navigate a VR environment, either directly from the activity of cells in the EC (recorded for medical reasons) (Jacobs et al., 2013), or in VR using a functional magnetic resonance imaging (fMRI) proxy (Doeller et al., 2010).

Besides, activity in the MEC can analogously represent abstract spaces, exhibiting multiple peaks and coding for sound frequencies in 1D sound space for rats (Aronov et al., 2017); and as demonstrated in an elegantly designed ”bird-space” task where humans had to observe birds morphing in a 2D continuous space and predict when their appearance matched the previously learned bird (Constantinescu et al., 2016).

The interaction between grid cells and other types of spatially tuned entorhinal cells - head cells, border cells, speed cells (Sargolini et al., 2006) - has also been of great interest.

Head-direction cells

Initially reported in the PrS of the rat (Ranck, 1984), head direction cells have since been found in the post-subiculum, PaS, thalamus and EC as well (Taube et al., 1990; Taube, 1995; Knierim et al., 1995). Individually, these cells maximally fire when the animal's head points in a particular direction, independently of the animal's location in the environment, with a Gaussian-like gradual decrease in activity as the head moves away from the preferred direction (Taube et al., 1990; Taube, 1995). At the population level, head direction cells cover the entire range of possible orientations within a given space (Johnson et al., 2005). Although head direction cells are primarily driven by vestibular and proprioceptive cues - they maintain their preferred direction even in the absence of visual stimuli, i.e. in the dark (Taube et al., 1990) -, they can be modulated by visual cues (Taube et al., 1990; Zugaro et al., 2001).

Boundary vector cells

Boundary vector cells, or border cells, fire preferentially when the animal is close to borders or boundaries within the environment. Initially found in the EC, they exhibit elongated fields along their preferred boundaries and increase their firing as the animal gets closer to them (Solstad et al., 2008). Border cells can also be found in the subiculum, and even though they share characteristics with the entorhinal border cells, they show a specific conjunctive coding of boundary proximity and head direction that might help the animal positioning relative to the boundaries

(Lever et al., 2009). These cells provide a spatial anchor to the cognitive map, and researchers believe they contribute to the formation of place and grid fields and the estimation of distance to boundaries (Hartley et al., 2000; Barry et al., 2006; Burgess, 2008).

Object vector cells

Although also present in the HPC, object vector cell firing fields have a much more precise spatial tuning in the EC. These neurons fire when the animal is at a particular distance and in a specific direction from a given object, irrespective of its location in the environment (Deshmukh and Knierim, 2011; Høydal et al., 2019). Similarly to the boundary vector cells, object vector cells are thought to act as an anchor in the cognitive map that points to relevant features in the environment, and it does so by interacting with the boundary vector cells, but also with place and grid cells.

Computational models have been proposed to explain their function and suggest they arise from integrating object- and space-related information within the hippocampal-entorhinal circuit (Byrne et al., 2007; Bicanski and Burgess, 2018).

Speed cells

For the first time found in the HPC of rats in the 1980s (McNaughton et al., 1983), speed cells, or velocity cells, can also be observed in the MEC and are generally known for linearly increasing their firing rate with the increase in the animal's running speed (Kropff et al., 2015; Hinman et al., 2016). Additionally, some velocity

cells are conjunctively modulated by the animal moving direction (Kubie and Fenton, 2012). Speed cells have been proposed to support navigation by interacting with grid and place cells to update their spatial representation (Hasselmo, 2008; Burak and Fiete, 2009; Hasselmo and Stern, 2014).

These cells have also recently been identified in the human EC and HPC to help with navigation similarly to other animals (Qasim et al., 2019; Stangl et al., 2021).

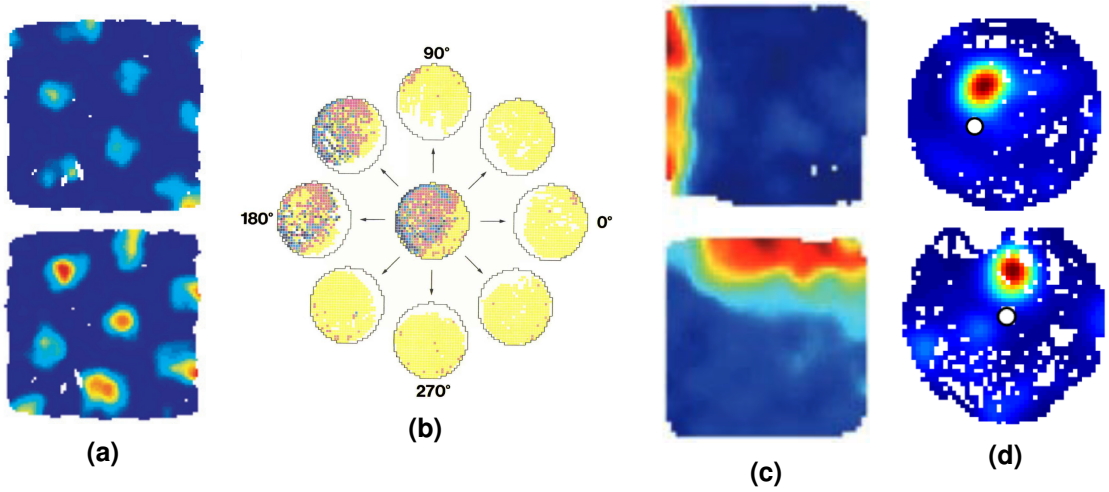


Figure 1.9: Spatial representations outside of the hippocampus. (a) two example grid cells with firing fields repeating in a triangular lattice (adapted from Hafting et al. (2005); (b) example head-direction cell with specific firing between 90-180 degrees (adapted from Taube et al. (1990)); (c) two example boundary vector cells with firing fields along the outer edges of the arena (adapted from Solstad et al. (2008)); (d) two example object vector cells whose fields fire at a fixed distance and angle from an object (adapted from Høydal et al. (2019)).

Section 1.3

Theories and models of hippocampal flexibility

Over the past few decades, neuroscience has made significant progress in elucidating the neural representations of space in the brain (see Section 1.2). However, while we now have precise knowledge of how individual neurons represent aspects of the spatial worlds (albeit in simple environments!), there is still a lack of formal understanding of how they are combined to control behaviour and what computations underlie these processes. To bridge this gap, many researchers have developed computational models that aim to explain how the brain integrates information to guide different kinds of flexible behaviour.

Here, we will focus on theories that try to understand the algorithms and computations underlying MB navigation, i.e. planning. Drawing from various fields - e.g. RL, optimal control theory, graph theory - we will start by exploring some of the most influential computational models of goal-directed navigation, such as the Dyna-Q, path integration and successor representation. Finally, we will dive into the realm of replay models in search of a better understanding of their various types, functions and impact on the creation and update of cognitive maps.

1.3.1 Planning in space

Often, models of decision-making in spatial environments borrow from the RL framework. In fact, Sutton and Barto (1998) dedicates an entire chapter of their book to discussing algorithms of learning and planning. And, before we dive into the realms of planning algorithms, I will first briefly introduce the concept of TD learning, as it is the learning method used by some of them.

Temporal Difference (TD) learning is an RL approach to learning the prediction of quantities that depend on the future values of a signal. It performs credit assignment to actions by comparing their expected and observed long-term value and by updating these expectations through RPE. It thus allows for online learning, and is particularly relevant in neuroscience due to its correspondence with midbrain dopamine signalling (Schultz et al., 1997). One downside of this approach is that it requires many experiences of actions to generate many of such updates, and so to perform long-term credit assignments. And this is where the Dyna algorithms can be of help, by alleviating the experience requirements with their feature of offline value updates (Sutton, 1991).

Dyna-Q algorithm (and variations)

One of their most influential algorithms for navigation planning is the Dyna-Q. Combining MF and MB learning, the Dyna-Q agent learns a model of the environment through experience, which it can then use to simulate the outcomes of different actions and update its value estimates offline, without actually performing those

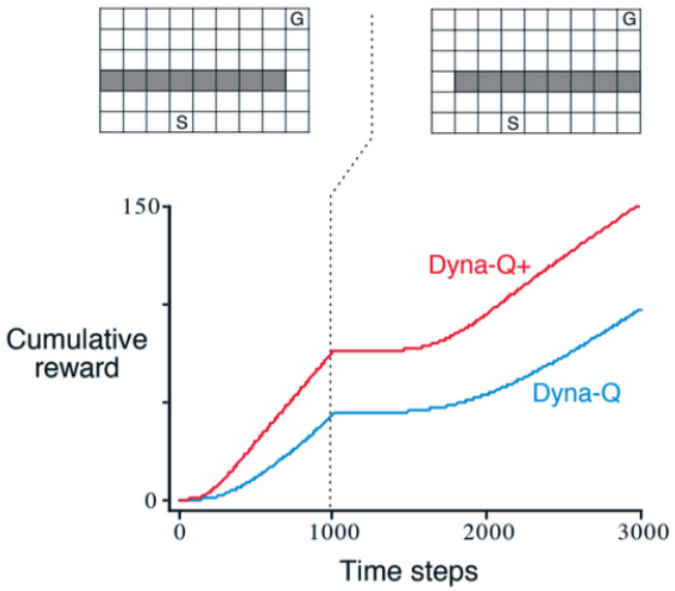


Figure 1.10: RL Dyna-Q planning. Dyna-Q+ outperforms Dyna-Q when a change is introduced to the environment (adapted from Sutton and Barto (1998)).

actions in the real world (Sutton, 1991). This allows the agent to learn faster and adapt more quickly to changes in the environment compared to purely MF methods. Additionally, a small exploration/exploitation heuristic can be added to the Dyna-Q, which is then called Dyna-Q+. In this version of the algorithm, the agent will keep track of the elapsed time since its last visit to each state-action pair and is encouraged to revisit the locations it has not visited for longer sporadically. This modification makes the model more successful in the face of dynamic environments, as the exploration incentive prevents the agent from getting stuck in previously learned trajectories (Figure 1.10).

The Dyna-Q architecture has successfully been employed in the modelling of navigation and of the two-step task (Gershman et al., 2014).

Later, prioritised sweeping was introduced as an extension of the Dyna-Q algorithm (Moore and Atkeson, 1993). The key idea behind it is to prioritise the

updates of state-action values based on the magnitude of their expected changes. So, the agent keeps an ordered queue of every state-action pair whose value would change if updated with respect to the size of the change; and sweeps through this list backwards from the goal or other updated states, or forward if focusing on states according to how easily they can be reached from the most frequently visited states under the current policy. In the case of stochastic environments, this agent should update each state-action pair with an expected update instead, or sample updates for a less costly solution (Sutton and Barto, 1998).

Real-time dynamic programming

Contrarily to the explored Dyna-Q algorithms, a real-time dynamic programming (RTDP) agent only updates the value of relevant state-action pairs - pairs that are more likely to be visited by the agent given the current policy -, and performs these updates while interacting with the environment, in real-time (Barto and Dayan, 1995). Respectively, these characteristics render it faster to converge and quicker to adapt to changes in the environment than the traditional Dyna. Yet, it can require more extensive exploration to assure convergence to optimal policy and, as it relies on an internal model to compute the expected values of the next states, it may update the value function suboptimally. Despite its limitations, The RTDP algorithm has been demonstrated effective in several domains, including robot navigation (Koenig and Simmons, 1996) and autonomous driving (Shalev-Shwartz and Wexler, 2016).

Rollout algorithm

A simple and yet very effective algorithm in sequential decision-making problems (e.g. navigation) is the rollout (Sutton and Barto, 1998). The idea behind it is to simulate sequences of decisions, like trajectories - i.e. the rollouts - starting from the current state until a terminal state, or until a maximum number of available steps, and following a given policy. The simulated trajectories are then used to estimate the expected cumulative reward of each possible action, and the action with the highest estimated value is selected as the best decision. While it can function with any model and policy, this rollout mechanism can be highly computationally costly in large or complex environments. Variations of this algorithm, such as the Monte Carlo tree search (MCTS) algorithm, combine a rollout and a tree search algorithms to selectively expand promising regions of the state space, thus reducing the size of the effective search space (Sutton and Barto, 1998).

Besides the navigation planning algorithms discussed above, a range of theoretical decision-making models has been proposed to explain the behavioural and neural underpinnings of goal-directed navigation.

Path integration

According to the path integration (PI) model of navigation, animals maintain a continuous estimate of their location relative to a starting point by integrating self-motion cues over time (McNaughton et al., 2006). Unlike landmark-based navigation (Epstein and Vass, 2014), this model allows the agent to calculate routes with-

out the use of landmarks. Instead, the agent continuously updates its estimated position relative to a reference point, by integrating self-motion cues derived from its movement through the environment. The desert ant is an excellent example of this form of navigation, as it can return directly to the nest after being hundreds of meters away from it in a featureless environment (Wehner and Srinivasan, 1981).

This and related models have been used to explain the firing properties of place and grid cells. For instance, in the boundary vector cell model (Barry et al., 2006; Hartley et al., 2000), hippocampal place cells receive input from cells whose firing activity is modulated by the presence and distance from boundaries in the environment (boundary vector cells), which allows them to represent the animal's position in relation to the boundaries around it and thus explains the stable firing patterns place cells exhibit even in the absence of landmarks (O'Keefe and Burgess, 1996). Other models have demonstrated how grid cell regular firing patterns can be explained by simple integration of self-motion cues (Burgess, 2008; Hasselmo, 2008). The authors propose entorhinal grid cells receive input from speed-modulated head-direction cells, which provide information about the animal's direction and velocity of movement. Subsequently, integrating this input over time allows grid cells to maintain a representation of the animal's position, updated on the base of self-motion cues.

Recent studies have provided further evidence for the role of path integration in navigation. Gil et al. (2018) showed that mice can navigate back to a home location in a VR setting in the absence of visual landmarks. Additionally, they found that the mice's navigation performance is correlated with the stability of grid cell repre-

sentations in the EC, suggesting a direct link between path integration and spatial representation. Similarly, Chen et al. (2019) demonstrated that path integration is essential for the formation and maintenance of place cell representations in the HPC, using a VR system that allowed for the dissociation of self-motion cues from visual landmarks.

Euclidean vs Geodesic navigation

The most straightforward strategy for navigation would be to travel along the shortest vector path between the starting and the goal points. Exploring this idea, Dordek et al. (2016) proposed a computational model of Euclidean navigation (or vector navigation), which plans routes by calculating the distances between locations as a combination of Euclidean distance and angular information. Although very efficient in open environments, this model quickly fails to provide the optimal answer (or a possible one) when a barrier is introduced and cannot reproduce the firing activity observed in more structured spaces (Figure 1.11(a)).

A more helpful strategy for navigating complex environments would perhaps follow a geodesic approach, which translates into travelling along the shortest possible route towards the goal location. Modelled by Gustafson and Daw (2011), this strategy incorporates knowledge about the environment's topology through the learning of a graph-like structure that encodes the presence of obstacles and the connectivity between different locations. This model was shown to account for the navigation behaviour and firing properties of hippocampal place cells of animals in complex environments (Figure 1.11(b)).

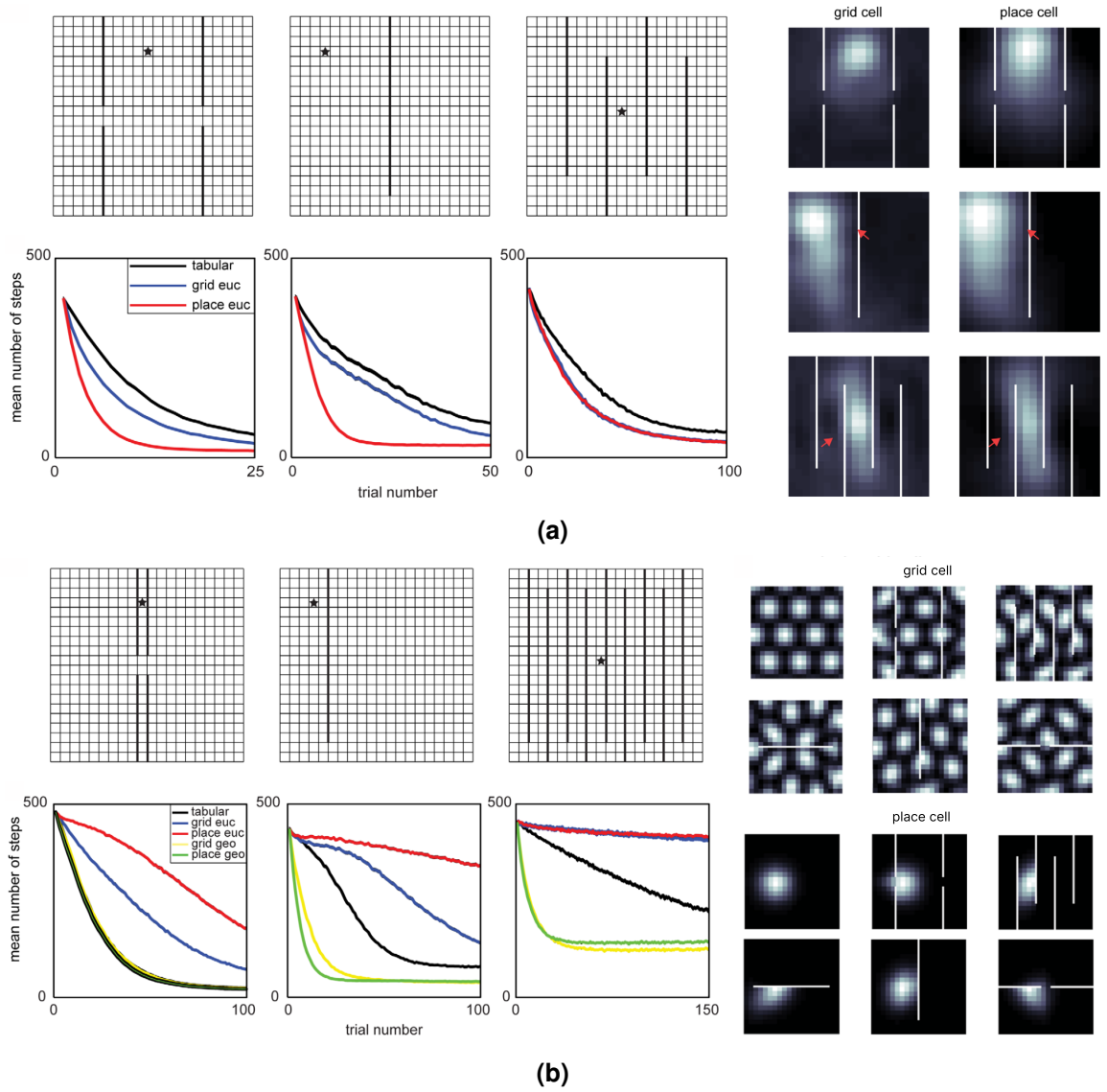


Figure 1.11: Euclidean vs geodesic navigation models. (a) modelling of a Euclidean strategy in a grid world navigation task shows benefits of spatial generalisation (left) and comparison of learned value functions using Euclidean place cells after 15, 25 and 50 learning trials (right); (b) model showing the value of geodesic representation for learning in a grid world with value function discontinuities in Euclidean space (left) and example geodesic place cell transformations in the presence of walls in the grid world (adapted from Gustafson and Daw (2011))

Employed in a RL context, Euclidean and Geodesic strategies of navigation have also been assessed in human spatial cognition. In a VR world, humans exhibited a mixture of these two strategies when navigating through a structured

space with various obstacles, mainly relying on Euclidean navigation in simpler spatial configurations, and the goal was visible and within easy reach; but switching to a primarily geodesic strategy when the goal was obscured, and a detour was necessary to reach it (Warren et al., 2017).

1.3.2 Successor representations

The successor representation (SR) was first introduced by Dayan (1993) as a framework capable of extending the concept of state adjacency from space to time. While in static tasks, generalisation is commonly achieved by conferring similar representations to states spatially near each other; the SR proposes this adjacency be defined in terms of time proximity over some behavioural axis for dynamic systems, which naturally require predictions to be made over time.

An SR agent learns a predictive representation of the environment that encodes the expected occupancy of each current state (s) (instead of the expected reward), given the current policy (Π). Credit assignment can thus be decomposed into two components - the expected future occupancy of each state (s') and the immediate reward associated with it ($R(s)$). This simple multiplication allows quick generalisation across reward locations within the same environment.

$$M(s, s') = \mathbb{E}\left[\sum_{t=0} \gamma^t \mathbb{I}(s_t = s') | s_0 = s\right] \tag{1.1}$$

where γ is the discount factor, $I(s_t = s')$ is an indicator function that equals 1 if $s_t = s'$ and 0 otherwise, and the expectation is taken over all possible trajectories

starting from state s and following policy Π.

Nonetheless, the SR model fails to generalise across environments due to its intrinsic policy dependence, which must be relearned in a new space. To overcome this limitation, Gershman (2018) extended the SR model by incorporating a non-parametric Bayesian approach that allows the agent to infer the underlying state space and transition structure of the environment from limited experience, thus enabling rapid generalization to novel environments. This Bayesian SR model proved to account for a range of navigation behaviours, such as shortcut discovery and detour behaviour. Additionally, it could explain the role of the HPC in constructing and utilizing cognitive maps.

On another variation of Dayan's model, Geerts et al. (2024) designed the probabilistic successor representation (PSR), a generalisation of the original SR that represents the environment as a probability distribution over future states, conditioned on the current state and context. The PSR captures the uncertainty in the environment's dynamics and efficiently adapts to changes in context by updating the probability distribution over future states, without requiring a complete relearning of the environment.

Proving its representational accuracy over the Euclidean and Geodesic strategies (Figure 1.12(b)), the SR model has been shown to explain the emergence of place cells and grid cells in the hippocampal formation (Stachenfeld et al., 2017). The hippocampal place cells are proposed to represent the SR of each state, allowing for quick computation of value functions and the simulation of future trajectories. Entorhinal grid cells, on the other hand, represent the SR eigenvectors

across states (Stachenfeld et al., 2017), which encode a low-dimensional set for the SR and recapitulate the hexagonal firing patterns of these cells that allow for efficient encoding and computation of predictive representations.

Furthermore, evidence has been found for neural representations of SR in human brain regions (Russek et al., 2017). The HPC exhibited activity patterns that correlated with the SR of the chosen state, suggesting that it encodes the expected future states given the current choice, and consistent with a role of HPC in representing the spatial and temporal structure of the environment. In the ventromedial PFC, firing patterns correlated with the value of the chosen state, suggesting that this region could use the SR to evaluate the expected future rewards associated with each choice and coherent with its role in value-based decision-making. Additionally, activity patterns correlated with reward prediction error (RPE) signal were found in the VS, suggesting the potential use of SR to update the value of each state based on the observed rewards in this area.

1.3.3 Models of replay

In Section 1.2.3, we identified replay events as a prominent type of representation in the HPC. In space, replay's capacity to recapitulate old (Diba and Buzsáki, 2007; Foster and Wilson, 2006; Louie and Wilson, 2001; Skaggs and McNaughton, 1996) and explore new trajectories (Gupta et al., 2010; Ólafsdóttir et al., 2015) suggests that this mechanism might play a key role in generating and sampling internal models of the world (Foster, 2017; Pezzulo et al., 2013). Its detection during be-

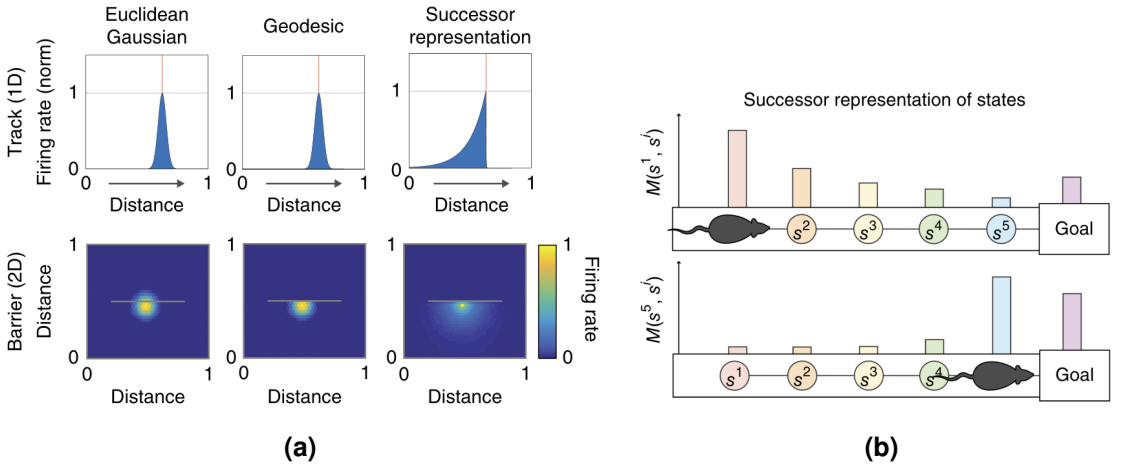


Figure 1.12: The successor representation model. (a) Euclidean, geodesic and SR model comparison of place field representations elicits SR place cell skewing indicates the anticipation of the successor state (top) and respects barriers in the environment (bottom). (b) Illustration of 0the SR state representation and place fields backward skewing (adapted from Stachenfeld et al. (2017)).

havioural (*online*) and sleep (*offline*) contexts and its variations in direction (*forward* vs *backward*) increase the complexity of this phenomenon, and contribute to the lack of a unified theory of replay function. Hippocampal replay is proposed to serve separate functions in different circumstances, from learning (Foster and Wilson, 2006) to planning (Diba and Buzsáki, 2007; Johnson and Redish, 2007; Pfeiffer and Foster, 2013; Singer et al., 2013), from spatial memory retrieval (Jadhav et al., 2012) and credit assignment (Kurth-nelson et al., 2016), to systems consolidation (Carr et al., 2011; McClelland et al., 1995). But, how exactly might replay facilitate the creation, use and adaptation of cognitive maps to efficiently and flexibly serve all these purposes?

1.3.3.1 Replay as a mechanism for planning in RL

In a model by Mattar and Daw (2018), hippocampal non-local place cell activations (including replay) are posited as the neural substrate for MB goal-directed planning. Besides predicting the need for replay, this model also accounts for the choice of memories that should be accessed at each time in the context of RL. The authors base their planning definition on the DYNA views of learning value from prior experience (see Section 1.3.1), while expanding on the trade-offs between MF and MB controllers. For this, the prioritization of memory access operations is framed in terms of action evaluation through Bellman backups. The utility of recapitulating a specific memory at a specific time is defined as the product of two intuitive terms - the gain and the need terms. Whilst the gain term prioritises states behind the agent when an unexpected outcome is encountered, the need term prioritises states ahead of the agent that are immediately relevant, thus establishing which memory ought to be accessed at a given moment for choosing the most rewarding future solution (as in some sort of modified prioritised sweeping mechanism).

Simulations of this model in a spatial navigation task demonstrate that an agent that retrieves experience in sequential order of utility accelerates learning while reproducing patterns resembling place cell replay (Figure 1.13). Moreover, the Mattar and Daw model recapitulates the existence and balance between forward and reverse replay (Diba and Buzsáki, 2007; Ambrose et al., 2016), its content (Pfeiffer and Foster, 2013; Ólafsdóttir et al., 2015; Singer and Frank, 2009) and the effects of experience (Foster and Wilson, 2006).

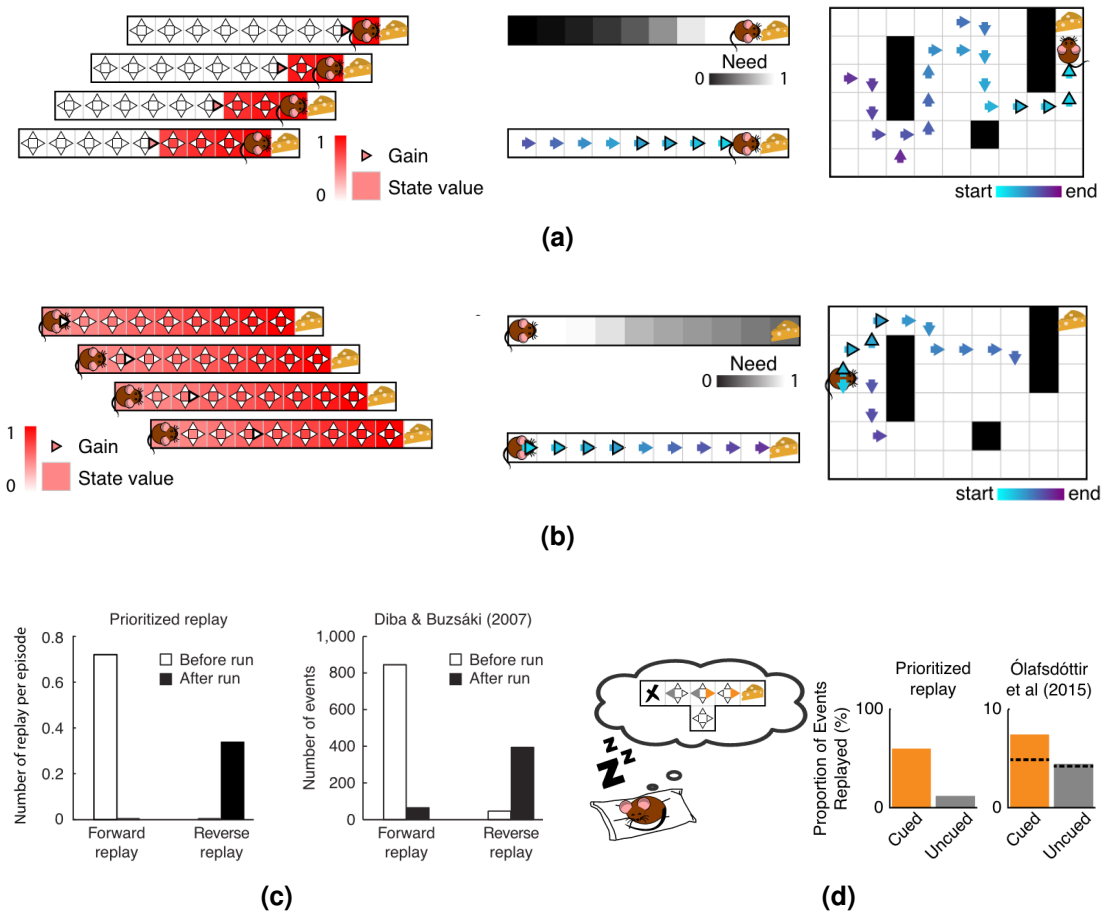


Figure 1.13: Replay through memory access prioritisation. Model examples of (a) reverse replay and (b) forward replay showing extended trajectories. (c) model recapitulation of Diba and Buzsáki (2007) correlation between forward and reverse sequences and stationary times before and after running, respectively (left), and recapitulation of the proportion of offline ripple events categorised into preplay of the cued and uncued arms of a T-maze by Ólafsdóttir et al. (2015) (adapted from Mattar and Daw (2018)).

Nevertheless, this theory does not discriminate between the role of episodic and semantic information in the computation of action values. It thus neglects the distinction between how MB and MF learning impacts the prioritisation process. Additionally, the authors leave the model untested in (1) spatial conditions of uncertainty and dynamical environments, which would require online deliberation features; and in (2) non-spatial settings, leaving open the question of how well it applies to abstract decision-making problems.

1.3.3.2 Online and offline replay contributions for planning

Online and offline replay are both thought to contribute to decision-making processes, although the underlying mechanisms remain unclear. With the purpose of differentiating the roles of online and offline events in flexible planning behaviour, Eldar et al. (2020) suggested a model where both types of replay serve distinct yet complementary functions in supporting planning.

Online replay is proposed to support on-the-fly planning by reactivating and chaining together recently experienced state-action sequences, thus helping the agent explore the consequences of different choices and update its value estimates in real-time. Offline replay, however, is sought to support consolidation and optimisation of the task's model during periods of rest, when the agent can access more distant experiences and recombine them to refine its internal model of the environment.

To guarantee the distinction, as much as possible, between MB and MF behaviour, the authors designed a novel state-based sequential decision-making task. Besides favouring a MB approach (93% vs 80% success rates for a MB approach vs the best MF algorithm), this task integrates a one-step and a two-step control contexts together with changes in the reward location and task structure in the face of which online MB reasoning is required for an accurate and quick adaptation.

Consistent with rodent replay studies that hypothesised prioritisation of trajectories that can soon be re-encountered, and for which one's policy can be improved

(Mattar and Daw, 2018); humans in this task also increase replay of surprising information, which in turn predicts relevant policy updates. Additionally, predominantly MB individuals replayed trajectories that reflected choices they wanted to reconsider, contrasting with rodent replay of deprecated trajectories (although for stable environments) (Gupta et al., 2010; Carey et al., 2019). More MF individuals, on the other hand, exhibited increased replay during rest periods and a lower success rate, arguing for incompatibility between offline policy generation and trial-by-trial adaptation.

1.3.3.3 Replay as a mechanism of compositionality

If we consider the HPC through the lens of a state-space compositional model, we can reconcile its two most striking roles: memory (Eichenbaum, 1999) and space representation (O'Keefe and Nadel, 1978). In recent theoretical work, Bakermans et al. (2023) propose a compositional model of HPC that supports flexible solutions in both spatial and non-spatial scenarios. They propose hippocampal cells to provide conjunctions of pre-learned primitives, specifying their arrangement in the current experience, which allows new rationalisation of hippocampal features and computations, such as replay.

In this theory, the role of replay is constructive. The compositional structure of memories allows extraction of common features and regularities of space one at a time, using replay to add them to existing memories and, thus, not overloading behaviour. Still online, replay is proposed to build new cortical primitives to help with poorly modelled experiences resulting from a mixed system that relies on

rapid compositional inference and modelling of transition structures. Offline, this replay should be involved in the same cortical primitive generation and in learning the policy on the compositional state-space (replay should generate new training policy maps and simulate trajectories over those maps).

Section 1.4

Thesis outline

In Chapter 1, I have provided a literature overview of the main topics necessary to follow and interpret the work of this project. We first approached the topic of *flexible behaviour*, from the origins of behaviourism with naturalists like Darwin and his pupils, all the way through Thorndike's Law of Effect, to then meet Tolman and his concept of cognitive maps. After briefly reflecting on Eichenbaum's reinterpretation of these maps, we continued diving into the goal-directed/habitual behaviour dichotomy and culminate with the discussion of recent and contemporary views on model-based decision-making behaviour. Secondly, we looked into the *hippocampus* as a substrate for the inherent flexibility of human and animal behaviour. Contextualising its role with a sequence of lesion and inactivation studies, we contemplate its anatomical and functional connectivities within and outside of the hippocampal-entorhinal complex, then exposing its neural signatures with a special focus on hippocampal spatial representations - from place cells to replay. Lastly, we reviewed the most relevant *computational models* of planning

and decision-making behaviour one should consider in the context of goal-directed navigation, with a particular emphasis on the successor representation framework, and reflected on the proposed potential functions of replay phenomena.

The remainder of this thesis will focus on the experimental work developed with the goal of understanding hippocampal neural representations in the navigation of a complex and dynamic world. The next chapters will be organised as follows:

Chapter 2 describes the methodological approaches used in this project, from the subject group, surgery procedures and histology, to the maze apparatus and task structure utilised, including the steps for data acquisition and processing prior to analysis.

Chapter 3 discusses the process of creation of optimisers for single mazes and sequences of maze structures. Nowadays, we possess extensive knowledge of how different regions of the brain support the representation of our internal models of the world, or cognitive maps. However, this information was often acquired in the context of simple behaviours in simplistic environments. Thus, we created a 49-state grid maze that can exhibit complex space structures and that, when appropriately designed, allows for testing of quantitative predictions against the behavioural and neural data collected. Through a series of iterations, I established a simple and intuitive way of deciding the environmental structures to test, and that were used for the experiments executed within this thesis.

Chapter 4 goes over the behavioural results of mice running a goal-directed navigation task on the designed 7x7 grid maze. There are several ways of navigating in an environment, but the more accurate your internal map of space is, the more successful and efficient you will be in maximising your return within a limited time window. I describe the basic features of the animals' behaviour while training and performing under this paradigm and quantify the contribution of planning and vector navigation strategies they use to maximise rewards in the maze over time.

Chapter 5 comprises the results of single-cell electrophysiology recordings in the mouse dCA1 region of the HPC. Place cells are commonly accepted as the ground of hippocampal spatial representation, with well-defined fields that tile the entirety of the environment; however, the extent to which they solely depict position in space or conjunctions of location and other related information remains unknown. I present the main features of space representation found in dCA1 of mice that navigated through a complex and dynamic environment and attempt to unravel the neural factors underlying such representations.

Chapter 6 focuses on the investigation of hippocampal ripple activity during task performance and prior and subsequent rest, or sleep, periods. Whilst being one of the most fascinating mechanisms for cognition support, the many kinds of hippocampal replay have been attributed to different functions over the past decades, from memory consolidation to planning. I analyse the content of online and offline replay events and investigate their involvement in flexibly supporting

planning and structural representations, respectively.

To conclude, in Chapter 7 I compile the evidence collected throughout the execution of this thesis, highlighting the *main contributions* of this project to the field of cognitive maps in the context of complex spatial navigation in mice. I also propose improvements and *future work* that could further expand on the results presented here.

2

General Methods

Section 2.1

Subjects

All procedures were performed per the UK Animal (Scientific Procedures) Act 1986 and according to the University of Oxford's Local Ethical Review Committee guidelines.

For the maze experiments, a total of 24 wild-type (WT) C57BL6 male mice, distributed into three groups of eight, were submitted to training in the 3x3 grid maze. The first cohort of eight mice was used for behavioural testing only in the 7x7 grid maze. From each of the following two cohorts, the best five animals (joint evaluation of performance and body weight) were chosen to proceed to electrophysiology testing in the 7x7 grid maze and submitted to surgery for silicon probe chronic im-

plantation (Section 2.6.2). Throughout training and behavioural testing, animals were housed in groups of up to four mice. The probe-implanted animals were single-housed from the day of surgery and throughout the entire testing period.

All subjects were maintained in a 12-hour light/dark cycle. Food was provided *ad libitum* at all times. When water-restricted, mice were kept above 85% of their initial weight. During training, the water-restriction procedure consisted of 6 days when the subjects received a minimum of $1mL$ of water per day while in the task, followed by one rest day with 1 hour of water *ad libitum*. During testing, mice were tested daily for up to 28 consecutive days, after which one rest day was mandatory with 1 hour of water *ad libitum*.

All experiments took place during the light phase.

Section 2.2

Elevated grid maze

The entire project of this thesis was developed using an elevated grid maze apparatus initially designed and proposed by Dr. Thomas Akam. The first versions of this maze - a 2x3 and a 6x6 - were built by Dr. Thomas Akam (PhD) and PhD student Michael Pereira in the laboratory of Professor Rui Costa, at the Champalimaud Foundation - Lisbon, Portugal.

Alterations made to the 6x6 maze in Lisbon to allow for imaging recordings and the experiments run by Michael and I greatly informed the construction of a

more robust and reliably functioning maze in Oxford. The necessary changes to implement were thought and discussed by Dr. Akam, Michael and I; and the design of the mazes was implemented by Dr. Akam.

2.2.1 7x7 elevated grid maze

For testing purposes, animals ran a goal-directed navigation task (Section 2.3) in a 7x7 elevated grid maze.

The maze comprised 49 octagonal towers (also referred to as platforms or nodes) arranged in 7 columns of 7 towers each, and 84 bridges (also referred to as links or edges) when fully connected, with total dimensions 1.38 x 1.38 x 2.10m. Its structure was built of extrusion rails [Misumi] and enclosed with black aluminium panels [Cut-Tec], assuring a Faraday-cage-like effect for electrophysiology recording purposes. The maze was always accessible by 3 sides to facilitate access to the equipment and subjects, with 4 removable doors on each side. The wall that did not provide access into the maze was, exceptionally, half black and half white. White acrylic mosaics [Trent Plastics] were fitted around the towers' bases as flooring, avoiding access to cables and tubing if subjects would ever be at floor level.

Tower centres were disposed 18cm from each other. Platform elevation was achieved using 30cm high extrusion rails [Misumi], and octagons composed of 2 layered acrylic plexiglass (one white and the other transparent) [Trent Plastics]. Each platform had one nose port at its centre for delivery of water reward and one

4cm frosted transparent cylindric rod [Sheet Plastics] to diffuse the light cueing the goal location. Underneath every platform was one PyControl poke board (Eurocircuits) equipped with an infrared (IR) beam to detect nose pokes, one solenoid for water delivery [The LEE Company] and one speaker and an LED to provide sound and visual cues, respectively. Water delivery was assured through a network of rubber tubes and valves [VWR] that connected a water-filled syringe to each platform's solenoid. Ethernet cables [Farnell] connected every poke board to the PyControl main boards (Eurocircuits) located on one side of the maze.

Bridges were made of transparent acrylic plexiglass [Trent Plastics], with dimensions 5x10cm, and with 1cm borders on each side to provide animal safety when travelling the maze. The addition or removal of bridges within the maze was done manually.

Video was acquired using a FLIR Point Grey Chameleon3 [1.3 MP Mono USB 3.0 Camera, 1/2", On Semi P1300 CMOS, 1280 x 1024 at 150 FPS, Global shutter, CS-mount camera] and a wide fisheye lens [1/2.7In CS mount 2.8-10.0mm F1.2-16C Aspherical 3 Megapixel Manual Iris IR Corrected]. Small infrared LEDs [RS Components] were installed on the maze walls to allow good filming quality in the dark.

The nodes and bridges of the maze were named according to an alphanumerical system that worked as a discrete coordinate system for further analysis of behaviour and neural activity in this setup (Figure 2.2).

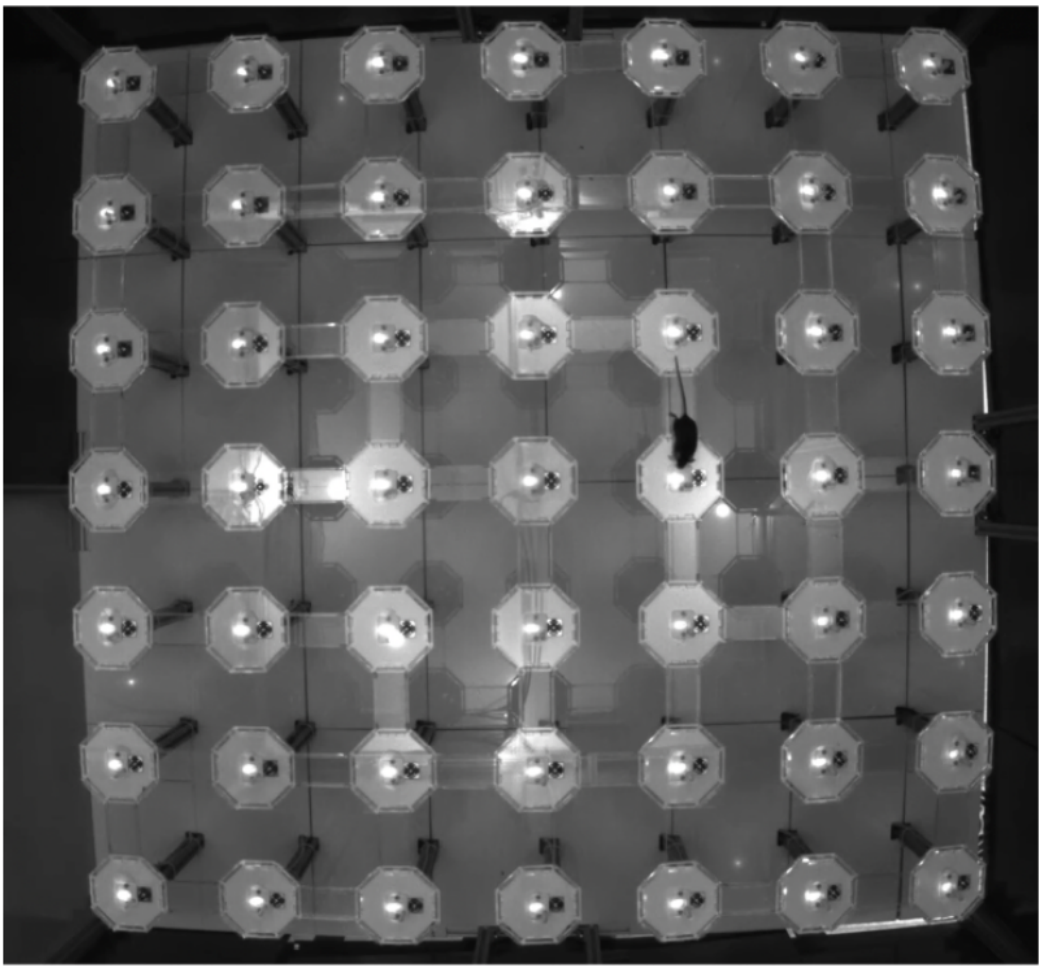


Figure 2.1: 7x7 grid maze apparatus. (a) photo of the maze setup from the top. It is comprised of 49 towers, and its structure is defined by the existing bridges that can be manually placed connecting the towers. (b) close-up photo at the platform level. Each platform is equipped with a pole that diffuses the light of an LED and a speaker for visual and auditory reward cueing. In the centre of each platform, a nose port can deliver water rewards if selected as the goal location.

2.2.2 3x3 elevated grid maze

Adaptation to setup and training of mice was done using a 3x3 elevated maze, or small maze. The design and equipment used on this maze are analogous to the 7x7 elevated maze described in Figure 2.1., with some exceptions - the maze is only accessible by one side with two doors, all the walls of the maze are black.

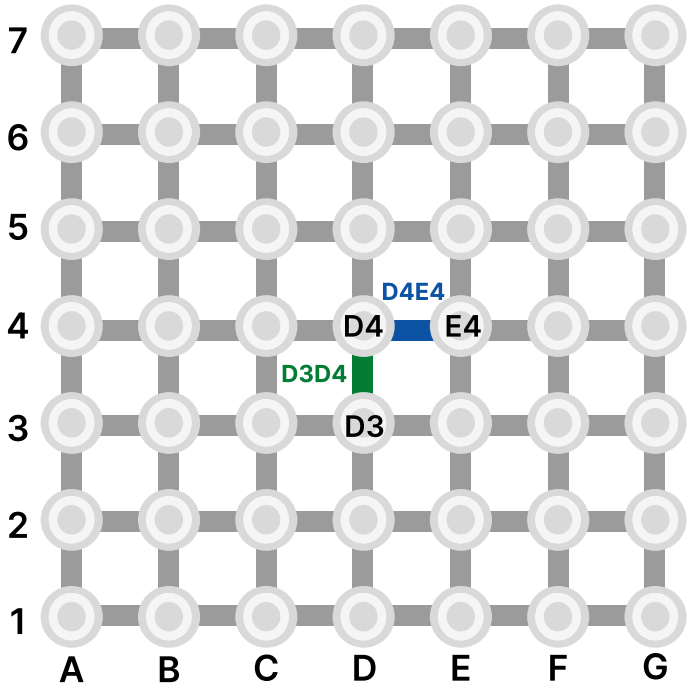


Figure 2.2: Discrete coordinate system in the 7x7 grid maze. To facilitate the identification of the animals' location in the maze, we established an alternative and discrete coordinate system. Columns are named using a letter from A-G and rows using a number from 1-7. Thus, we can refer to a node by the combination of its column and row position (e.g. D3, D4, E4) and to a bridge by combining the name-code of its neighbouring nodes in a crescent manner (e.g. the bridge between D3 and D4 is D3D4 and the bridge between D4 and E4 is D4E4).

This maze had a total of 9 towers and 12 bridges when fully connected, and overall dimensions 0.66 x 0.66 x 0.8m.

Section 2.3

Goal-directed navigation task

Mice were trained and tested under a goal-directed navigation paradigm.

On each trial, one of the platforms was randomly selected as the goal location. This goal location was cued with a brief (500 ms) white noise and with the light of an



(a) (b)

Figure 2.3: 3x3 grid maze apparatus. (a) view from the exterior of the setup; (b) view of the inside of the maze, which is analogous to the 7x7 grid maze but comprised solely by nine towers.

LED diffused by a 4cm cylindrical rod, to allow visualisation from afar. Having seen the light cue, the animals must travel to the goal location to collect a water reward. The visual cue was switched on until the animal poked in at the goal location port. After poking in at the goal port, a water reward was released - a variable volume during training; 2μL during testing - and a reward consumption timer was initiated. Once the subject had last poked out of the port for 500ms or more, an inter-trial interval (ITI) timer of 500ms would start, after which a new reward location was randomly selected and cued, signalling the beginning of a new trial.

The task was run in the dark, with only a very dim light present in each trial due to reward visual cueing. A new trial began when, and if only, the animal had completed the previous trial. There was no penalty of any kind for the subjects.

As sessions had a limited duration of 30 minutes, mice were highly motivated to complete the trials as quickly as possible in order to collect as much reward as possible, i.e. achieve a higher number of trials.

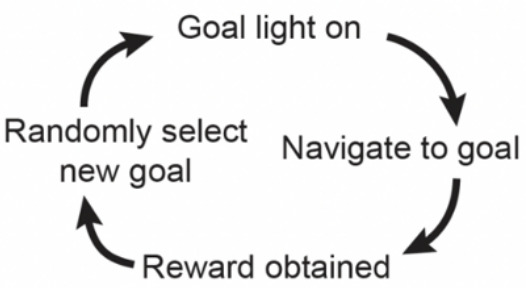


Figure 2.4: Goal-directed navigation task. A trial in the task starts with a light cue switching on at a randomly allocated goal location. The mouse then needs to navigate to it, collect the reward by poking the port at the correct platform, after which a 500ms ITI is initiated. A new trial starts with the visual cue of a new randomly selected reward location.

The task script was custom-coded in Python language to communicate with PyControl equipment and GUI (`https://pycontrol.readthedocs.io/en/latest/user-guide/programming-tasks/`) (Akam et al., 2021).

Section 2.4

dCA1 electrophysiological recordings experimental schedule

Two cohorts of five mice ($n = 10$, total) performed a goal-directed navigation task (Section 2.3) on a 7x7 elevated grid maze (Section 2.2.1). Tested maze structures were chosen according to the optimisation process described in Chapter 3, and

the final sequence of evolving structures is depicted in Section 3.3.2.2.

Every day, mice will experience a sequence of three rest and two maze sessions (Figure 2.5). In the morning, they start with a rest/sleep session in a covered plant vase (height = 16cm), followed by the first maze session of the day (maze A). Immediately after experiencing the first maze, animals are placed back into the plant vase for the second rest session and subsequently return to their home cage. At this point, we (experimentalists) manually change the maze link determined by the chosen sequence of maze structures - we either add or remove one link of the maze structure. In the afternoon, subjects restart with the second maze session of the day, where they experience the newly changed structure (maze B); followed by the third and last rest session of the day. The following day, mice will perform the same experimental schedule, but this time experiencing the last maze structure of the day before as the first maze structure of the day (maze B); this structure will again be changed by one link, and the new structure will be presented to the mice in the afternoon (maze C). This schedule repeats over days, allowing the mice to experience a series of 24.889$\pm$4.114 (mean$\pm$SD across mice) different maze structures and enforcing a gradual structural change of the maze.

All rest and maze sessions had a 30-minute duration. All rest and maze sessions were recorded using a 128-channel silicon probe (Section 2.6.2.1).

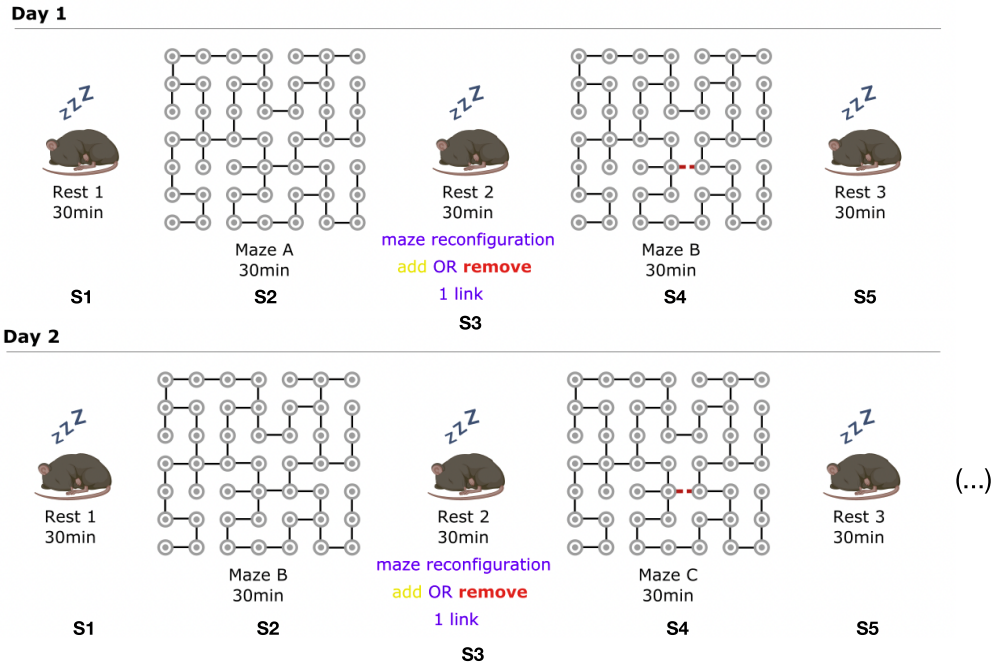


Figure 2.5: Hippocampal recordings experiment schedule. On each day, mice experience five 30-minute sessions - three sleep (S1, S3, S5) and two maze sessions (S2, S4). Day 1: in the morning, mouse experiences a sleep session (S1) inside a tall vase filled with bedding, then a maze session with structure A (S2) and a new sleep session (S3); in the afternoon, the mouse experiences a maze session with a new maze structure B (differing by 1 link from the morning structure) (S4) followed by the last sleep session of the day (S5). Day 2: the schedule repeats, but the first structure the animal sees in maze S2 is the one last seen the day before (maze B) in S4, and the afternoon maze (S4) will suffer yet another link change into maze structure C. This slowly evolving maze schedule repeats until the end of the experiment, so that on each day the animal experiences two different structures - the one they last saw on the day before and a new one.

Section 2.5

Data acquisition

In all testing experiments, the session's events and state information were acquired in a .txt format using PyControl boards (`https://github.com/pyControl`) (Akam et al., 2021). The sessions' videos were recorded in a .mp4 format using a Bonsai

(Lopes et al., 2015) custom-written workflow. For the electrophysiology experiments, neural data were acquired in .dat and .npy formats using Open Ephys acquisition hardware and software (`https://open-ephys.org/`) (Siegle et al., 2017).

During training in the 3x3 maze, animals were filmed for monitoring purposes only and videos were not recorded. For all the experiments depicted in this thesis, we used a standard Dell desktop computer enhanced with a 1TB SSD internal drive and a NVIDIA® GeForce RTX™ 3080Ti graphics card to allow for high-throughput neuronal and video signal data collection and Kilosort cell sorting in between sessions (to inform probe lowering movements).

Section 2.6

Surgery

2.6.1 General

Mice were injected with buprenorphine (0.08 mg/kg), meloxicam (5 mg/kg) and glucose saline (0.5ml) immediately before surgery. Isoflurane (3% induction, 0.5-1% maintenance) was used throughout surgery for anaesthesia purposes. Marcaine (max 2 mg/kg) was injected into the scalp before placing mice into the stereotaxic frame [Kopf Instruments]. Subjects were maintained at approximately 37 degrees Celsius with a rectal probe and a heating blanket [Harvard Apparatus].

Subjects were given additional doses of gelatine with meloxicam each day for

3 days after surgery, and were closely monitored for 7 days post-surgery.

2.6.2 Silicon electrophysiology probe implant

A small screw was chronically implanted in the subject's skull at the far most right side, close after the lambdoid suture, for probe grounding purposes Figure 2.6. A cranial window of approximately 1000x500 μm was drilled, with its centre targeting just above the left dCA1 region (AP: -2, ML: -1.5 from bregma). One 2x64-channel silicon probe stack (assembly details in Section 2.6.2.1) was chronically implanted just above the left dCA1 region (AP: -2, ML: -1.5, DV: -0.7 from brain surface). This allowed for posterior gentle lowering of the probes in search for new cells and higher cell yields, which was performed at times when the number of cells detected by the probes would drop significantly. The probes' red ground cables were soldered to the ground screw implanted in the subject's skull.

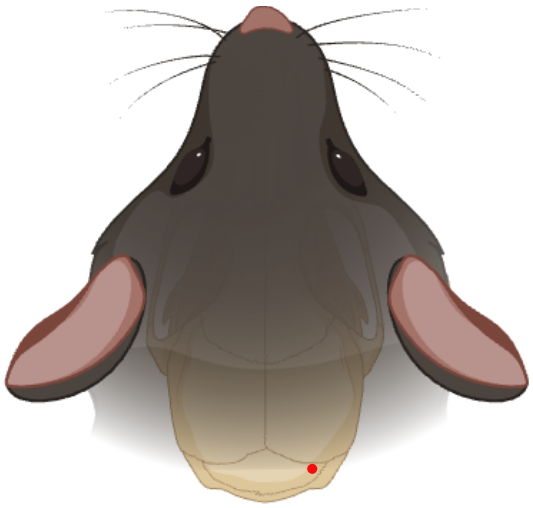


Figure 2.6: Schematic of the probe ground screw location in the mouse skull.

2.6.2.1 2x 64-channel silicon probe stack assembly

We used a 3D printed support (custom-designed by Dr. Thomas Akam) to assemble two 64-channel 6-shank silicon probes [Cambridge NeuroTech silicon probe F] (`https://www.cambridgeneurotech.com/neural-probes`) parallel to each other.

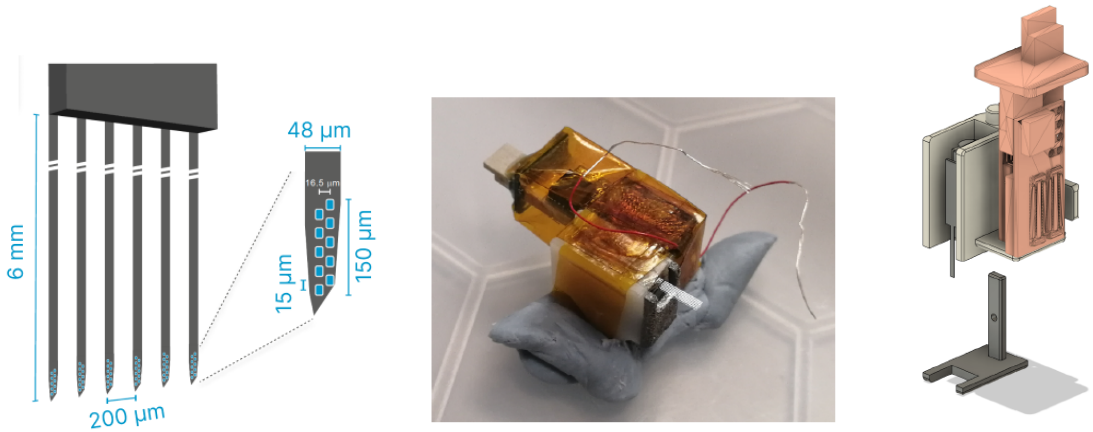


Figure 2.7: 128 channel dual mini-amp single nanodrive implant. (a) Cambridge neurotech probe schematic, illustrating the disposition of the 6-shank 64-channel silicon probe. (b) 2-probe stack and dual mini-amp single nanodrive implant mounted using a (c) custom-made (by Dr. Thomas Akam) body to allow probe recovery and reuse at the end of the experiment.

Section 2.7

Data pre-processing

2.7.1 Video processing

High-throughput video analysis of 7x7 maze sessions was done using DeepLabCut v2.1.9 toolbox (`https://github.com/DeepLabCut/DeepLabCut`) (Mathis et al.,

2018). For network training purposes, a total of 400 frames was uniformly extracted (5 frames/video) from 80 sessions of the 8 mice (10 sessions/mouse) used for the 7x7 maze behavioural experiment. Eight points per animal were consistently and manually labelled in each frame (Figure 2.8). These data were then split into a training and a testing datasets, and a ResNet-50 was trained using the training dataset for 1,000,000 iterations, or until the loss plateaued. The labelling accuracy of the trained network was then evaluated using the test dataset. Once sufficiently accurate, the trained network was used to analyse all the 7x7 maze videos used for the purpose of this work and whose results are depicted in this thesis. This resulted in the chosen labels being applied to the analysis videos, with the estimated x- and y-coordinates and a confidence score for each label. Confidence scores above 85% were considered satisfactory.

The 7x7 maze videos recorded a 140x140 cm arena at a relatively high resolution of 1280x1024 px, which initially rendered a 2-hour analysis time per 30-minute video. Resolution and temporal downsampling of the video prior to analysis with DLC proved inefficient, reducing training accuracy and confidence scores to unacceptable levels. The DLC dynamic-cropping feature proved to be the best solution - especially recommended for large frames where the animal occupies a small fraction of it. At the video analysis stage, we configured a dynamic cropping margin of 20 with a detection threshold of 0.5, and this was sufficient to accelerate analysis to 17 minutes per 30-minute video while maintaining an accurate position estimation.

The processing of DLC video data results was done with custom-written code in Python. Labelled point *mid1* (in between the front paws) was assumed to be

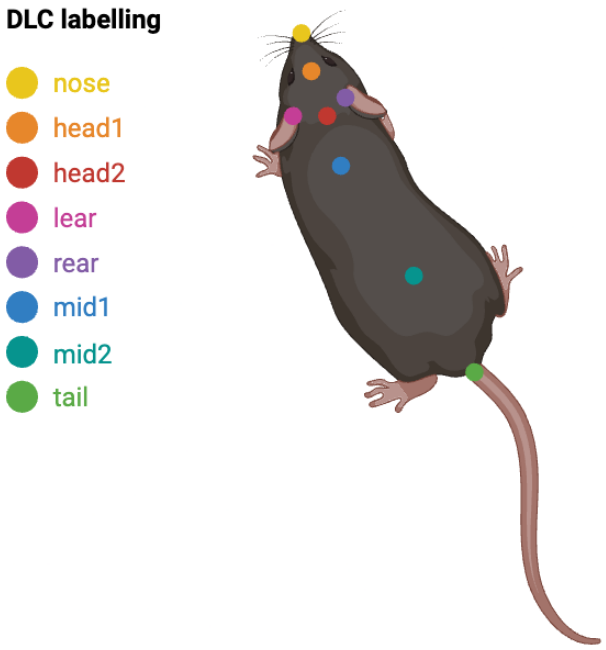


Figure 2.8: Schematic DLC labelling of mouse parts. Illustration of the eight mouse body parts identified in DLC at the frames labelling stage for tracking.

the animals' centre of mass, thus providing its location in the environment - we consider a mouse has willingly entered a new location in the maze once its front paws touch that location. Low confidence score estimations of *mid1*-point ($< 99\%$) were not considered and alternatively linearly interpolated. Pose estimations were then smoothed across time using a Gaussian smoothing ($\sigma = 5$) to avoid recurrent jittering in the animal's position.

To account for the video image distortion caused by the fisheye feature of the camera, we performed a direct linear transformation of the image coordinates (inspired on `https://github.com/JD-Canada/Tracker3D/blob/master/DLT.py`). We combined the x- and y-coordinates (in pixels) of every node's poke-port of every experiment, or in between adjustments of the camera position in the setup, along

with the real dimensions of the nodes and bridges of the maze (in cm) and allowed for the adjustment of the video coordinates to the real physical coordinates of the space.

2.7.2 Electrophysiological signal processing

Kilosort 2.5 Python package (`https://github.com/MouseLand/Kilosort`) (Pachitariu et al., 2024) was used for spike sorting of the neural signal recorded from mice in the 7x7 grid maze. Example individual-session ephys recordings were sorted in between maze sessions to inform the lowering of the probes on each experimental day. At the end of the experiments, same-day and -animal ephys files (2 maze and 3 sleep ephys files per animal) were stitched together using a custom-written Python code that generates (1) one large ordered concatenated ephys file and (2) an additional .txt file that contains the (a) collection order, (b) bit size and (c) percentage of the whole day performance information of each individual session ephys files. Results of the concatenated files are more reliable than those of individual sessions, and in theory allow for direct comparison of neuron activity across sessions within the same day. We did not sort session files across days, and so we could not perform any type of analysis that would assume the same neurons across different days.

2.7.3 Signal alignment

For the 8 mice that ran a goal-directed navigation task in the 7x7 grid maze, we collected PyControl and video data. For the ephys experiments, we collected PyControl, video and electrophysiological data from 10 mice. It is impractical to achieve simultaneous start times for all the systems when running our experiments. Additionally, each system operates independently and maintains its own time reference. Hence, aligning timestamps across systems becomes imperative at the data preprocessing stage, and can often be a challenging task.

To set all the necessary systems to the same reference frame, we used the Rsync tool and module provided by PyControl (`https://pycontrol.readthedocs.io/en/latest/user-guide/synchronisation/`). Briefly, the PyControl, PointGrey camera and OpenEphys operate 1kHz, 60Hz, and 30kHz, respectively. Using the *Rsync* tool, PyControl outputs sync pulse trains with random inter-pulse intervals that are recorded by the video and ephys systems in their own time reference, and can later be used to convert times from one system into another's reference frame with the *rsync* module code.

Section 2.8

Data analysis

All behavioural and electrophysiological analysis from grid maze experiments were performed using custom-written Python code and will be described in more detail in Chapters 4 to 6 of this thesis.

Section 2.9

Histology

Electrophysiology experimental mice were terminally anesthetised at the end of the respective experiments with sodium pentobarbital and transcranially perfused with saline and then 4% PFA solution.

For the first cohort of 5 mice that ran this experiment, 50μm coronal brain slices were cut using a vibratome and covering the entire hippocampal formation region. The slices were then stained with crystal violet and later imaged using a Zeiss Axioscan 7 Microscope Slide Scanner to determine the placement of the probes (Figure 2.9). For the second cohort of 5 mice that ran the experiment, the brains were not stained and the imaging for probe trace detection (Figure 2.10) was done using a serial-section 2-photon imaging microscope and the

BakingTray acquisition software (`https://sainsburywellcomecentre.github.io/OpenSerialSection/acquisition/`, `https://github.com/SWC-Advanced-Microscopy/BakingTray`) at the Sainsbury Wellcome Centre - UCL, in London.

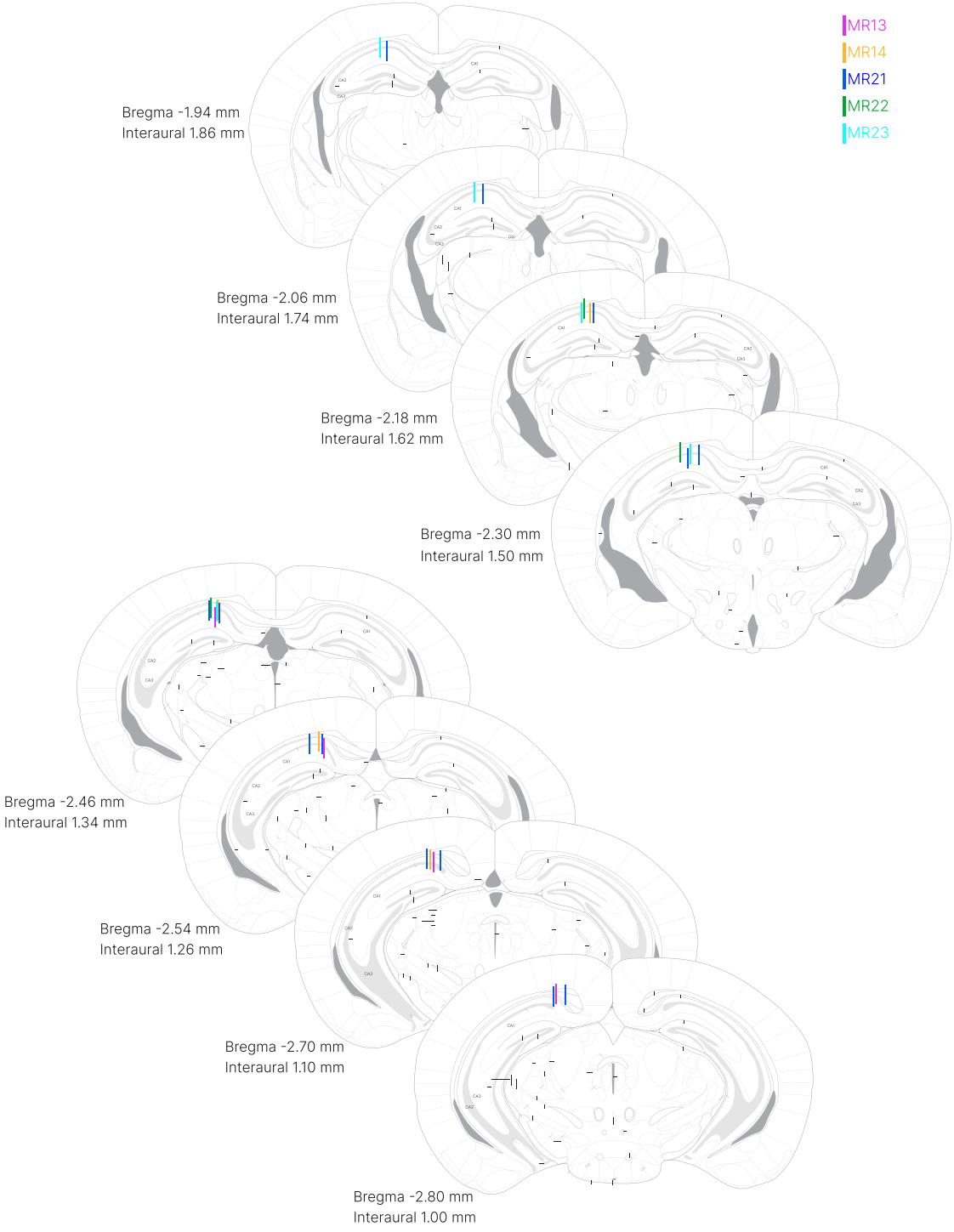


Figure 2.9: Histological assessment of fibre placement - first recordings cohort. Location of the probes implanted in dCA1 of the first group of five mice recorded in the 7x7 grid maze. The traces identified post-experiment are identified with a different colour for each mouse.

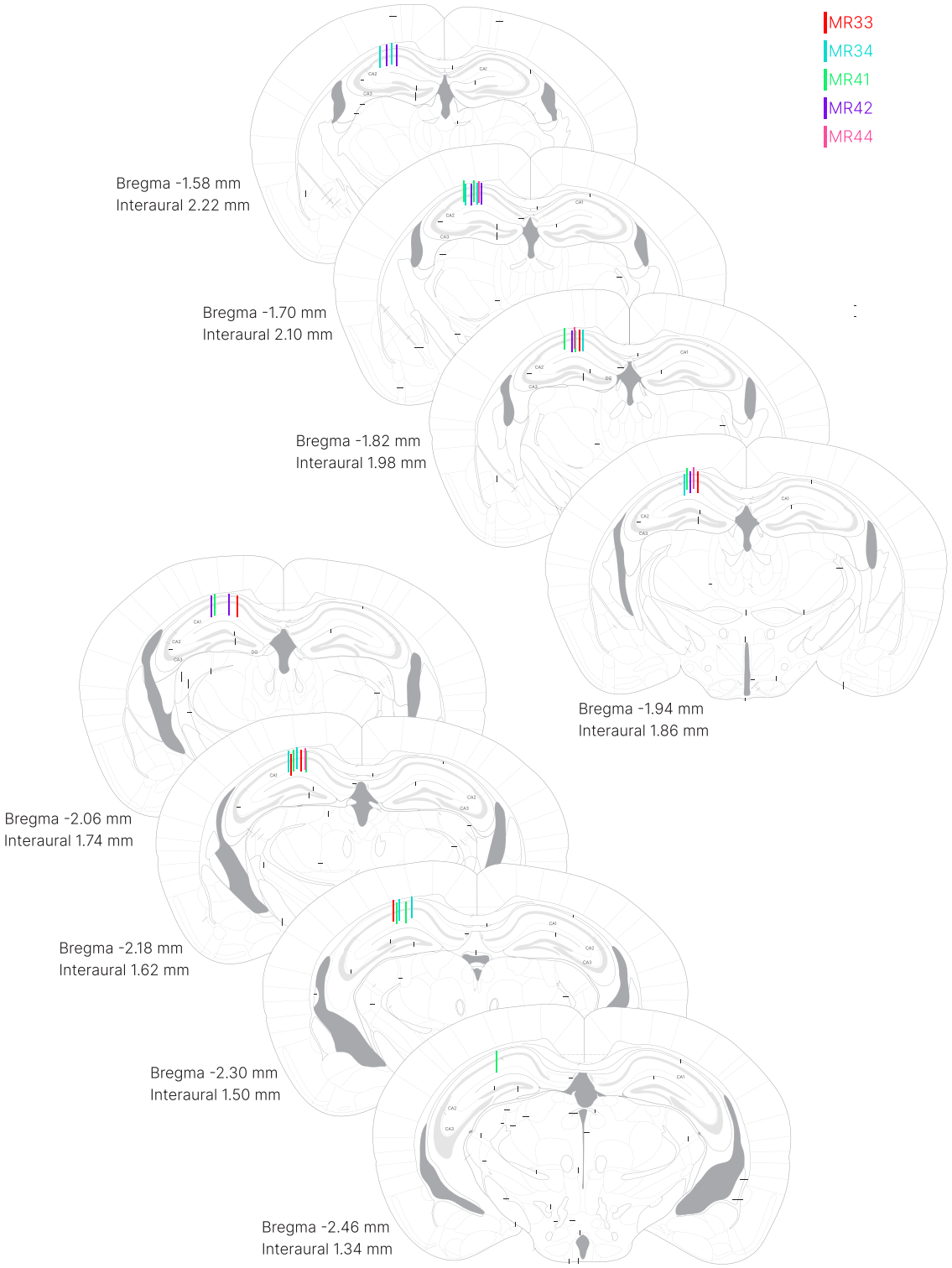


Figure 2.10: Histological assessment of fibre placement - second recordings cohort. Location of the probes implanted in dCA1 of the second group of five mice recorded in the 7x7 grid maze. The traces identified post-experiment are identified with a different colour for each mouse.

3

A sequence of maze structures optimiser for complex navigation testing

Section 3.1

Introduction

Rodent spatial navigation is a powerful model system for studying the neural basis of cognitive maps (Tolman, 1948a), due to its ethological validity and our precise knowledge of how neurons represent space. Over the past decades, scientists have built a comprehensive repertoire on how individual neurons precisely represent aspects of the spatial world - place cells (O'Keefe and Dostrovsky, 1971),

grid cells (Hafting et al., 2005), object vector cells (Høydal et al., 2019), boundary vector cells (Solstad et al., 2008), head-direction cells (Taube et al., 1990), among other helpful representations in space.

Nevertheless, all this knowledge and maps of representations have been built using simple and, many times, deterministic environments. From open arenas (O'Keefe and Dostrovsky, 1971; Hafting et al., 2005; Taube et al., 1990; Solstad et al., 2008; Kropff et al., 2015; Høydal et al., 2019) to open water (Morris et al., 1982; Dupret et al., 2008; Gallagher et al., 1993) and Barnes mazes (Harrison et al., 2006; Barnes, 1979; Fox et al., 1998), using variations of T-mazes (Bannerman et al., 1999; Deacon et al., 1983; Packard and McGaugh, 1996), Y-mazes (Conrad et al., 1996), M- and W-mazes (Kim and Park, 2008), plus mazes and radial arm mazes (Olton and Papas, 1979), all the way to linear tracks (Diba and Buzsáki, 2007) and loops (Sun et al., 2020), we could perhaps summarise more than 95% of the work done to this day in the study of spatial cognition. Although we should celebrate all the beautiful discoveries researchers have achieved over the years using such apparatus, one should also recognise limitations in these setups as to the complexity of behaviours they allow us to understand.

Recently, more complex mazes have appeared. Some examples are the hairpin maze (Derdikman et al., 2009), which consists of a zig-zagged stack of interconnected maze segments with equal shape and orientation; a seven-node tree-like maze (Grieves and Jeffery, 2017) where trajectories between most combinations of every two locations can occur through two different paths; and the honeycomb maze (Wood et al., 2018), composed by 37 platforms placed together in the shape

of a honeycomb and that can be individually manipulated to control the choices of the animal at each point in the maze. Alternatively, a new era of spatial tasks has emerged with the use of VR. VR allows a world of infinite choice in the layout and complexity of spaces to test. These can be used with humans to allow simultaneous collection of neural activity while they navigate intricate worlds (Howard et al., 2014); they can even be used with rodents, teaching them to run for rewards in endless corridors that can be as simple as a linear track (Gabriela Michel et al, unpublished) or incorporate extraordinary features such as wormholes (Mattias Horan et al., unpublished).

While the new maze paradigms described enable increased complexity of behaviours and representations when compared to their predecessors, they do not allow the design of large state spaces with complex and dynamic structure. Or, when comprising a large space, they do not allow the behavioural flexibility we wish to study and understand in the mouse. On another topic, VR setups provide a breadth of choices when it comes to the design of the environmental structure, and certainly allow for online dynamics to occur. However, when applied to rodents, these require the animal to be head-fixed, imposing a range of movement constraints, which renders the entire behaviour less naturalistic than that of a freely moving navigation task. In fact, several studies have reflected on the difference in stability and precision of neural representations in VR vs real-world navigation, with the amount of multisensory cues being reduced and the absence of vestibular and proprioceptive cues in the first one when compared to the second (Chen et al., 2013; Cushman et al., 2013; Minderer et al., 2016).

For these reasons, and because we wanted to investigate the hippocampal representations of complex and dynamic environments, we designed a 7x7 elevated grid maze that allowed the construction of intricate world structures, as well as its modification. With the right configurations, the implicit geometry of these environments allows for the testing of quantitative predictions at the behavioural and neuronal levels. What is then the right structure to test, and how can we systematically define it?

In this chapter, we share the process of building an optimisation algorithm for generating the structures that allowed us to conduct the experiments of this project. As the grid maze can, in essence, be interpreted as a graph network - the platforms are its nodes, and bridges are the edges -, we draw from concepts of graph theory, such as centrality, to build our simulator (Wilson et al., 2013). Additionally, we take advantage of a few generic pre-existing optimisation algorithms throughout the process of our search. In the end, we optimised single maze structures, as well as sequences of slowly changing structures (1 edge at a time) to obtain the sequence of mazes tested in this thesis.

Section 3.2

Methods

The single maze and sequence of mazes' optimised simulators described in this chapter were obtained using Python custom-written scripts. For defining the maze's

grid-like structure and calculating many of its mathematical properties, we have used NetworkX, an open-source Python library (`https://networkx.org/`) for the creation and manipulation of complex networks. This package provided a wide range of already-built algorithms for network analysis, such as measures for centrality, and allowed for straightforward visualisation of the maze structures which can be approximated by network graphs. Additionally, NetworkX integrates well with other useful data analysis libraries like NumPy and SciPy.

All optimisers were built considering we were using a 7x7 grid maze apparatus. The optimisation process was also highly informed by the experience and intuition gained while working with Michael Pereira on the 6x6 grid maze at the Champalimaud Foundation, in Lisbon, Portugal.

3.2.1 Optimisation tools

For constructing our single maze and sequence of mazes optimiser simulators, we relied on other optimisation algorithms at some of the intermediate steps. In this section, we introduce each of these auxiliary algorithms and their functions.

3.2.1.1 Genetic algorithm (GA)

Genetic algorithms are a powerful tool for solving search and optimisation problems within the field of "evolutionary computation" (Sivanandam and Deepa, 2008). Invented and developed by John Holland and colleagues in the 1960s and 70s, the GA was first presented as an abstraction of biological evolution (Holland, 1992), which has proven to be a valuable source of inspiration for solving various com-

putational problems, from protein engineering to predicting financial markets dynamics (https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=630a7d2f0506cb5a01b09f07eef8a3b5a3af0387). Some of the most demanding computational problems lie with (1) the search through an ample space of options, (2) the flexibility to perform well in volatile environments and the capacity to generate (3) innovative and (4) complex solutions. In biological settings, evolution efficiently tackles these exact points: it searches among a huge sea of genetic possibilities to generate the fittest organisms, a measure that varies with the organism's lifetime and strongly depends on the external environment, many times reaching the most complex and effective solutions through the simplest random variations (Mitchell, 1998).

A simple variation of the GA, and the one we employed for the effects of this project, relies on the following steps (Figure 3.1):

1. Generation of multiple random chromosomes - in our case, the chromosomes, or maze vectors, are binary vectors (1x98 edges) that represent m maze structures where bridges exist in positions with value 1 and do not exist in positions with value 0. These maze vectors were obtained through a transformation of the mazes' adjacency matrices (49 nodes x 49 nodes).

2. Rank the chromosomes by fitness - we defined fitness as the capability to discriminate between Euclidean and geodesic navigation strategies within each maze structure.

3. Apply random crossover to every $m/2$ pairs of ordered chromosomes with

randomly selected crossover points that could range from positions 1 to 97 in the chromosome.

4. Insert random mutations in the chromosomes (vectors) with probability of mutation $P_{mutation}$ of 0.002, to increase structure variability.

5. Repeat steps 2 to 4 for g generations.

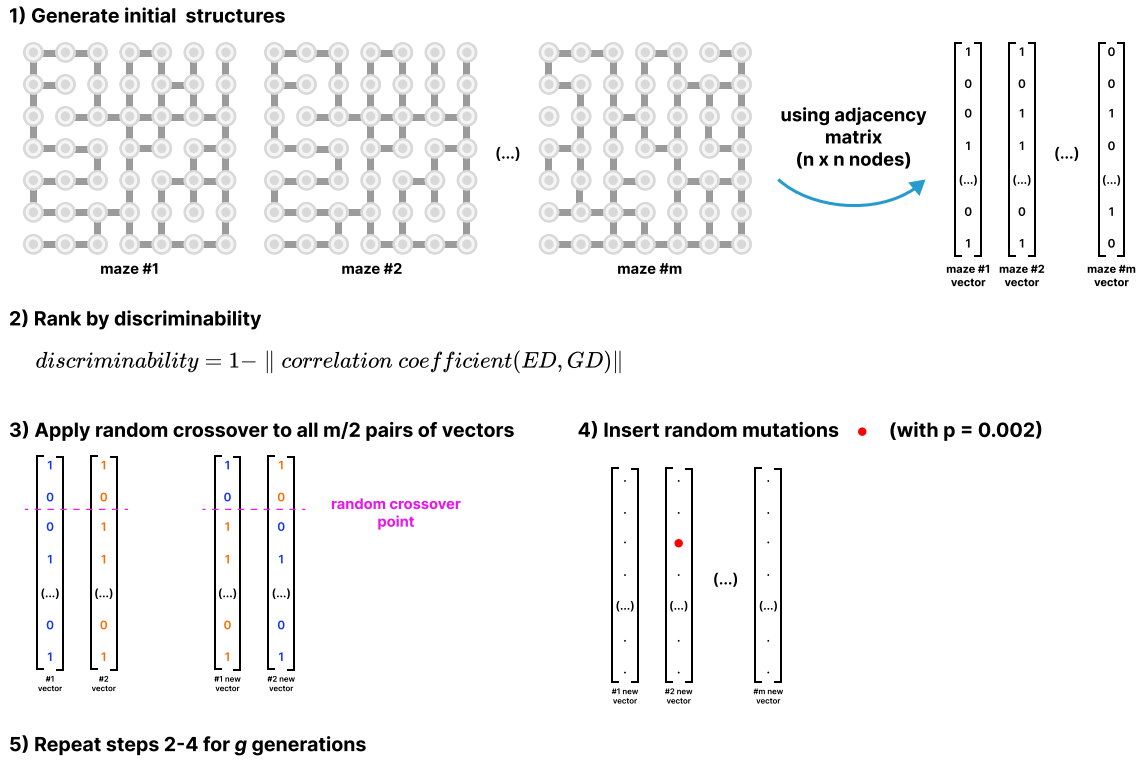


Figure 3.1: Schematic of the genetic algorithm steps used for the individual maze structure optimisation.

3.2.1.2 Pareto Front

Known as the set of all Pareto-efficient solutions, the Pareto Front provides a solution that cannot be improved for one of the considered objectives without penalising another. It was initially used by Vilfredo Pareto in the study of economics (Pareto,

1896), and it is widely applicable to other fields, such as engineering and biology, where you often need to perform multi-objective optimisation (Deb, 2001).

Section 3.3

Results

3.3.1 Single maze optimiser

We sought to be able to distinguish between model-based (or planning, or geodesic, or shortest-path) and model-free (or vector, or Euclidean) navigation strategies (Figure 3.2), both detected to be represented in the hippocampus (Howard et al., 2014), to allow the testing of quantitative predictions against the neural data (Stachenfeld et al., 2017).

In a squared grid maze, the situation when the maze is fully connected will be the case where Euclidean and geodesic navigation strategies will be mostly correlated (Figure 3.3, left). In that particular case, optimally travelling from any start location to any goal location on the maze will take identical routes for both considered models, as the maze structure is composed solely of multiple unit squares.

By removing some of the links within the maze, and consequently reducing the number of unit squares available, we are decreasing the correlation between Euclidean and geodesic navigation strategies (Figure 3.3, centre) and increasing its navigational complexity. However, optimising for the decorrelation of these two

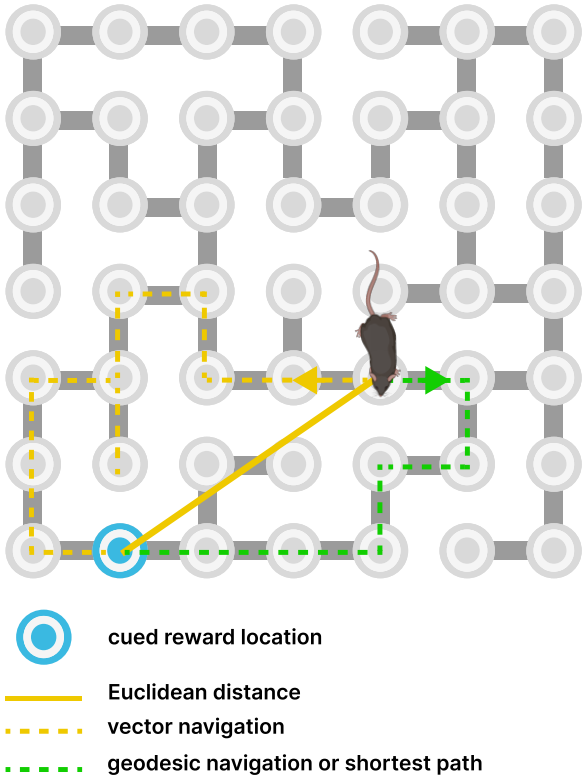


Figure 3.2: Maze example of vector and geodesic navigation strategies on the maze. From the animal's current position (mouse cartoon), it will perform the trajectory in dashed yellow if following a strictly Euclidean/vector navigation strategy, whereas It will perform the trajectory in dashed green if following a strictly geodesic navigation strategy, i.e. if the mouse is using its knowledge of space to optimise for the shortest path to goal (blue node).

terms alone will lead to the emergence of "snake-like" structures (Figure 3.3, right). Although highly efficient in distinguishing between Euclidean and geodesic navigation strategies, these are very much resembling of linear track structures, which have been widely used in the study of space representation throughout the last decades and are not representative of an intricate world model for travel. Hence, we want to avoid linearised maze structures as much as wholly connected ones, aiming for a balance between structural complexity and disambiguation of vector navigation and planning.

In brief, we wanted to optimise each single maze structure to:

1. discriminate as much as possible between Euclidean and geodesic navigation strategies;

2. avoid fully connected and linearised structures;

3. avoid structures with fully connected outer edges (to encourage mice to explore the entire maze).

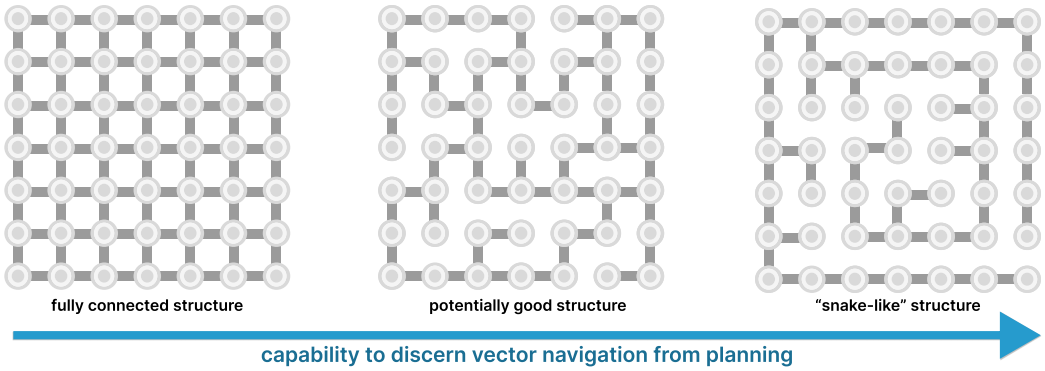


Figure 3.3: Variation in maze structure with its capacity to discriminate between vector navigation and planning Increase in discriminability from fully connected (left) to a "snake-like" structures (right).

3.3.1.1 First iteration - genetic algorithm and Pareto fronts

We started our search for adequate maze configurations by combining a genetic algorithm (Section 3.2.1.1) and the Pareto efficiency concept (Section 3.2.1.2) into a single maze structure optimiser. This optimisation process can be separated into two stages: a first stage that generates a pool of maze configurations and a second that determines the best mazes out of that pool.

In the first stage, we used the genetic algorithm to generate the initial pool of mazes. Importantly, for any initial vector structures to be valid, they needed to guarantee access to all 49 nodes of the maze. In the second optimisation stage, we wanted to identify the best structures from the total number of structures obtained

using the genetic algorithm. Structures were considered suitable if they maximised two different factors as much as possible: the euclidean-geodesic (EG) discriminability (Equation (3.1)) and the centrality factor, defined as the ratio between the mean and variance of betweenness centrality. Such multi-objective optimisation was processed through the set of Pareto efficient solutions for these factors, also called Pareto front.

$$\textit{Discriminability } (EG) = 1 - \left\| \begin{array}{c} \textit{correlation} \\ \textit{coefficient } (ED,\, GD) \end{array} \right\| \tag{3.1}$$

Although the maze structures found in the Pareto front presented the best trade-off between the EG discriminability and centrality, these did appear suboptimal given our requirements. Nonetheless, mazes outside of the Pareto front and situated within the red elliptical area depicted in Figure 3.4(a) were manually identified by their characteristics - not snake-like, not completely connected - as examples of potentially good structures (Figure 3.4(b)). The cost function (CF) in Equation (3.2) reflects an approximation to the area inside the red ellipse, where 95% of the good mazes were located. However, this was a rather complex CF, so we sought to simplify the single maze structure optimiser.

$$-8.31f^2 - 0.169fc - 0.00272c^2 + 5.25f + 0.0968c = 1 \tag{3.2}$$

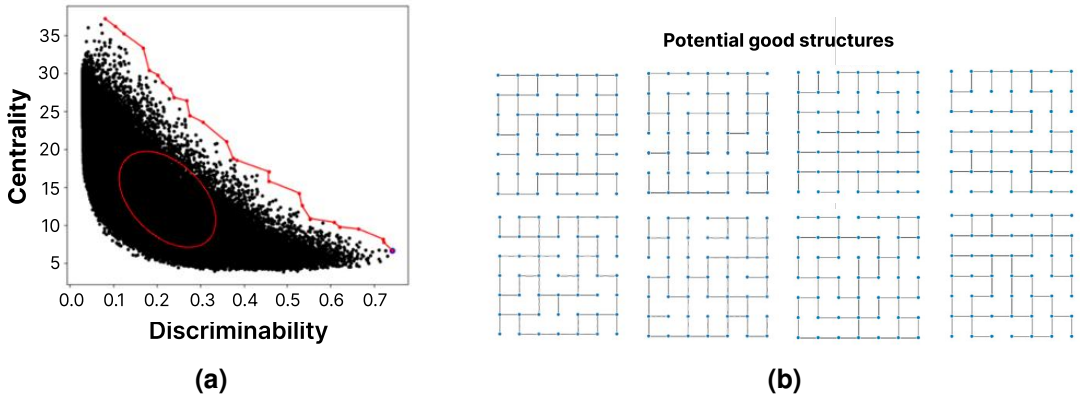


Figure 3.4: Individual maze structure optimisation - first iteration result. (a) pool of all generated maze structures using the genetic algorithm (Section 3.2.1.1), in black, and the ones resulting from Pareto multi-optimisation criteria (Section 3.2.1.2) for the discriminability and centrality factors (mazes in red line). The red ellipse contains inside it a pool of (b) potentially good maze structures. Images extracted from the custom-coded simulator.

3.3.1.2 Second iteration - exhaustive search from random maze

Another approach to our optimisation problem was to perform an exhaustive search through the space of possible mazes and to alternatively consider factors other than centrality to be maximised jointly with EG discriminability. We kept the Pareto efficiency search as a method for optimising multiple factors simultaneously.

This time, we started by (1) randomly selecting one initial structure and by (2) exploring all the possible search space of different maze structures that differed one link from the initial configuration. (3) Applying the Pareto efficiency method, we kept the structures that lay on its frontier, and with each one of those, we repeated steps (2) and (3) until the Pareto front had converged - i.e. until the structures on the two consecutive Pareto fronts were precisely the same. The new factor to trade off alongside EG discriminability was the coefficient of average Euclidean and average geodesic path lengths ($avgEUC/avgGEO$). We hypothesised that

this latter measure would, through controlling the length of possible trajectories, help avoid unit squares and long linear formations within the structures.

Implementing the optimiser described above proved computationally very heavy, as it allowed for the exponential increase in the number of maze structures that composed each new Pareto front. Additionally, the new $avgEUC/avgGEO$ factor failed to generate suitable configurations free of single unit squares or long linear substructures within the mazes (Figure 3.5).

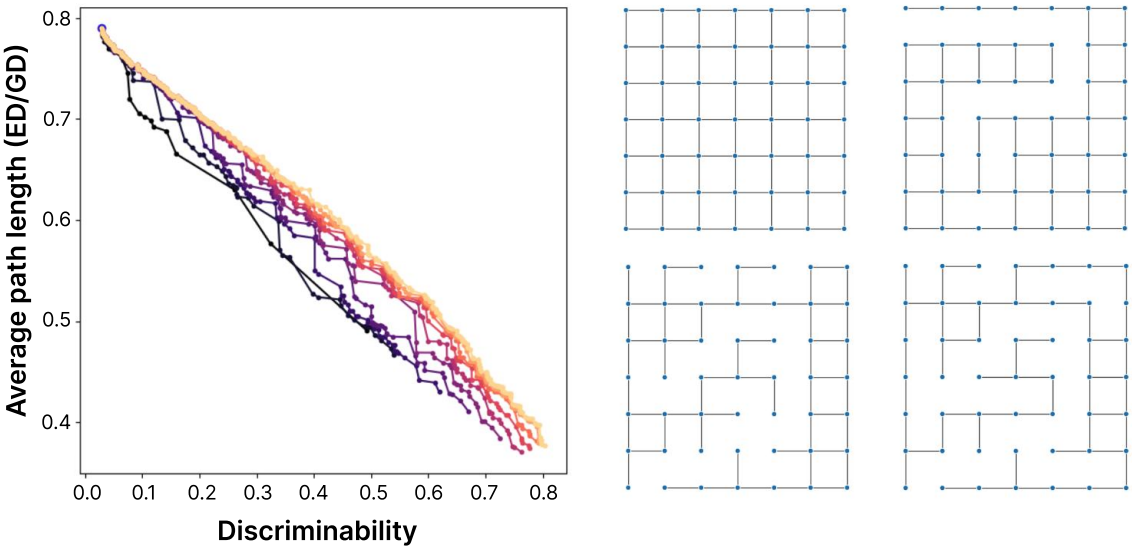


Figure 3.5: Individual maze structure optimisation - second iteration result. (a) group of maze structures in consecutive Pareto fronts optimising for the ratio of average Euclidean over average geodesic path lengths ($avgEUC/avgGEO$) and discriminability exhibiting more structures per front as the iterations progress (from black to yellow). (b) example mazes resulting from the last Pareto front, which did not satisfy the requirements. Images extracted from the custom-coded simulator.

3.3.1.3 Third iteration - exhaustive search with a maximum limit of structures per step

The next modification of the optimisation process was to keep only the 10 best structures of each Pareto front generation (Figure 3.6), therefore limiting the exponential growth in available structures at each generation, and consequently speeding up the search time. Besides retesting the first combination of factors to optimise - EG discriminability and centrality -, we also tested for a new combination of factors - EG discriminability and the coefficient of longest Euclidean and geodesic path lengths ($longEUC/longGEO$).

Both factor pairs proved inefficient at providing reasonable maze structure solutions. While the centrality factor trade-off managed to minimise the already mentioned single-unit squares problem, it resulted in mazes with an extremely linearised or “snake-like” structure, thus not complying with the complexity required for our prediction testing (Figure 3.6(a)). The $longEUC/longGEO$ factor, in its turn, returned mazes with really long linear sections, in a ”snake-like” or “tree-like” format (Figure 3.6(b)). In common, both factor pairs led to structures that were mainly completely connected on the maze’s outer edge and simultaneously exhibited multiple high-dimensional connected components (CC) (Figure 3.7).

Having a fully connected (or close to) outer edge is a highly undesirable characteristic in a maze structure since mice preferentially navigate around the edges and closest to the walls when in large open environments such as the 7x7 grid maze. Thus, to ensure full coverage of maze space by mice, we must guarantee

that this is not a feature of selected structures. Conversely, we should limit the dimension of connected components within the maze - open and closed (Figure 3.7) -, as these are proportional to the length of linear sub-trajectories within the maze, and we want to avoid highly linearised structures.

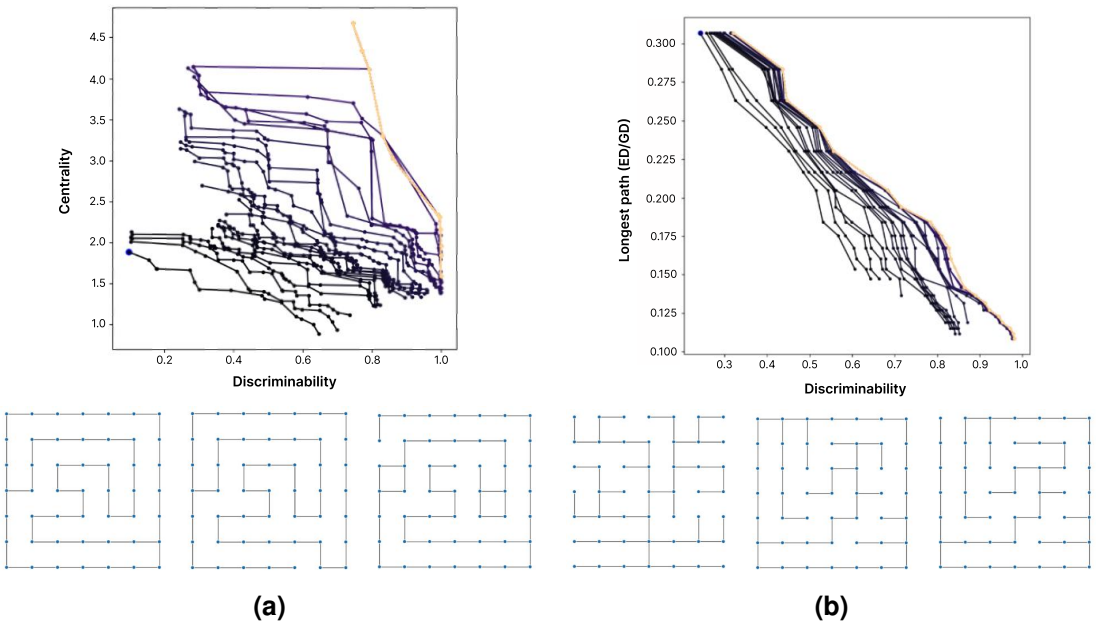


Figure 3.6: Individual maze structure optimisation - third iteration results. (top) group of successively optimised Pareto fronts for discriminability and (a) ratio of longest Euclidean over longest geodesic path ($longEUC/longGEO$) or (b) centrality factor, and resulting non-satisfactory example maze structures (bottom) found in the last Pareto front (yellow). Images extracted from the custom-coded simulator.

3.3.1.4 Fourth and final iteration - maximising EG discriminability and simple rules

The learning attained through the iterations described from Section 3.3.1.1 to Section 3.3.1.3 thoroughly informed the final version of the single maze structure optimiser.

Finally, we could obtain adequate structures by maximising a simple term - the

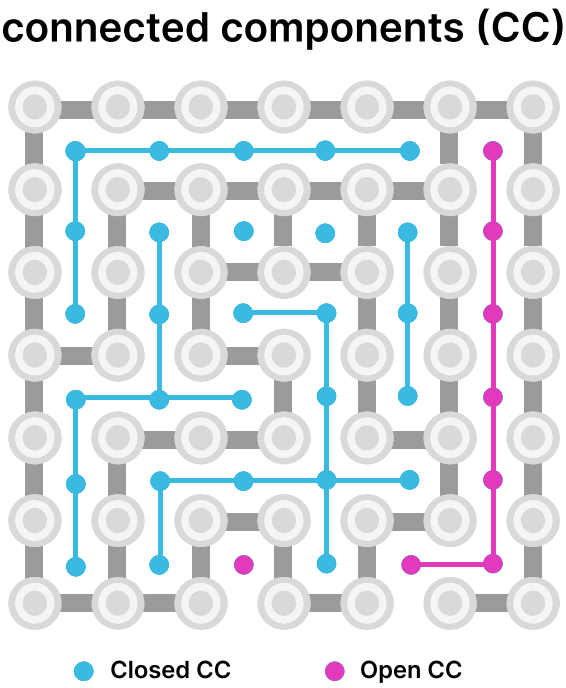


Figure 3.7: Illustration of the open (pink) and closed (blue) connected components concept.

EG discriminability - and adding two intuitive rules relative to the outer edges of the maze and the connected components within it. We decided to set the maximum number of connected outer edges to 75% (18) of the total possible number of connected outer edges (24), to discourage mice from exploring solely along the walls. Limits to the length of open and closed CC (Figure 3.7) were also enforced to avoid long linear substructures inside the maze and allow the structure to become more intricate and complex. The open CC maximum length was 4 and the closed CC maximum length was 8. These values were defined based on trial and error and observation of multiple other values.

3.3.2 Sequence of mazes optimiser

Since we wanted to study the formation and update of hippocampal representations in complex and dynamic environments, besides guaranteeing that we designed adequate structures, we had to carefully decide where the changes in structure would be introduced over time. To achieve the latter goal, we built a sequence of mazes optimiser.

In the case of maze sequences, besides guaranteeing the quality of each individual maze structure, as described for the single maze optimiser (Section 3.3.1), we wanted to select the changing link at each step such that:

1. it maximised the difference in possible geodesic trajectories between consecutive structures;

2. it was not the same link that had been changed in the previous step, or a link in close structural proximity to the previous one, to avoid the repetition of identical or similar structures.

3.3.2.1 Building a sequence of maze structures

The cost function

The sequences defined with this algorithm were thought to be composed of good individual maze structures only - following the rules explained in Section 3.3.1.4 - and to introduce one single change at a time - to add or remove one link - in the location where it would generate a higher global impact on the maze structure. That is, the change should be introduced in the location that would maximally decor-

relate all the possible trajectories between that structure and the previous one, following a geodesic navigation strategy. We called this between-maze discriminability (BD) and we had to maximise the CF in Equation (3.3) to generate good sequences of mazes. The other term of this cost function was the EG discriminability mentioned in the single maze structure optimisation process (Section 3.3.1), here renamed to within-maze discriminability (WD) for coherence purposes.

The basis for creating an adequate sequence of maze structures started with (1) choosing one good initial maze using the single structure optimiser described in Section 3.3.1.4 and then (2) doing an exhaustive search through the space of all possible single link changes (addition and removal) to that structure, (3) calculating the cost function value for each of them (Equation (3.3)) and (4) selecting the structure with the highest score. However, maximising this cost function did not suffice to find the best sequences, as some concerning algorithmic behaviours arose, which forced us to implement a couple of intuitive rules to the process (similarly to what had happened with the single maze structure optimiser, see Section 3.3.1.4).

$$\textit{Cost Function} = Weight * WD \pm (1 - Weight) * BD$$

where

$$WD = 1 - \left\| \begin{array}{c} \textit{correlation} \\ \textit{coefficient } (ED,\, GD) \end{array} \right\| \tag{3.3}$$

$$BD = 1 - \left\| \begin{array}{c} \textit{correlation} \\ \textit{coefficient } (GD_{i+1},\, GD_i) \end{array} \right\|$$

Extra rules

It is easy to imagine that, if removing one link in the structure represents the highest value in cost function, then adding the same link in the next step of the sequence might again maximise our measure of reference. Thus, to avoid consecutive changes of the same link (Figure 3.8), we introduced a freezing rule prohibiting the change of a link that has just been added or removed in the next 10 consecutive steps of the optimisation process. Interestingly, implementing this rule led to a new undesired algorithmic behaviour of mainly changing links in close structural proximity to each other (Figure 3.9(a)). Keeping this behaviour would result in fairly similar consecutive structures, impeding the study of structural evolution over time. To prevent this, we added a second rule that allowed changes to happen only if at a minimum of 3 steps away from the previously changed link (Figure 3.9(b)).

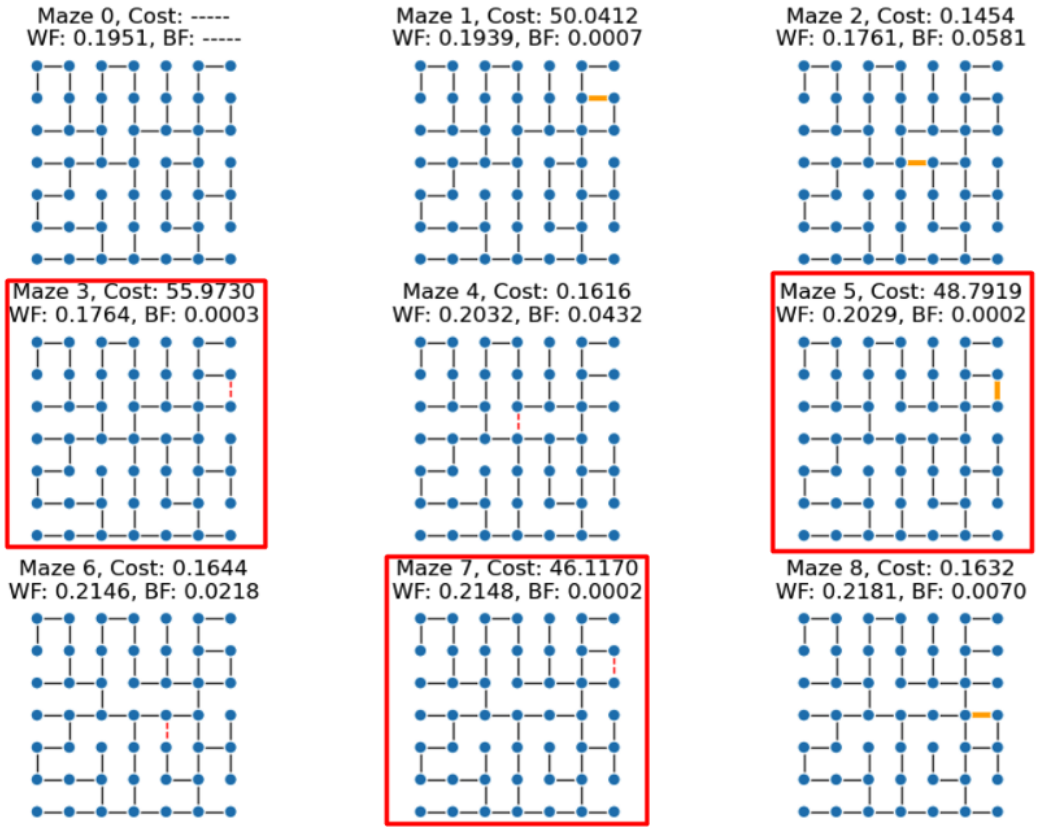
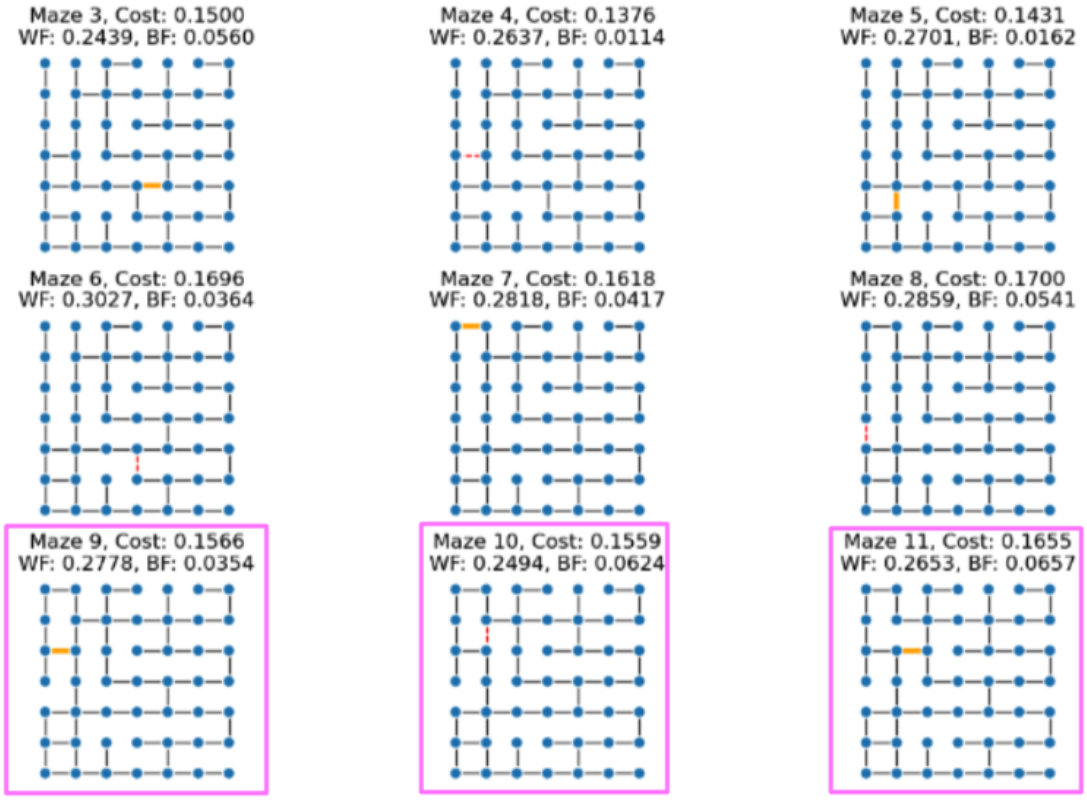
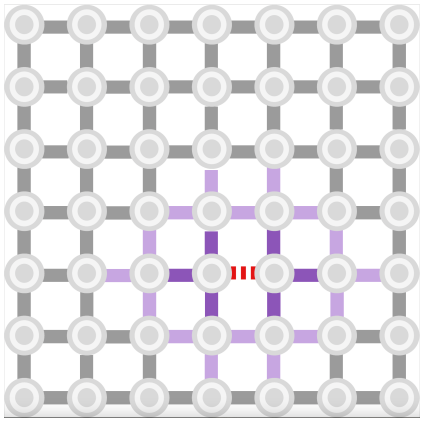


Figure 3.8: Additional rules for optimising a sequence of maze structures (I). A change-freeze of 10 modifications was introduced to avoid the same link from changing in consecutive or close to consecutive modifications of the structure.



(a)



(b)

Figure 3.9: Additional rules for optimising a sequence of maze structures (II). (a) To avoid the consecutive link changes to take place in the vicinity of each other, (b) a security distance was introduced that only allowed the change of links that were further than 3 steps away from the previous changed link.

3.3.2.2 The chosen sequence

Using the sequence of mazes optimiser described in Section 3.3.2.1, we were able to choose the testing sequence (Figure 3.10 to Figure 3.13). The sequence combined high values of WD and BD for each individual structure and link-change, respectively, and respected all the extra rules implemented in Section 3.3.1.4. This sequence was used in its entirety by the two cohorts of mice ($n = 5$ plus $n = 5$) that were implanted in dCA1 for the collection of electrophysiology data; it was used up until the 21st maze structure to test the behaviour-only cohort of mice ($n = 8$).

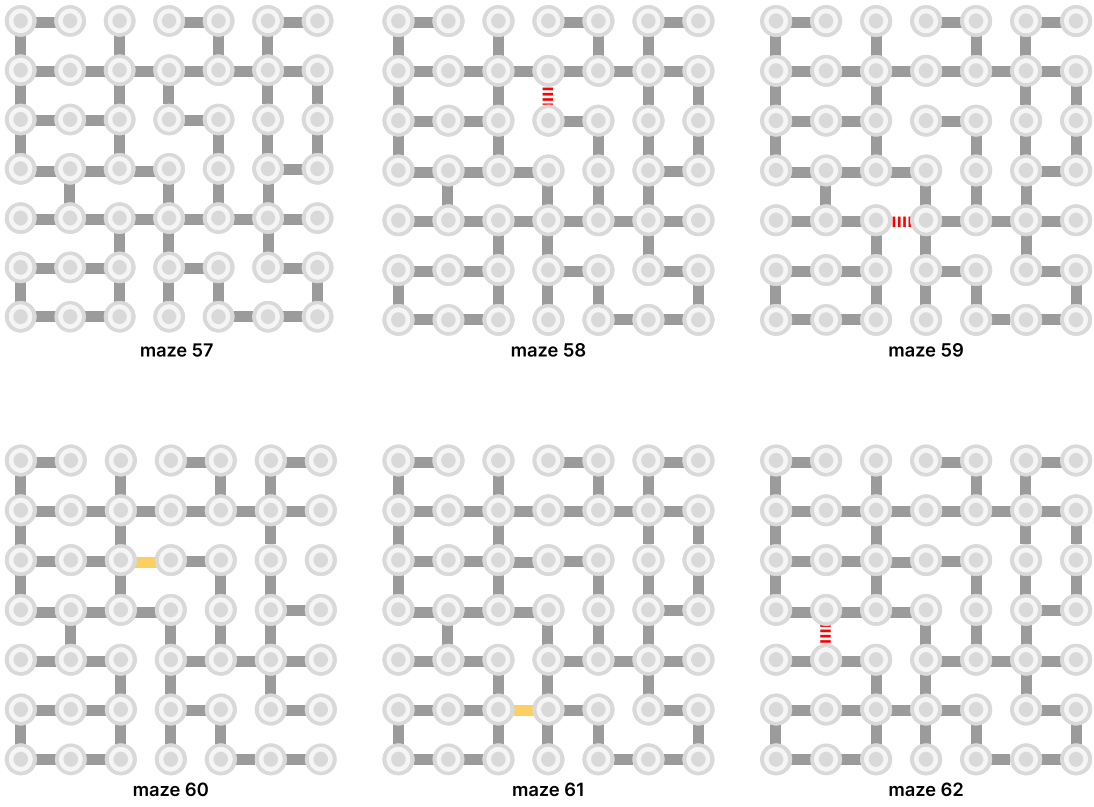


Figure 3.10: Chosen sequence of maze structures used in this project's experiments (I) From maze 57 to 62, inclusive.

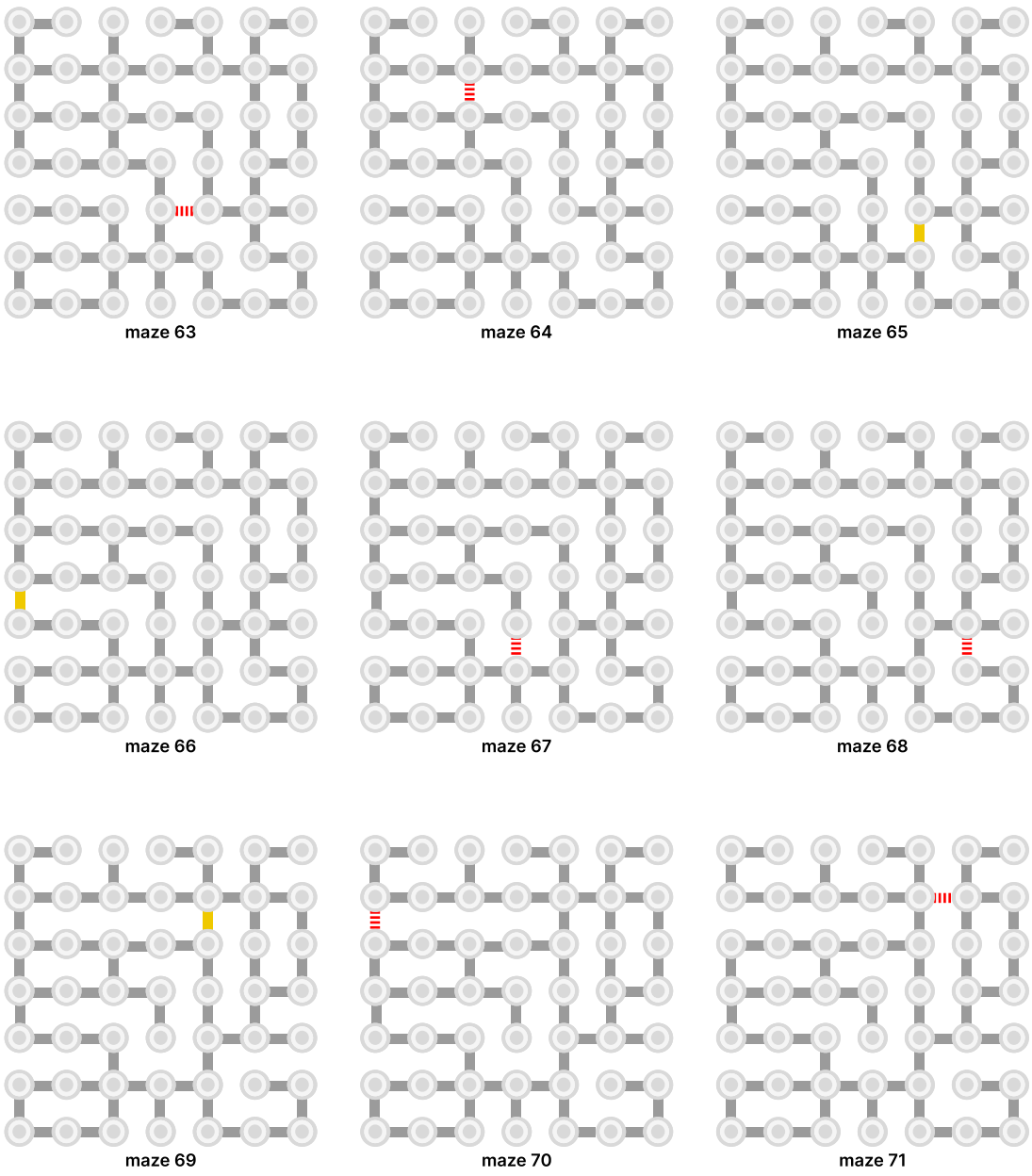


Figure 3.11: Chosen sequence of maze structures used in this project's experiments (II). From maze 63 to 71, inclusive.

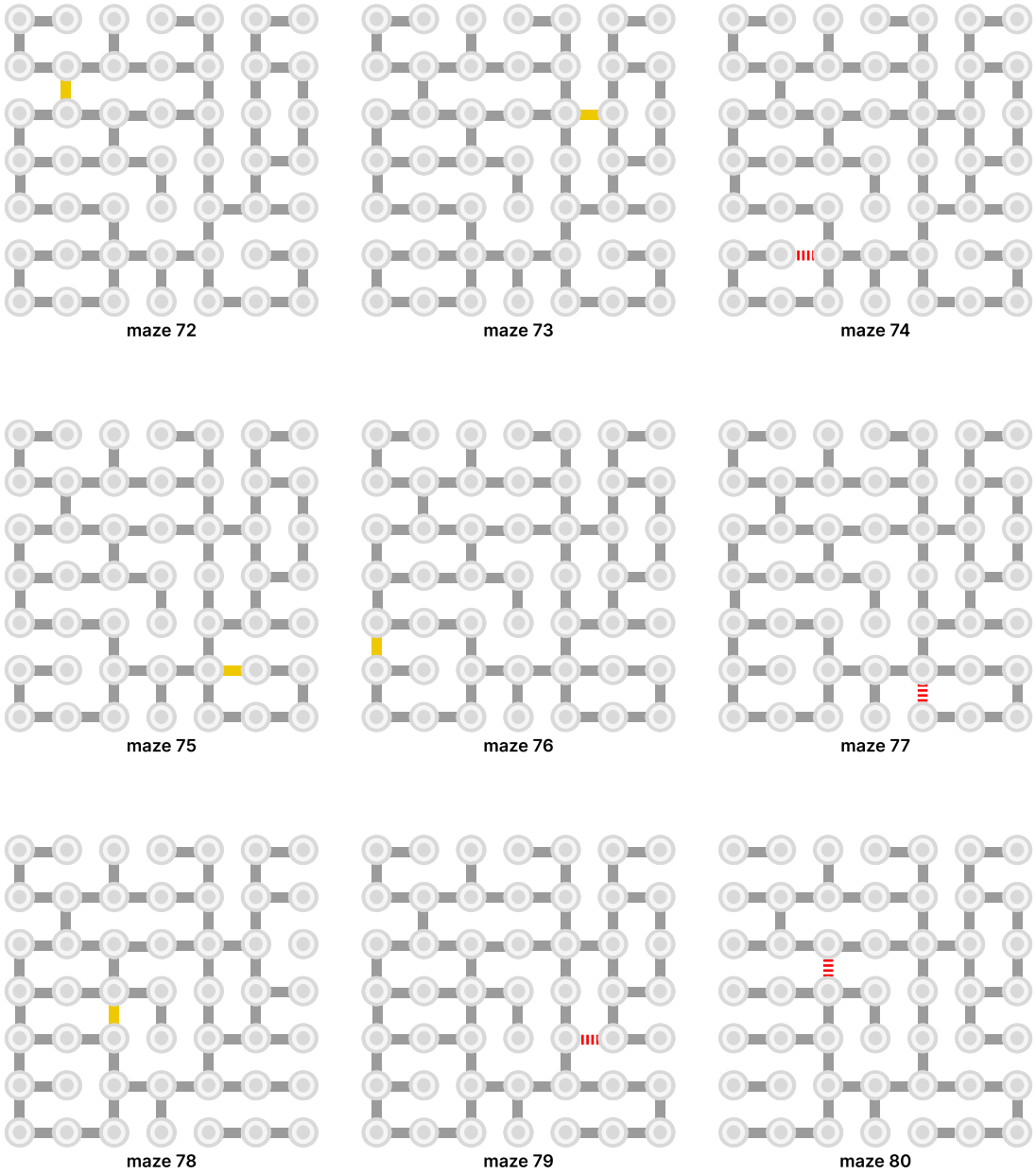


Figure 3.12: Chosen sequence of maze structures used in this project's experiments (III). From maze 72 to 80, inclusive.

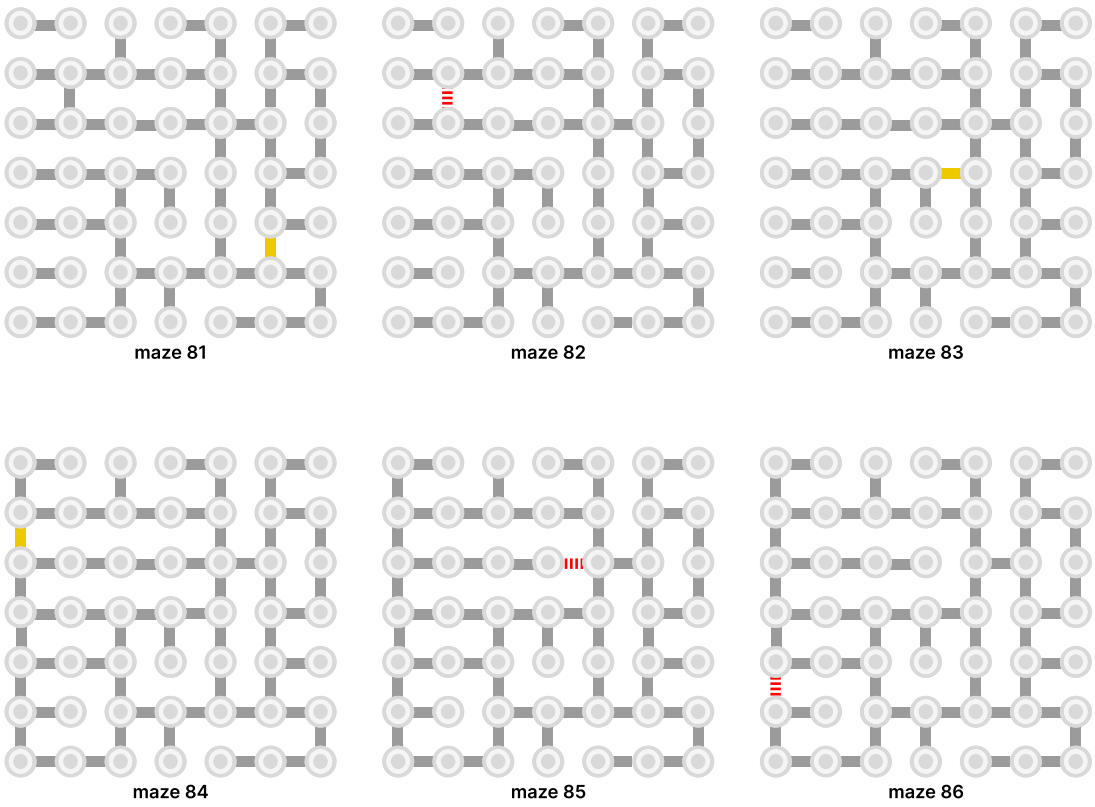


Figure 3.13: Chosen sequence of maze structures used in this project's experiments (IV). From maze 81 to 86, inclusive.

Mice experienced the following maze structures:

- **Behavioural cohort:** mazes 57 to 75.
- **First recordings cohort:** mazes 57 to 85.
- **Second recordings cohort:** mazes 57 to 86.

Section 3.4

Discussion

In this chapter we presented the iteration process for creating a sequence of maze structures optimiser that allowed us to run a goal-directed navigation task in a 7x7 grid maze (maze and task in Sections 2.2.1 and 2.3). The first stage was the optimisation of the individual maze structures, which should allow for clear discrimination between Euclidean and geodesic strategies of navigation.

Attempts to build this optimiser included variations of the algorithms used for generating intermediate state structures and of the number of maze structures considered at the Pareto front; and (re)testing the EG discrimination, centrality, average- and longest-EG path coefficient factors. Regarding the number of structures considered at the Pareto frontier step, no benefit was found for considering more than 10 structures - considering fewer structures and increasing the number of generations of the optimisation process was often preferable, as it rendered the process less computationally heavy while offering a sufficiently broad pool of results; considering more structures and fewer generations decreased the desired variability while increasing the computation time to an unpractical time window. As for the factors considered for trade-off, none of the combinations mentioned originated satisfactory configurations, as these were often too connected or too linearised. Whereas the over-linearised structures would not allow the collection of

rich behavioural data, the partially fully connected ones would prevent the discrimination of the model-based and model-free components of the animal's navigation strategy.

The winning combination for generating individual maze structures consisted of maximising the EG discriminability factor with a few additional intuitive rules - a limit to the number of edges in the maze, a limit on the number of total possible outer edges in the maze and a limit on the length of the closed and open connected components in the structure. As the first rule prevented the maze from becoming too connected, the second was imperative to incentivise mice to equally explore the entire maze. Mice are prey animals and previous studies have reported their preference for navigating along the walls of big open environments. Since we want to track and study intricate navigation behaviours, we want to avoid as much as possible to bias the animal trajectories in any way - in this case we want to avoid them navigating along the walls as much as possible to reach any position in the maze. The connected components rule was successfully introduced to avoid long linearised substructures within the maze that the animals would have to travel in its totality before being able to drop off onto new trajectories.

We reused these rules for simulating sequences of maze structures that differed by one link from each other. In this case, the cost function considered had two terms - the WD (previously known as EG discriminability) and the BD, that decorrelated the difference in geodesic trajectories between every two consecutive structures. Besides maximising this CF, a new set of simple rules had to be implemented - a link-freezing rule and a minimum interlink distance for consec-

utive changes. Respectively, these rules prevented consecutive changes to consecutively target the same link or neighbouring links. Repeated changes in the same area of the maze would make it difficult to assign responsibility for any local changes in representation and could ultimately result in mice building a two-fold template of the environment, eliminating the need to update space representation to instead infer in which of the states they were.

In sum, the sequence of maze structures simulator allowed the execution of this project by supporting the choice of adequate world configurations for mice to navigate, which thus enabled the quantification of the animals' choices in the maze and classification of their behavioural strategy (Chapter 4 for details). This discrimination property can, in principle, allow for the testing of quantitative predictions of hippocampal cell representations such as the ones generated in the context of the successor representation framework (Stachenfeld et al., 2017). Moreover, the individual maze optimiser described in this chapter has already been used by another student in the lab to optimise new structures for their experiment, and will hopefully be a useful tool for others to come.

4

Mice quickly integrate new spatial information into their navigation strategy under a goal-directed paradigm - and other behavioural signatures

Section 4.1

Introduction

Navigating through complex and dynamic environments requires animals to employ flexible and adaptive behaviours to reach their goals efficiently. Contrarily to the case of simple and deterministic spaces, complexity requires subjects to rely on their cognitive abilities to plan and execute optimal routes based on their internal representation of the environment and the expected outcomes of their actions. This goal-directed navigation behaviour is closely tied to the concept of cognitive maps (Tolman, 1948a) and the distinction between MB and MF learning (Daw et al., 2005).

Tolman's theory of cognitive maps suggests that animals form comprehensive mental representations of their environment, which allows them to navigate flexibly and generalise their experience in the face of novelty (Tolman, 1948a). If so, then flexibility relies on the integration of spatial and non-spatial information, such as the location of rewards and the relationships between different maze sections (Eichenbaum and Cohen, 2014).

In the light of RL, a MF behaviour fails to adapt to change due to the lack of an internal representation of the environment and its relations; whereas a MB behaviour makes use of a detailed representation of space and the relationship between its states, actions and outcomes. As the latter aligns better with the cog-

nitive map theory, providing adaptability; the former can be useful in situations of habit, of overexposure to the environment, of static contingencies, or when one can perform accurately by solely relying on sensory inputs (e.g. vision, sound). In the context of navigation, behaviour can be interpreted as a sequence of choices where, at every decision point, one must weigh the benefits of following an accurate but often laborious MB strategy or employing a quick but often error-prone MF tactic. The answer seems to be using a hybrid system (Daw et al., 2011; Hasz and Redish, 2018).

In this chapter, we investigate the behavioural features of goal-directed navigation of a complex maze in mice. After explaining the training stage animals experience in a small 3x3 grid maze, we look for behavioural signatures of these animals' performance while navigating the large 7x7 maze, such as reward distribution and speed. Using a mixture of strategies model, we quantitatively assess mice navigation tactics, looking to understand the contributions of vector and shortest path strategies for efficient navigation of the space. Given the intricate structure and the random allocation of rewards across trials, we predict it will be beneficial for the animals' performance to create an internal model of space instead of simply following the head direction to the goal. Furthermore, subjects experience single changes to the maze structure daily, with one bridge being added or removed according to the chosen sequence (Section 3.3.2.2). We thus wish to comprehend how these changes affect their behaviour, and how long it takes for mice to adapt to the new spatial contingencies.

Section 4.2

Methods

4.2.1 3x3 grid maze training schedule

All mice were trained in a 3x3 elevated grid maze (Section 2.2.2) prior to exposure and testing on the 7x7 grid maze (Section 2.2.1). This training period was essential for the animals to (1) habituate to the setup, (2) create an association between light/sound cues and reward and (3) reach a good overall performance, reflected by an increase in the number of trials completed and a decrease in the number of pokes in unrewarded ports (Section 4.3.1).

For the behavioural cohort ($n = 8$) and the first dCA1 probe-implanted cohort ($n = 5$), the training schedule consisted of two 30-minute sessions per day, for 18-21 days. Mice were initially exposed to the entire 3x3 maze in a fully connected structure on the first day, followed by a gradual expansion from 3 connected nodes to 5, 7 and finally 9 connected nodes (not in a fully connected form), at a subject-dependent rhythm (Figure 4.1).

After 18 days, subjects in the behavioural cohort were transferred onto the 7x7 grid maze for testing. The first cohort of dCA1 probe-implanted mice suspended training after 21 days for surgery purposes and resumed it for an extra 15 days after recovery.

We then optimised this training period for the second cohort of dCA1 probe-

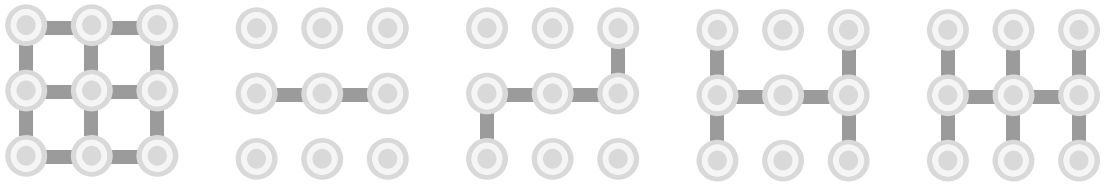


Figure 4.1: Initial 3x3 maze training schedule. Mice from the behavioural ($n = 8$) and first recordings ($n = 5$) cohorts were trained following this schedule. They started by experiencing one day of the fully connected maze, progressing onto the 3-node line and continuing at variable rates through the other configurations, in the order depicted in the figure, from left to right.

implanted mice ($n = 5$) by modifying the sequence of maze structures the animals experienced (Figure 4.2). Firstly, we eliminated the fully connected maze experience, as it highly biased the animals to explore the maze preferentially relying on the outer edges, as these are closer to the walls of the setup. Secondly, we eliminated the 3-node stage of the training, as it required the mice to more constantly turn around on themselves due to the shortage of available options - which they preferably avoid doing - and did not seem to bring any advantage compared to starting in a 5-node structure. Each mouse experienced two 30-minute sessions a day.

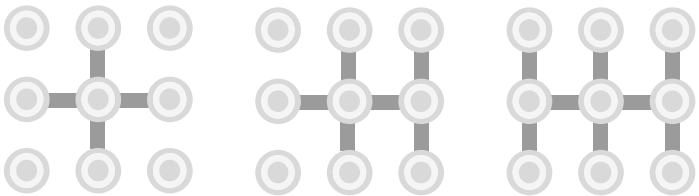


Figure 4.2: Optimised 3x3 maze training schedule. The second recordings cohort ($n = 5$) followed this training schedule, experiencing at first 5 connected nodes, then 7 and finally 9, but never in a fully connected configuration.

4.2.2 Mixture of strategies model

For quantifying the influence of planning and vector (or Euclidean) navigation on the subjects' strategy we fitted a mixture of strategies model to their behaviour.

$$\begin{aligned} Q_{net}(a_i, s_t, g_t) = \alpha \cdot Q_{\text{vector nav.}}(a_i, s_t, g_t) \\ + \beta \cdot Q_{planning}(a_i, s_t, g_t) \end{aligned} \tag{4.1}$$

The Q_{net} comprises a weighted mixture of two components - the *vector navigation component* Q_{vec} and the *planning component* (shortest path) Q_{pla}. The Q_{vec} reflects the vector navigation rule of preferring the option with the smaller angular distance to the goal location (Equation (4.2)), taking value of the $\cos(\theta)$, where theta (θ) is the angle of the animal to the goal at a given decision point. The Q_{pla} assumes the mice have a cognitive map of the environment and can be approximated by evaluating the animals' choices according to the shortest path navigation rule (Equation (4.3)), taking values $+1$, -1 or 0 if the choice at the decision point brought the animal closer, further or at an equal distance to the goal, respectively.

$$Q_{\text{vector nav.}}(a_i, s_t, g_t) = \cos(\theta) \quad , \quad \theta = \angle a_i, g_t, s_t \tag{4.2}$$

$$Q_{planning}(a_i, s_t, g_t) = \begin{cases} +1 & \text{if } a_i \text{ is on the shortest path between } s_t \text{ and } g_t \\ 0 & \text{if } a_i \text{ is on a path with equal distance} \\ -1 & \text{otherwise} \end{cases} \tag{4.3}$$

The choices of the mice are then analysed by maximum likelihood fitting, and the

probability of each action is defined as in Equation (4.4).

$$P(a_i|s_t, g_t) = \textit{softmax}(Q_{mix}(a_i, s_t, g_t)) \tag{4.4}$$

where softmax denotes a softmax function with fixed temperature parameter $\tau = 1$. This temperature value allows interpretation of the coefficients obtained by maximum likelihood fitting of mice behaviour as indicators of the extent to which the respective component explains the observed behaviour.

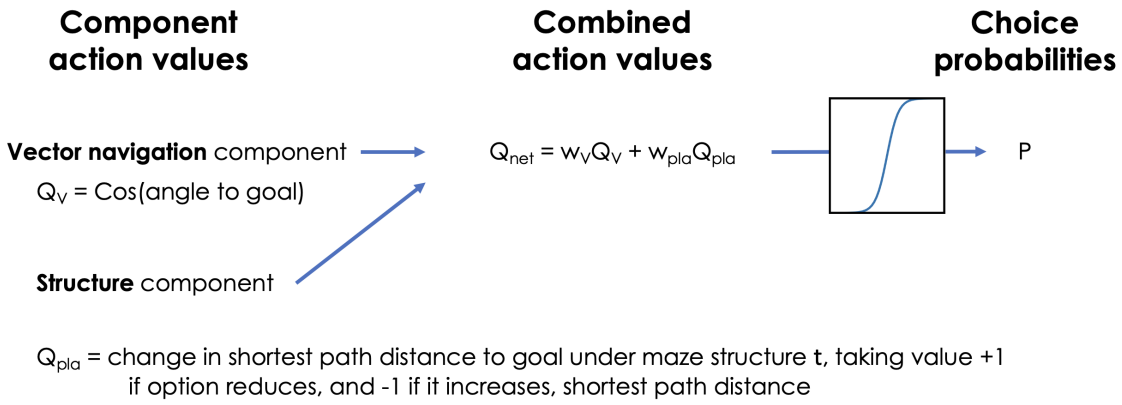


Figure 4.3: Vector and planning strategies model cartoon.

Analogously, we used the same mixture of strategies model to assess how quickly new information was incorporated into mice's strategy following changes in the maze structure. This was done by adding a Δ *structure component* $Q_{\Delta S}$, which reflected the difference in values between the current and the previous structures, i.e. the change in values due to the single link added or removed immediately before the current session. In this case, the fitting was applied to data from the first session after each change.

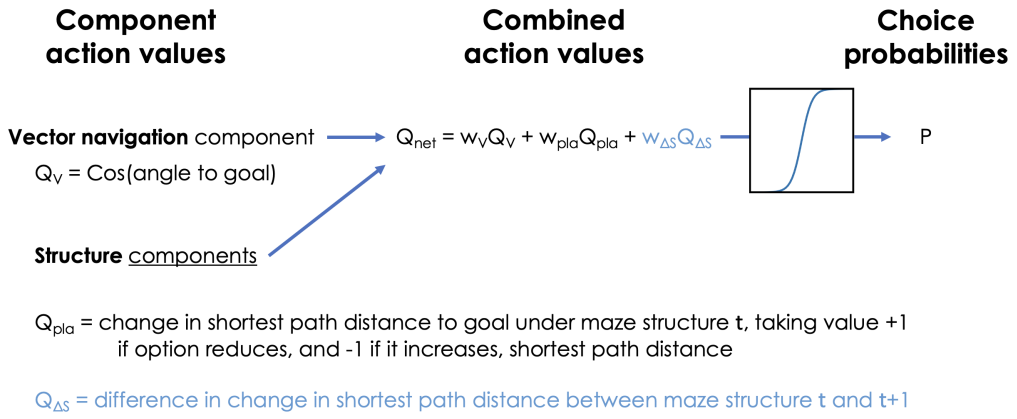


Figure 4.4: Vector, planning and structure change integration strategies model cartoon.

Section 4.3

Results

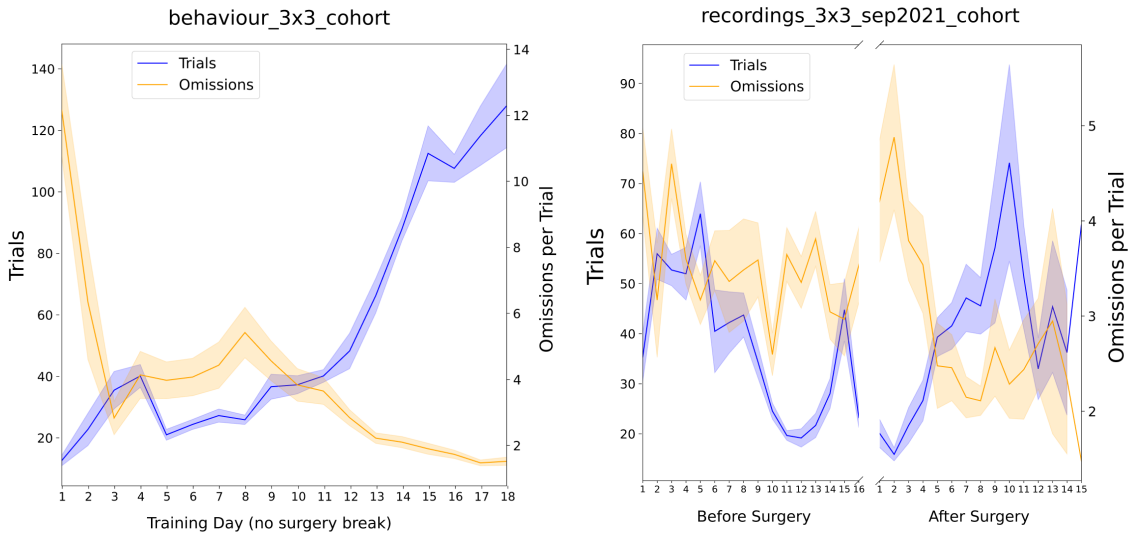
For this section's results, we will consider data from the three different cohorts of mice (total $n = 18$) that ran the 7x7 grid maze: 8 mice that ran behaviour without implants and 2 distinct cohorts of 5 mice each that were implanted in dCA1 and ran the goal-directed navigation task tethered.

4.3.1 Mice training in a 3x3 grid maze

Before every experiment in the 7x7 grid maze, mice are trained in a smaller (3x3) version of it (details in Section 2.2.2). We called this the training stage, and it was done with the purpose of (1) habituating the animal to the setup, (2) teaching mice to associate the light cue with the reward location, and (3) getting subjects to

perform at a satisfactory level, i.e. performing $>$1 trial/minute and minimising the number of pokes on non-rewarded ports.

Training duration and schedule were explained in Section 4.2.1 of this chapter. The first training schedule was indeed less efficient than the second one - cohorts of behavioural and first recorded animals exhibited a learning stage twice as long (Figure 4.5) as the animals from the second recordings cohort, which ran the improved training schedule on the 3x3 maze (Figure 4.6). Besides showing the number of trials during the training stage (before and after surgery when applicable), we establish a measure of omissions. These correspond to the number of pokes in unrewarded ports that the animal makes in each trial until it reaches the reward. In the early stages of training, before the animals grasp the meaning of the cue and, therefore, the task, they will often poke every single port on their trajectories until they collect the reward. This number of unrewarded pokes drastically decreases once the animal understands the cued goal-directed paradigm, while the number of trials per session increases. After surgery, animals return to the training setup for a quick recapitulation of the task, and habituation period with the implants (Figure 4.5(b) and Figure 4.6). For the second cohort (Figure 4.6), subjects' performance was significantly impaired due to construction works happening in the building during training.



(a) (b)

Figure 4.5: Training performance on the 3x3 grid maze - behaviour and first recordings cohorts. Variation of the number of trials (blue) and omissions (orange) performed by the mice on the (a) behavioural and (b) first recordings cohorts across training days, before and after surgery, when applicable. The training protocol can be observed in Figure 4.1.

4.3.2 General behavioural signatures of mice navigation in a 7x7 grid maze

After finishing the training stage in the 3x3 grid maze, mice ($n = 18$) navigated a goal-directed task on a sequence of predefined maze structures that changed by one bridge every day (Section 3.3.2.2).

In total, we collected data from $37,573$ trials throughout 979 sessions. Subjects navigated to maximise the rewards collected in each session (1 reward/trial), performing a total of 2087.39±686.78 (mean±SD across subjects) trials. Per session, the number of trials varied between 1 and 90 (1 for one animal on the first day in the 7x7 maze), with each mouse performing 38.38±17.45 (mean±SD across

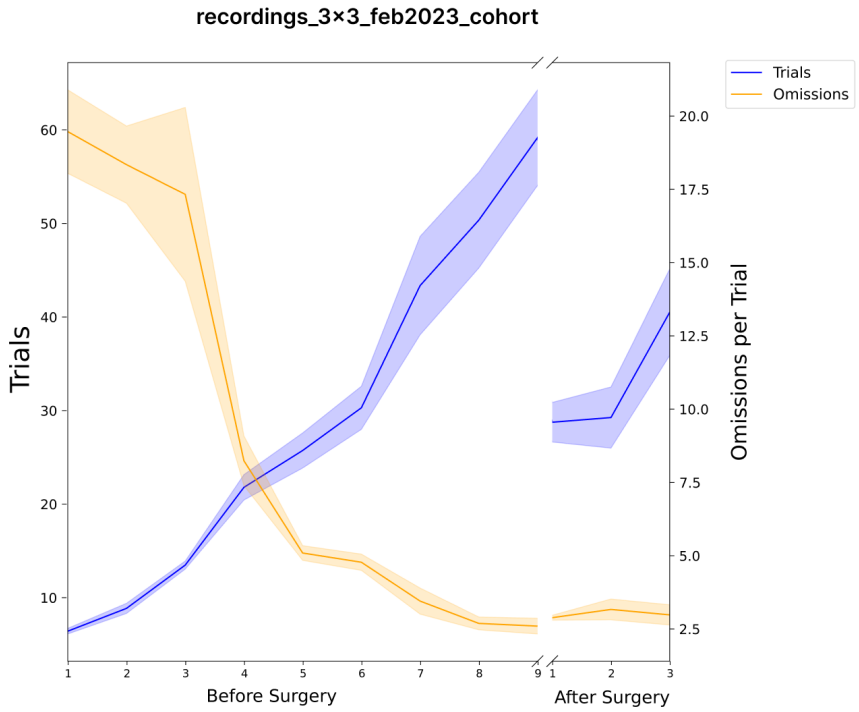
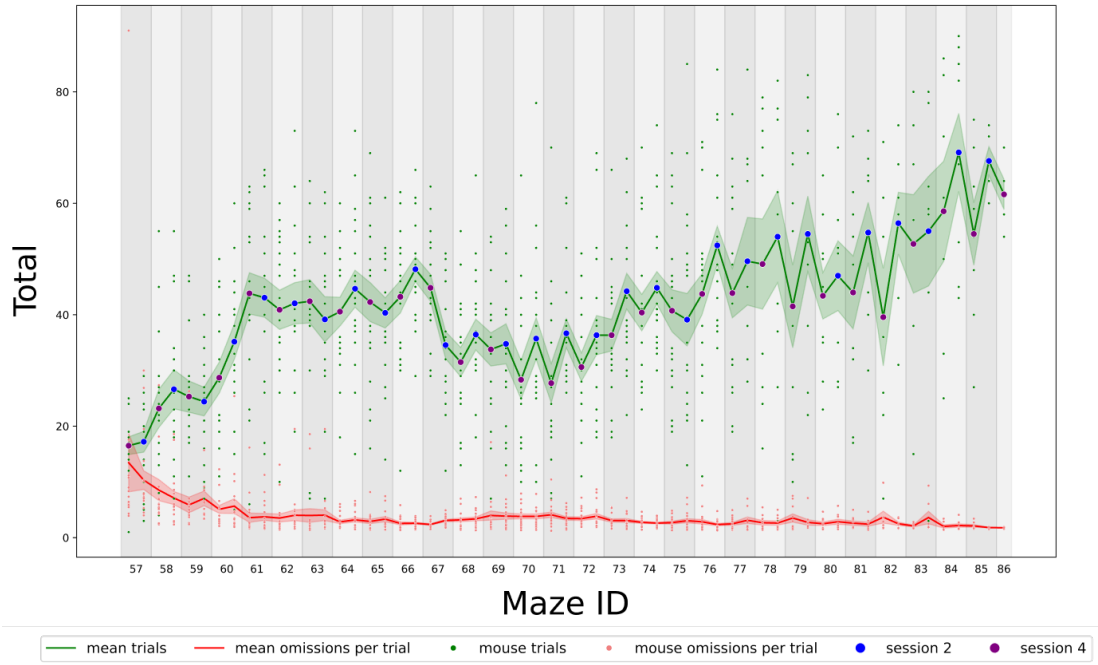


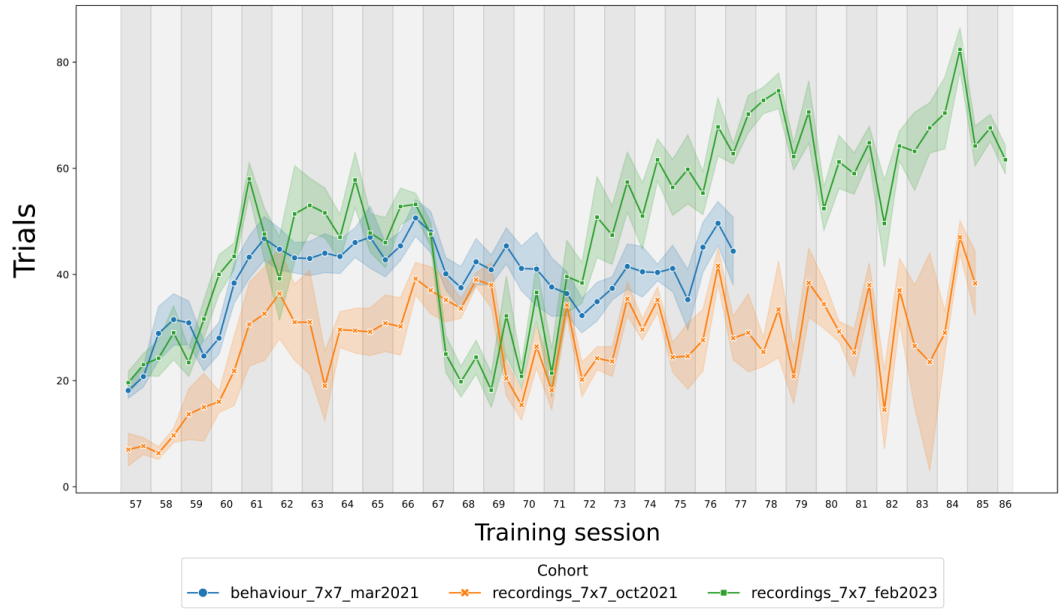
Figure 4.6: Training performance on the 3x3 grid maze - second recordings cohort. Variation of the number of trials and omissions performed by the mice on the second recordings cohort across training days, before and after surgery. The training protocol can be observed in Figure 4.2.

sessions) trials, in a total of 54.39±5.75 sessions (mean±SD across subjects).

Observation of the animals shows that, for the first couple of days after being transferred to the 7x7 maze, they are not necessarily engaged with the task for most of the session, preferring to explore the entire space rather than maximising the number of rewards. However, three to four days suffice for the mice to resume goal-directed navigation behaviour (Figure 4.7). Such is reflected in a gradual increase in the number of trials concluded per session up to a first plateau of around 45 trials/session, and by a consistent decrease in the number of omissions per trial, a behavioural feature more dominant in the 3x3 maze before mice grasp the task. This reflects quick generalisation from the 3x3 space.



(a)



(b)

Figure 4.7: Overall performance on the 7x7 grid maze. (a) representation of mice number of trials (green dots) and mean±SEM across animals per session (green line and shade), with the indication of sessions 2 (blue dot) and 4 (purple dot) overlaid; representation of mice number of omissions (i.e. unrewarded pokes performed on the way to the goal) (red dots) and mean±SEM across animals per session (red line and shade). (b) representation of the number of trials (mean±SEM across animals) across sessions by cohort - first and second recordings cohorts (orange and green, respectively), which experienced 29 and 30 maze structures, respectively; and the behavioural cohort (blue), which experienced 21 maze structures in total.

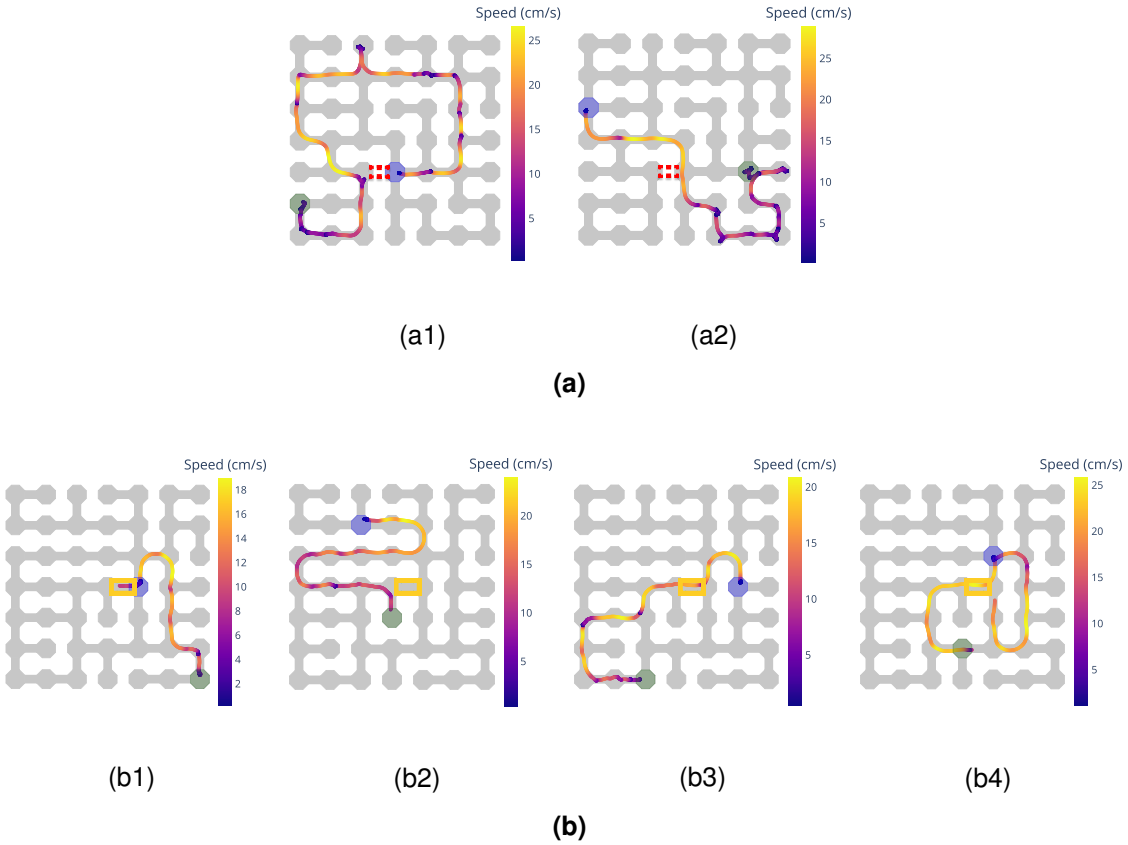


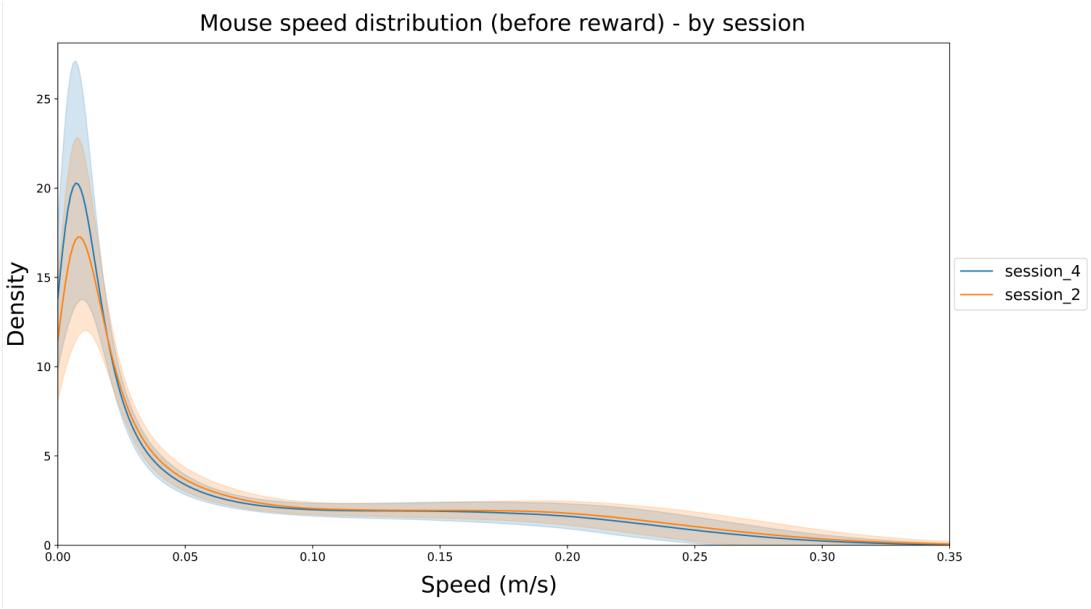
Figure 4.8: Behavioural trials in the 7x7 grid maze. The plots presented show examples of navigation trials, representing the mouse velocity (increase from dark to light colours) throughout its trajectory from the start (green node) to the goal (blue node) locations. (a) shows example mouse MB11 within the first session after the bridge in red had been removed. In trial 8/24 (a1), the animal exhibits a sudden velocity reduction when approaching node C3, perhaps indicating an attempt to use the bridge on the way to the goal, and having to find an alternative route instead. In trial 14/24 (a2), the mouse finds its way to the goal passing by a node adjacent to the removed bridge, but without movement deceleration. (b) shows example mouse MR33 within the first session after the bridge in yellow had been added. The mouse does not integrate the added link (trial 14/57) (b2) until later in the session (trial 28/57) (b3), even though it visited the link early in trial 7/57 (b1). In the subsequent session (b4), the mouse traverses the newly added link (trial 13/57). More examples in Appendix A.

In Figure 4.8, we can observe example trials of mice running in the maze in the first and second sessions after removing (top row) or adding (bottom row) a bridge. We can observe a drastic deceleration of the animal's speed when approaching the location of the removed bridge in the first session after removal (trial

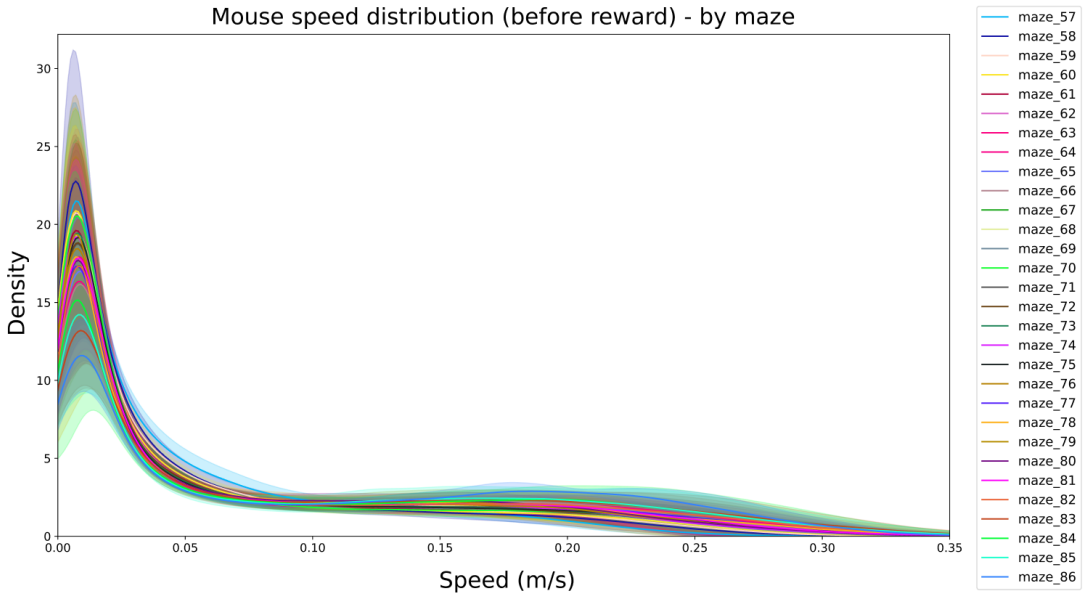
8, Section 4.3.2), when the bridge was part of a direct trajectory from the starting point (in green) to the goal (in blue). In the second session with the same maze structure, the animal is able to perform through alternative ways around the bridge without speed reduction (Section 4.3.2). When a bridge is added to the maze structure, the mice can take a long time to notice it, as the paths they preferred to take before remain unaltered (Section 4.3.2); they might start incorporating it in their trajectories halfway through the session (Section 4.3.2, trial 28/57), but will traverse them more often in the subsequent session (Section 4.3.2).

Overall, we observe a tendency for mice to navigate at higher speed in the second sessions after the change, compared to the sessions when the bridges are introduced/added (Figure 4.9(a)), albeit non-significant. Moreover, they also seem to exhibit continuously higher velocities across the experimental days, i.e. across maze structures (Figure 4.9(b)).

One of the unique characteristics of this task is the random allocation of reward in every trial. This means that, in each trial, the goal location is surprising for the animal, and thus cannot be inferred based on prior information - e.g. two consecutive rewards can be cued in opposite corners of the maze, and so can they be cued on the exact same platform (Section 4.3.2). Over time, and across all animals, the reward distribution appears very even throughout the maze, as expected (Figure 4.10(b)).

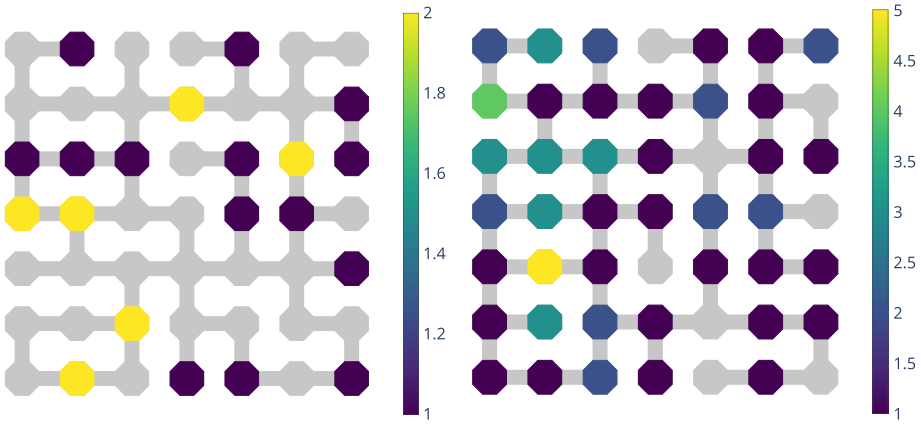


(a)

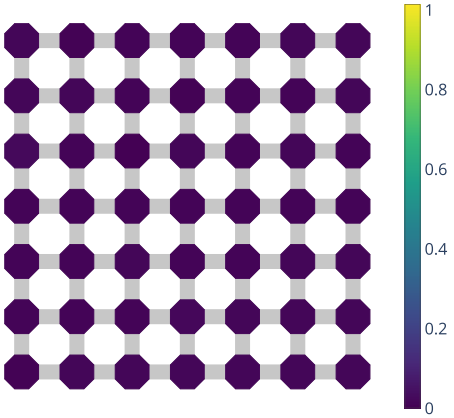


(b)

Figure 4.9: Variation of mouse velocity across sessions. Velocity is calculated for the period between trial start and poke-in at reward location (when all mice are invariable stationary consuming reward and when their movements are not related to the task, as they don't have the information of the next goal location yet). Thus, these variations in speed reflect solely the running velocity during active trial periods and moments of stillness due to adjustment of routes and/or disengagement. (a) mouse velocity tends to increase in sessions 2 compared to sessions 4, which correspond to the second and first time the animals' experience the same maze structure. (b) mouse velocity tends to increase throughout the experiment, i.e. across maze structures.



(a)



(b)

Figure 4.10: Reward distribution across the 7x7 maze structure. From dark to lighter colours there is an increase in the quantity of observed reward distribution. (a) example reward distribution across space of two sessions (MB14, maze 58 and MR33, maze 79); (b) overall reward distribution ratio for all nodes used across all the sessions.

Lastly, to investigate mice trajectories, we computed the time spent in each location of the environment (nodes and bridges). Across all sessions, subjects show an even distribution of time spent across every node and every bridge of space, with less time attributed to bridges than nodes, which aligns with the fact that the bridges are a mere passage spot on the way to platforms, where animals stop for reward collection and/or rearing (Figure 4.11(a)). Time spent on the maze locations is therefore influenced by the reward distribution of the session. The

strikingly high occupancy of node G1 (bottom right) relates to the fact that this is the start and end location of every session, for every subject. Hence, animals exhibited a strong preference for occupying this platform in moments of break from the task.

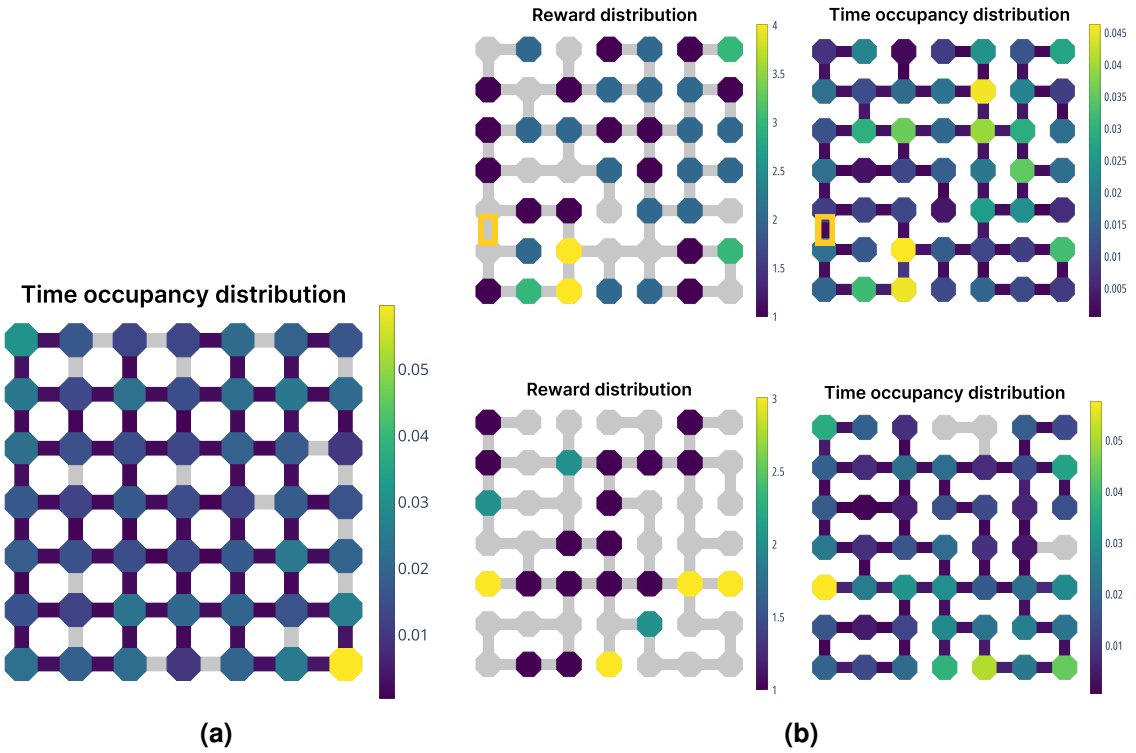


Figure 4.11: Time spent occupancy distribution across the 7x7 maze. From dark to lighter colours there is an increase in the quantity of observed time occupancy distribution. (a) overall time occupancy distribution across the maze for all sessions; (b) two example sessions time occupancy distribution (right column, top and bottom), which appear to correlate with the reward distribution of the corresponding sessions (left column, top and bottom).

More interestingly, we also calculated the occupancy in terms of the number of visits the mice paid to every location, i.e. incrementing one visit to a given location every time the animal entered that location from another area. While providing the overall preference of traversed locations throughout the experiment for all the mice used (Figure 4.12(a)), it reveals more interesting spatial patterns if we con-

sider individual maze structures. Figure 4.12(b) is an example of how impactful a maze change can be in shaping mice's favourite trajectories in just one session, while Figure 4.12(c) elicits the preferred routes' transformation that happens within two sessions of experience in the same structure. From day one, animals start developing preferred subsections in the maze, which include traversing some of the inner platforms and bridges of the maze more often than others. Such preferences slowly morph across time, with the slow evolution of the maze structure (Figure 4.12(b) and Figure 4.12(c)). Curiously, animals rarely (if ever) use the entirety of their favourite substructures in one go.

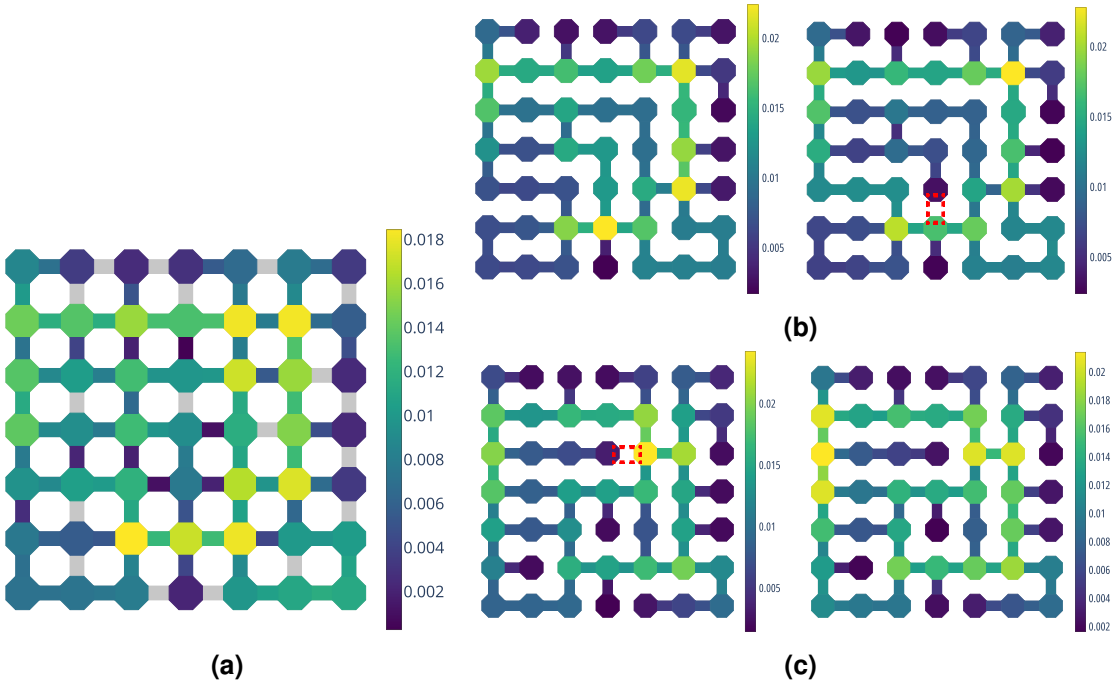


Figure 4.12: Number of visits occupancy distribution across the 7x7 maze. From dark to lighter colours there is an increase in the number of observed visits occupancy distribution. (a) overall visits occupancy (entries coming from a different location) distribution to each location of the maze; (b) drastic change in visits occupancy distribution pattern of example mouse from maze 66, S2 (left) to maze 67, S4 (right), due to the removal of the link in red; (c) consolidation of the visits occupancy distribution pattern of example mouse from maze 85, S4 (left) to maze 85, S2 (right), eliciting the effects of experiencing the same structure twice.

4.3.3 Mice incorporate spatial changes in their navigation strategy within the first session of experience

Firstly, we wanted to understand what strategy mice are using to goal-directedly navigate such a large and complex environment. More specifically, we ought to quantify the influence of vector (or Euclidean) navigation and shortest path (or planning) on the animals' strategy. For that, we analysed and modelled the choices taken by individual mice at every decision point using a mixture of strategies model with two components - a vector and a shortest path navigation components (Section 4.2.2). The weight of each component will inform the extent to which the respective component informs the subjects' strategy.

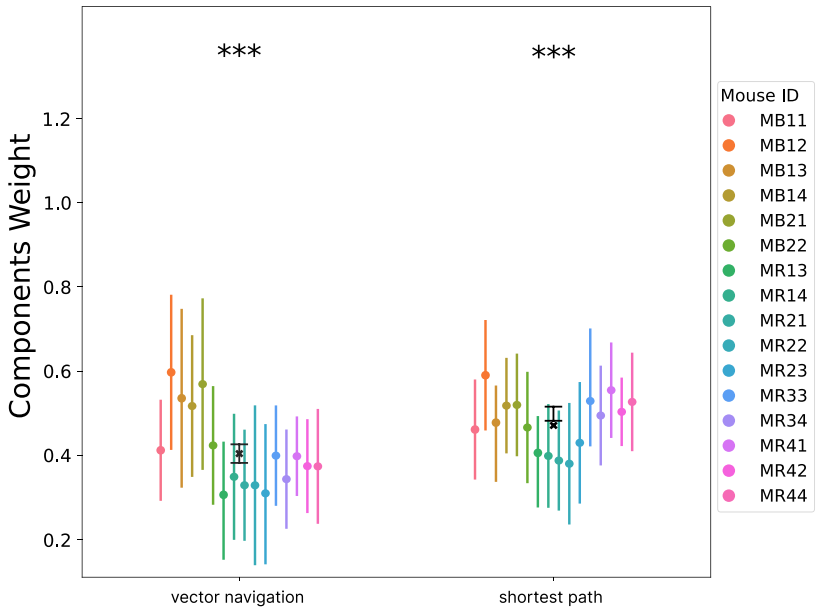


Figure 4.13: Mice exhibit a mixture of vector and shortest path navigation strategies for solving the task. Representation of the individual animals' weights for the vector and shortest path navigation components (mean±SD across sessions, in colours), along with the overall weights of these components (mean±SEM across animals, in black). Both vector and planning navigation components' weights are significant (p-value=2.202e^{-295} and p-value=0.0, in order) and the following pairwise correlation - vec_nav-sho_pat:1.771e^{-18}.

Looking at behavioural data from all animals - we can observe that, although the load of the shortest path component (Q_{S0}) is higher than that of the vector navigation component, mice seem to strongly rely on both strategies to solve the task and efficiently collect rewards across the maze (Figure 4.13). This remains true throughout the entire experiment; however, mice tend to rely more and more on their knowledge of the spatial structure as time goes by than on their angle to goal information (Figure 4.14).

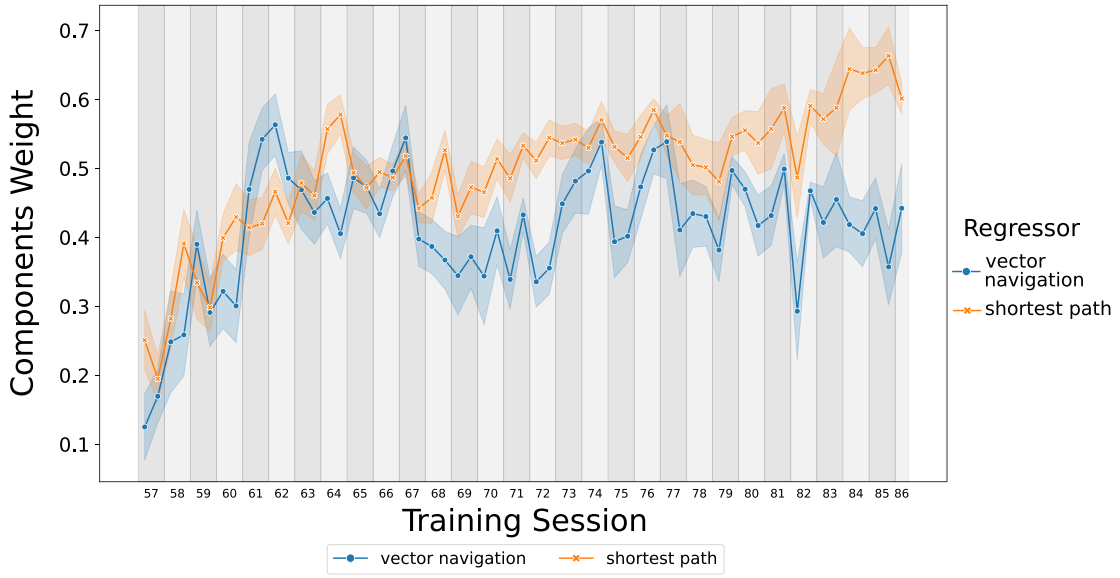


Figure 4.14: Vector vs shortest path navigation across maze structures. Animals keep a mixed navigation strategy throughout the whole experiment; however, after they experience the first five maze structures, they consistently more strongly rely on their knowledge of space than on their angle to goal information.

Additionally, we were interested in understanding how quickly the mice incorporated new information into their navigation strategy following single-link changes in the maze structure. For that, we analysed the first sessions of each mouse after a link had been added or removed from the maze by fitting the same mixture of

strategies model with an extra term - the "Δ structure component", which reflected the changes in value to options between the current and the previous maze structures. We expect that if mice have not integrated information about the change into their route-planning strategy, only the previous structure planning component (Q_{St}) will load, whereas if subjects have fully updated their planning model to use the new structure, both components should load equally.

We found that within the first session the animals see that a change has occurred in the maze, they are able to update their navigation strategy and incorporate the new spatial information into their map of the environment (Figure 4.15).

Lastly, we investigated whether the addition or removal of a bridge would impact the mice's strategy similarly, and so divided the "Δ structure component" into two - the "Δ structure-added component" and the "Δ structure-removed component" for the cases when a bridge had been added to the maze structure or removed from it, respectively -, and applied the same maximum likelihood fitting to the new Q_{net}.

In Figure 4.16, we can observe that both changes in value caused by adding or removing bridges from the maze structure are efficiently incorporated into the subjects' planning strategy within the first session they are exposed to this change. However, the loading of the "Δstructure-removed component" weight is significantly higher than the one of the "Δstructure-added component", indicating that removing a bridge from the environment will more strongly inform the mice planning strategy/ is more quickly incorporated into the mice's planning strategy than the appearance of a new bridge.

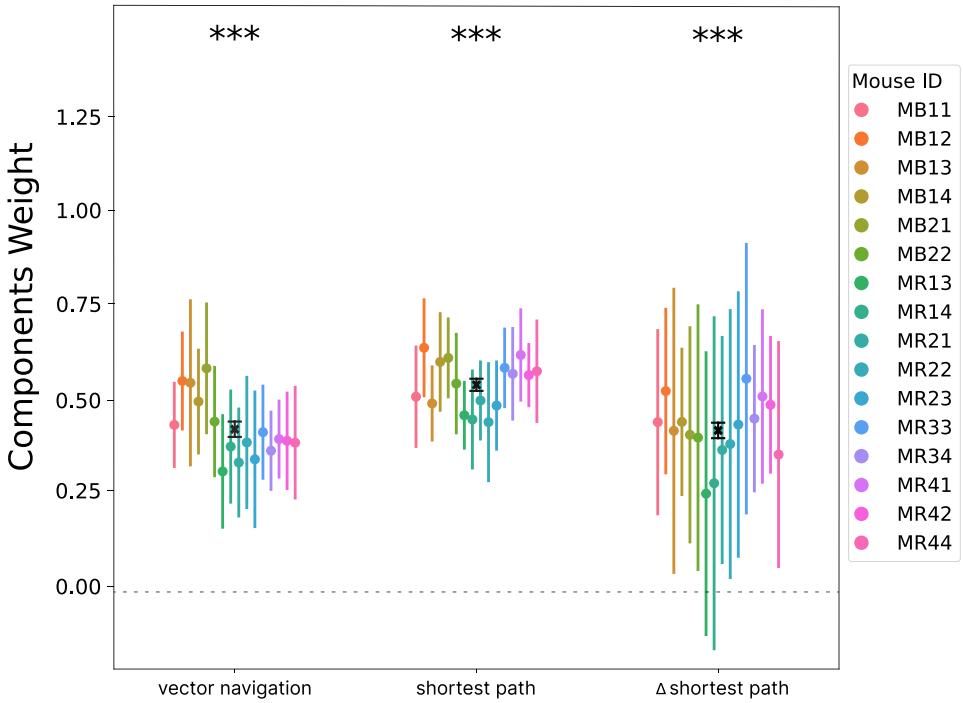


Figure 4.15: Mice integrate the changed-link information into their navigation strategy within the first session of experience. Representation of the individual animals' weights for the vector navigation, shortest path navigation and difference in structure components (mean±SD across sessions, in colours), along with the overall weights of these components (mean±SEM across animals, in black). The three components show significant positive weights (p-value=$6.769e^{-154}$, p-value=$1.858e^{-270}$ and p-value=$1.283e^{-78}$, in order) and the following pairwise correlations - vec_nav_short_path: $3.156e^{-26}$, vec_nav minus Δ short_path: 0.887, short_path minus Δ short_path: $6.153e^{-11}$.

Section 4.4

Discussion

Goal-directed navigation is an example of sequential decision-making behaviour. Following an RL perspective, it can be decomposed into a series of choices between the different options presented at each given location (state) and time. These choices are modulated by the environment - its size, structure complexity and dy-

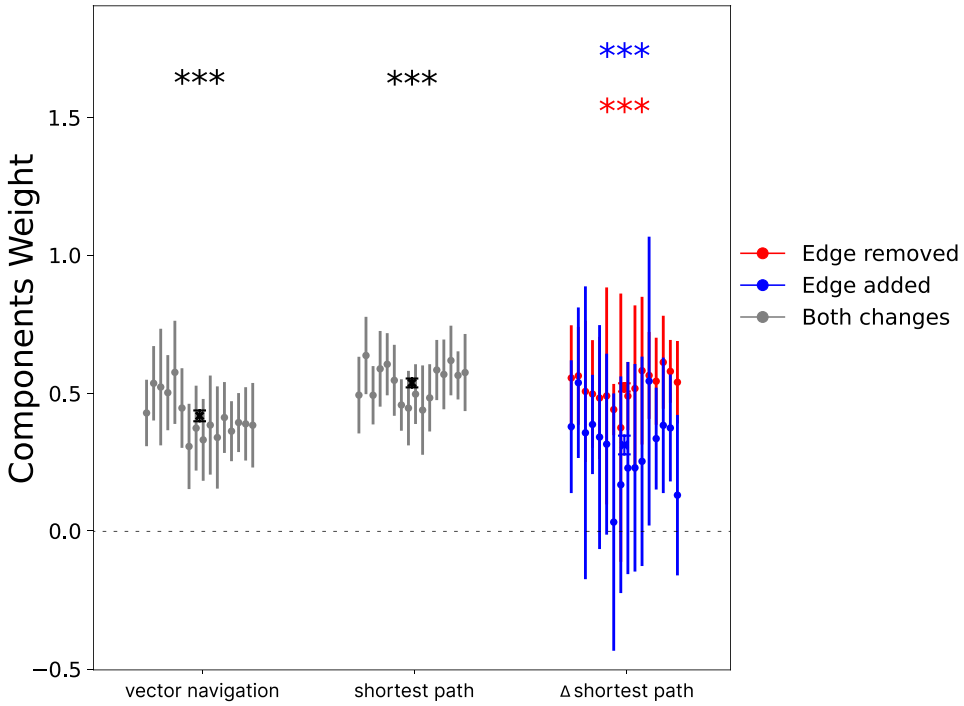


Figure 4.16: Mice integrate the changed-link information differently depending on the nature of the change. Representation of the individual animals' weights for the vector and shortest path navigation components (mean+SD across sessions, in grey) and respective overall weights (mean±SEM across animals, in black X - p-value: $8.850e^{-150}$, p-value: $1.333e^{-209}$, in order). The difference in structure component appears split into two components - a difference resulting from adding or removing one link (mean±SD across session, in blue and red respectively) and corresponding overall weights of these components (mean±SEM across animals, in an X). Both added- and removed-link components are significantly positive (p-value(Δsho_pat_added): $2.410e^{-19}$, p-value(Δsho_pat_removed): $1.294e^{-73}$), with a stronger influence in strategy caused by removed links: pairwise correlation Δsho_pat_added - Δsho_pat_removed: $6.317ee^{-11}$.

namics - and the received outcome (rewards). Here, we sought to investigate how mice navigate a large and complex state-space in the shape of a grid, and to understand their behavioural strategy in light of the MF vs MB dilemma (Daw et al., 2005) when facing spatial changes.

After establishing an efficient training protocol for this task (Section 4.2.1), in Section 4.3.2, we provided evidence of the animals' quick task generalisation onto

the larger environment, with a clear increase in the number of trials (and decrease in the number of omissions) after three to four days in the 7x7 maze (Figure 4.7). This observation suggests that mice develop a model of the task during the training stage, which allows them to easily apply the same knowledge to a similar space, albeit with a completely distinct dimension and structure, and that keeps changing from day one. But do they also integrate a model of space?

The random allocation of rewards in this task precludes the prevalence of habitual behaviour, and thus we expect that, in order to efficiently navigate this changing world, mice should build a flexible map of space. The fact that the mice modulate their velocity in the vicinities of recently removed bridges and manage to create alternative paths around them, or incorporate newly introduced bridges into their trajectories hints at the learning of such a map. An increase in overall speed in the second sessions of change experience (s2) when compared to the first sessions of change experience (s4) (Figure 4.9) could support the idea of information integration soon after the change has been introduced, thus allowing for better performance on the second session.

However, as it is presented, this analysis is confounded by the fact that sessions 2 are the first sessions of the day, and so animals are generally more deprived of water and thus more engaged in the task. Nonetheless, we quantitatively demonstrate that animals significantly integrate this novel information into their strategy within the first session after the change has been introduced (Figure 4.13). Such a quick policy adaptation could favour efficient navigation and encoding of the subsequent change, as, like this, animals need to update one-link change impact on

the world structure at a time.

Additionally, bridge removal seems to have a more significant impact on the subjects' strategy than addition (Figure 4.15). Intuitively, this makes sense - for instance, when cooking, if we run out of olive oil, we might replace it with another available oil, like sesame; however, if we have always used sesame oil and someone brings olive oil into our kitchen, it might take a while before we realise a new option is available and we don't immediately feel the need to use it in our recipes. In fact, the same happens with mice in this space - as illustrated in Figure 4.8, they realise early in the session the limitations of having one travelling option removed, adjusting around it, whereas the new link gets gradually incorporated throughout the session, and after the animal has seen it. Nonetheless, we are limited by our power to realise when the animals actually see the added links for the first time - assuming this happens when they first traverse it can be erroneous, as the animal might have seen it from afar prior to the visit, and we do not track mouse eye movement.

With the purpose of understanding mice's inherent strategy, we observed their occupancy throughout the maze and experimental time. While time spent in each location was a rough reflection of the reward distribution within each session and across sessions (Figure 4.11), the number of visits to each location revealed a strong preference for substructures of nodes and bridges predominantly located in the inner parts of the maze (Figure 4.12). Suggestive of the establishment of habitual behavioural patterns within each structure - routes -, quantification of the animals' vector and geodesic components to navigation confirmed the prevalence

of a dual control system where both strategies are employed for efficiently solving this task (Figure 4.13). While these results elicit that subjects are relying on their knowledge of the environment's structure for maximising reward rates, we recognise it will be cheaper to employ this model of space only when the trial cannot be solved purely based on the visual input (i.e. in the case of very short distance trials, where the animal can see the path without the need to integrate further information).

More detailed investigation and modelling of this behaviour are being executed by another PhD student in the lab, Chongyu (Xiao) Qin, and has so far provided evidence for the existence of a more intricate navigation behaviour in this task that cannot be described solely on the grounds of Euclidean and geodesic navigation trade-offs. Through dimensionality reduction of the trajectories performed by the mice, Qin et al. have found elementary behavioural components of navigation that align with the hints of route formation obtained via the occupancy analysis. Also, animals have been proven to find more and more optimal routes throughout the experiment, consistent with the increase in trial speed observed across mazes (Figure 4.9(a)). Currently, we are investigating the possibility of compositional planning happening within this paradigm, with routes interpreted as behavioural states and sequential choices taking place between these routes.

We could further investigate a combination of velocity decreases in the vicinity of removed edges and head-direction, for probing the existence of VTE behaviour. VTE events are commonly found at decision points in space (e.g. at bottleneck locations, or at points when the animal needs to rethink its decision, like when facing

a recently removed bridge), and are thought to align with moments of deliberation and planning, which can only be attained in the presence of an internal model of the environment (Redish, 2016).

To conclude, we analysed several features of mouse behaviour in the context of a complex route-planning task. We provided evidence for generalisation across environments and structures, and demonstrated that animals incorporate spatial changes into their navigation policy within the first session of experience. The model used for assessing the rate of information integration is a simple, elegant one that simultaneously allows for the quantification of the contribution of two distinct navigation strategies previously observed to be represented in the HPC (Howard et al., 2014); and that have been modelled and find support in cell representations of cognitive maps (Dordek et al., 2016; Gustafson and Daw, 2011). Here, Euclidean and geodesic strategies serve as proxies for MF and MB learning, with results indicating that the animals employ a mixture of both strategies, with a slightly higher weight on the geodesic strategy overall.

5

Hippocampal representations in a complex and dynamic maze environment

Section 5.1

Introduction

The hippocampus is critically involved in the generation and maintenance of our internal models of the world (Eichenbaum and Cohen, 2014; Behrens et al., 2018). From a declarative memory perspective (Cohen and Squire, 1980; Eichenbaum, 1999) to a spatial navigation-oriented one (O'Keefe and Nadel, 1978), several studies have provided evidence that the hippocampus is necessary for the formation of

new memories (Scoville and Milner, 1957; Morris et al., 1982) and for learning and associating relevant components (Deacon et al., 2002) in spatial and abstract scenarios. Furthermore, the hippocampus is now viewed as a fundamental region for flexible responses to the dynamics of the world we inhabit, facilitating policy adjustments (Miller et al., 2017) and reshaping our knowledge to attain future changes (Hassabis et al., 2007). However, to comprehend the mechanisms behind such functions, one must first get a clear picture of how these abstract and spatial features are encoded in the hippocampal brain region (more in Section 1.2.3).

Considering the hippocampus alone, multiple types of cells have been identified and characterised to represent diverse features of the environment. The most celebrated of these neurons are the place cells. Found in the dorsal HPC (O'Keefe and Dostrovsky, 1971), place cells are known for their restricted activity to a specific location of the environment, thus covering the entire navigation space with their fields, like a map (O'Keefe and Nadel, 1978). Typically encountered in the dCA1, dorsal CA3 (dCA3) and DG regions of the hippocampal formation (Leutgeb et al., 2004, 2007), place cells are modulated by the size of the environment (Mehta et al., 1997; O'Keefe, 1976; Muller and Kubie, 1987), exhibiting even a multifield representation system in larger spaces (Fenton et al., 2008). Although usually considered a precise spatial code, place representations are flexible and can partially or globally remap in the face of spatial shape changes (Lever et al., 2002). To help with spatial navigation, cells in the HPC can also represent task-specific events (Sun et al., 2020) that allow the tracking of progression within a given circuit, or specifically code for the distance to more valuable locations (Gauthier and Tank,

2018; Sarel and Finkelstein, 2017).

Additionally, regions such as the EC and subiculum support this organisation of spatial knowledge with another set of notable representations - grid cells (Hafting et al., 2005), band cells (Krupic et al., 2012), border vector cells (Solstad et al., 2008), object vector cells (Høydal et al., 2019), head-direction cells (Taube et al., 1990), and more (Section 1.2.3.5).

Inspired by these precise neural codes of space, theories have suggested distinct mechanisms by which these representations can be combined to control behaviour (Section 1.3.1). The SR framework is an example of a model that accounts for a direct interaction between behavioural policy and state representation, contrary to what is observed in Euclidean (Dordek et al., 2016) and Geodesic models (Gustafson and Daw, 2011). Revisited by Stachenfeld et al. (2017), this framework interprets the hippocampus in the view of RL as a predictive map, where place cells code for what is likely to happen next given the present strategy, and so place fields fire with respect to the likelihood of visiting a particular location in the near future, respecting obstacles and accounting for travel direction.

While over the past decades we have acquired a precise knowledge of how individual neurons represent aspects of the spatial world and this has inspired the creation of new hippocampal theories; our knowledge is limited to simple environments and, consequently, to simple behaviours (see Section 1.2.3).

In this chapter, we look to contribute to the unmasking of hippocampal representations in a complex world. We describe how dCA1 neurons represent mouse goal-directed navigation in complex and large state spaces. Moreover, animals

are exposed to a slowly evolving maze structure where single bridge changes are introduced each day, albeit at locations that globally disturb the available paths for navigation (Section 3.3.2.2). Since mice can adapt to these changes within the first session of experience (Section 4.3.3), we expect them to show incorporation of the new spatial information into their hippocampal maps at a similar rate.

Section 5.2

Methods

5.2.1 Linearisation of the maze structure

In Chapter 4, mouse behavioural trajectories and choices in the maze were often investigated not using the actual space coordinates (i.e. meters) but instead using a code name to refer to each platform and bridge separately (Figure 2.2). To analyse the hippocampal neural activity in this environment, we introduce yet another nomenclature of space.

To implement a linearised, or *skeletonised* version of the maze structure, we further divided the maze locations, or states, into multiple positions. Each platform is composed of 5 different positions, from 0 to 4, which correspond to the centre of the node, and four equally distant locations arranged in a squared manner, each in the middle of the hexagon sides that cannot connect to bridges (Figure 5.1). Each bridge is composed of 3 positions, from 0 to 2, one of them in the centre

of the bridge, and all equally distant from each other (Figure 5.1). This version of the maze was initially used for the calculation and illustration of hippocampal firing activity, providing a more detailed view of spatial representations during navigation, and hinting at the presence of a directional component in these maps (Section 5.3.2).

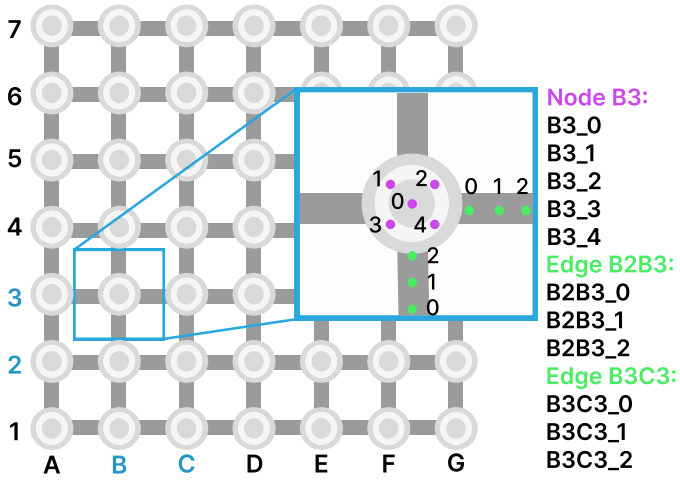


Figure 5.1: The "skeleton" maze structure. Illustration of the detailed linearisation of the maze structure. Each node is split into five locations and each bridge into three locations, as depicted in the image with the corresponding names.

5.2.2 Dimensionality reduction of neural space - PCA and NMF

With the aim of investigating what are the factors underlying the hippocampal neural representations in the used paradigm, we applied different dimensionality reduction techniques to our spike data - principal component analysis (PCA) (Cunningham and Yu, 2014) and non-negative matrix factorisation (NMF) (Kim and Park, 2008).

For each session and animal, we built a matrix $F(n,(s,a))$ that contained the values of the firing rates (Hz) of each neuron n for all the state-action (or location-

direction) pairs (s, a) of the current session maze structure. We normalised the matrix by row, such that the firing rates of each neuron summed to one across the entire space. Lastly, we performed PCA and, alternatively, NMF to the normalised matrix, and the resulting factors were plotted back on maze space (e.g. Section 5.3.3, Figure 5.2).

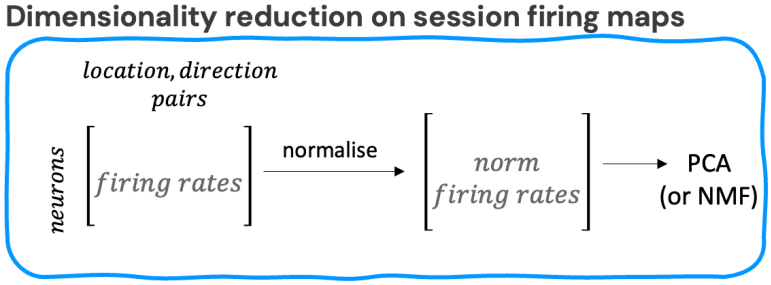


Figure 5.2: Schematic of the neural space dimensionality reduction analysis steps.

Section 5.3

Results

For the purpose of the analyses presented in this section, cell data was processed as described in Section 2.7.2 and no further curation was performed manually on those results. Hence, results and speculation about single cells should be taken with caution.

5.3.1 Complex hippocampal maps in a complex environment

We recorded dCA1 neural activity from 10 mice while running a goal-directed paradigm on the 7x7 grid maze. On average, we recorded a number of pyramidal cells per day varying between 2 and 109 per session, with 25.06$\pm$18.98 (mean$\pm$SD across mice) putative pyramidal cells per session for the first cohort, and 48.02$\pm$14.87 (mean$\pm$SD across mice) putative pyramidal cells per session for the second cohort. In total, we recorded 8869 putative pyramidal cells across all mice, over 506 sessions and $21{,}162$ trials.

One of the main goals of this project was to investigate hippocampal neural representations in large and complex state spaces, such as the 7x7 grid maze. Therefore, we started by looking into single sessions' individual cell average firing maps.

Initially, we wanted to translate single-cell spiking activity into firing maps by identifying the animal's real-space coordinate location in space at the time of the cell firing. However, when trying to do so, we were confronted with a technicality inherent to the setup structure that would not allow the representation of such maps - the fact that in a grid maze there's a significant amount of space that the mice never traverse in between all the regions of interest. This causes the firing activity to drop abruptly around the bridges and platforms that the animals visit, and so the visual representation of the neural activity in those locations was hard/impossible using these coordinates.

To overcome this, we decided to use a discretised version of the maze, where

we refer to the locations either by their platform or bridge code (Figure 2.2). Additionally, we created a *skeletonised* version of the maze with which we can identify more fine-grained locations within each platform or bridge (Figure 5.1), and therefore get more detailed spatial representations of the environment.

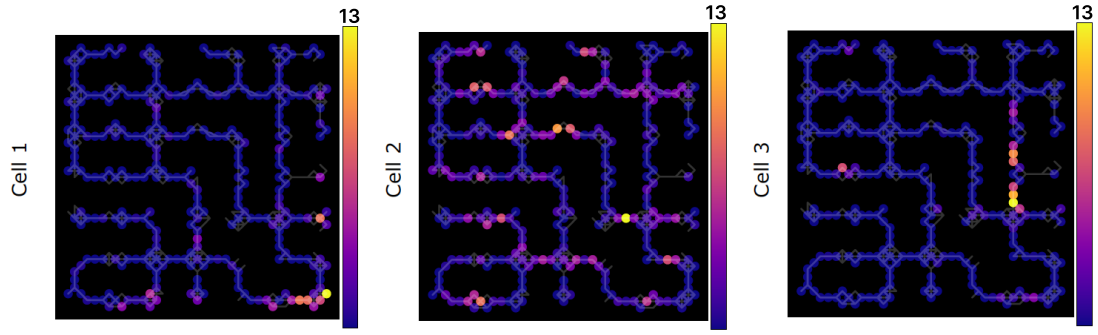


Figure 5.3: Detailed dCA1 neuronal activity session firing maps. Three cells recorded in dCA1 while the animal performs a goal-directed navigation task on the 7x7 grid maze. (left) example of a rare single-field firing cell - this cell represents the location through which the mouse enters and exits the maze every day. (centre) example of a commonly found multiple-field firing cell, representing scattered and structure locations across the maze. (right) example of a cell that fires with increasing activity along vertical corridors of the maze, and along each bridge within that corridor. The maps presented respect a minimum 0.2 correlation coefficient between an odd-even trial split of each session, precluding the depiction of single trial representations. More examples in Appendix A.

Considering the hippocampal putative pyramidal cells identified by Kilosort, we calculated the average session firing maps for each cell in the *skeletonised* version of the maze structure. In Figure 5.3, we can observe some example cells from different animals and spatial structures. Some of the recorded cells exhibit single localised firing fields on the maze, eliciting the need to represent specific locations within the structure as, for example, the platform where mice get placed on and picked up from at the start and end, respectively, of every session (Figure 5.3, cell 1). Nonetheless, most hippocampal cells present multiple fields scattered across portions of space (Figure 5.3, cell 2).

More interestingly, some cells display a gradual increase in firing rate along given subparts of the maze, e.g. along corridors and even along single bridges within those corridors (Figure 5.3, cell 3).

5.3.2 Directional maps

To inspect whether these firing patterns were in a way related to the animals' running direction - motivated by the successor representation theory types of representations found in linear track environments (Stachenfeld et al., 2017) -, we split the neural activity according to the subjects' running direction when passing a given location. That is, when considering the non-*skeletonised* version of the maze, for every given location - platform or bridge - we separately calculated the firing rate activity for when the subject was running in each of the applicable directions - North (N), South (S), East (E), West (W).

For this analysis, we only include neurons that fired between 2000-10,000 spikes per session and had a minimum of 10 spikes per location-direction pair. All the maps represented in this section have been validated to represent the actual average activity of the session through data splitting (odd vs even trials) - only data that holds a minimum 0.2 correlation coefficient across both groups of trials is presented.

In Figure 5.4, we present examples of highly directionally selective cells that represent given locations of space only when the animals are running in a specific direction(s). We find cells that fire specifically when the animals run in each

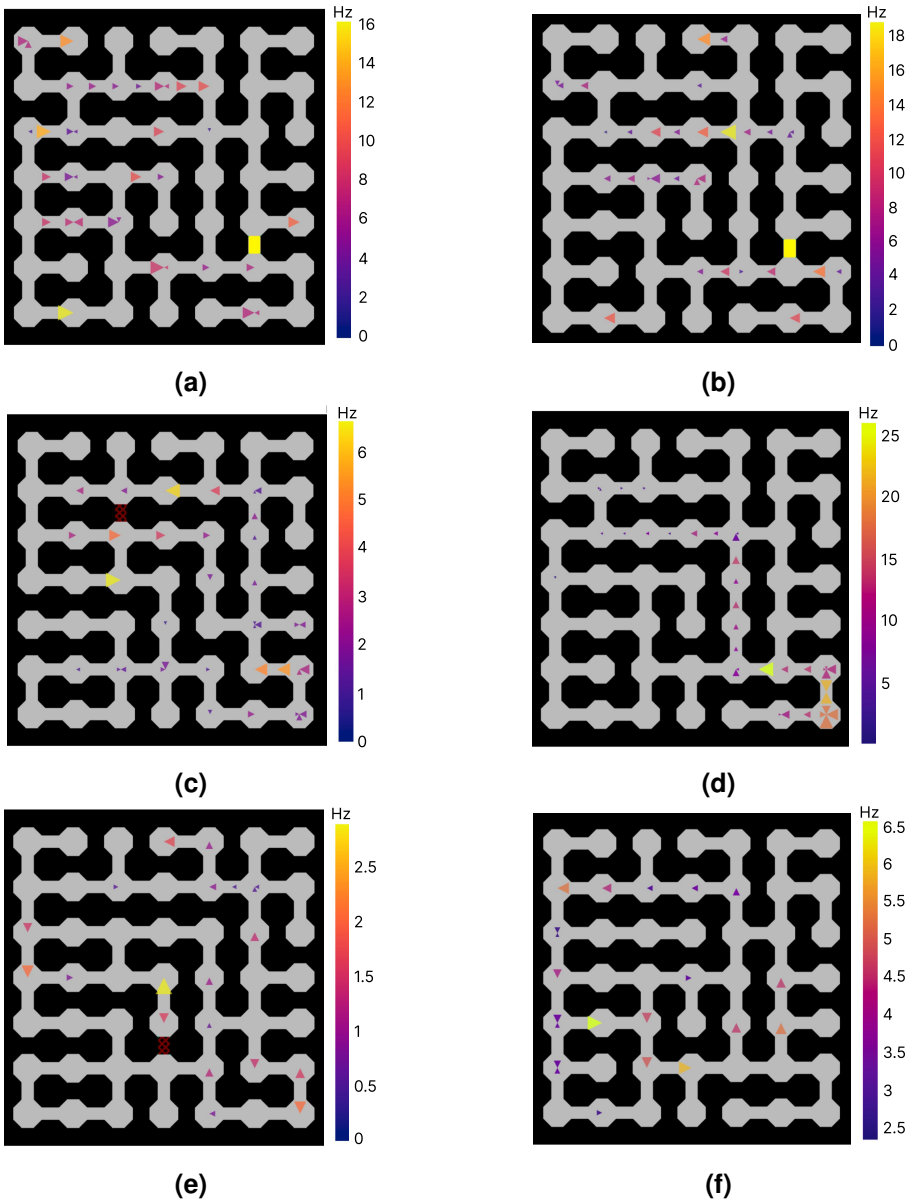


Figure 5.4: Directional session firing maps of dCA1 neuronal activity. In these maps, colour represents the average firing rate at the location independent of direction (increasing activity from dark to light colours), and arrow size indicates directionally selective firing rate proportion in the direction the arrow is pointing to. (a) example East firing cell; (b) example West firing cell; (c) example cell whose fields are extended representing partial routes; (d) example cell whose firing field extends over a trajectory used to go inwards from the session starting location; (e) example ”loop” cell directionally representing an inner structure of the maze; (f) example loop cell with fields along an inner circle structure that allows easy access to every location in the maze. The maps presented respect a minimum 0.2 correlation coefficient between an odd-even trial split of each session, precluding the depiction of single trial representations. More examples in Appendix A.

direction (Figure 5.4(a) and Figure 5.4(b)) in interaction with specific bridges or nodes. We have also consistently identified - across animals and maze structures - cells that fire in loops along sub-structures of the environment either clockwise or anti-clockwise. Again, these loops are represented not in their entirety or with great spatial detail, but in particular “anchor points” in space (Figure 5.4(e) and Figure 5.4(f)).

Furthermore, other neurons represent trajectories or parts of trajectories that are not necessarily associated with the central loop structures. In Figure 5.4(d), a cell represents a trajectory that extends inwards across the maze from the entry and exit location i.e. the place from where the mice start and get picked up at the beginning and end of every session, respectively. The cell in Figure 5.4(c) is another example of a ”trajectory” cell.

Alternatively, neurons can exhibit highly specific firing maps whose underlying components cannot be as easily identified or explained by the eye as direction or location specificity (platform vs bridge), or coherent sequences of actions (into routes). Nevertheless, being able to categorise the underlying factors constructing these maps is a crucial step for better understanding hippocampal representations and updating our model of how this region works and interacts with the rest of the brain.

5.3.3 Dimensionality reduction of the hippocampal neural space

Thus, we sought to discover and summarise the main features that composed the hippocampal representations observed for each maze structure by performing dimensionality reduction analysis of the neural data.

We attempted two different dimensionality reduction algorithms - NMF and PCA. There are advantages and disadvantages to using each of these methods. If, on the one hand, NMF will return only positive components, therefore facilitating the interpretability of the resulting components comparatively to PCA whose solutions have negative and positive components; on the other hand, PCA guarantees the components obtained to be orthogonal to each other, simplifying interpretation and translating in need for fewer components to efficiently represent the data when compared to NMF.

Figure 5.5 presents underlying factors of hippocampal neural activity for location-direction pairs by maze structure. These factors were obtained using a PCA or NMF approach (Section 5.2.2), for which neurons were only considered if they fired between 2000-10,000 spikes/session. In Figure 5.5(a), we can observe, for each structure, one example PCA result, with the positive and negative components represented in blue and red, respectively. We point out that these factors have strong directionality associated and, in most cases, are organised around substructures of the maze.

In yellow, we highlight parts of these components that present themselves in the

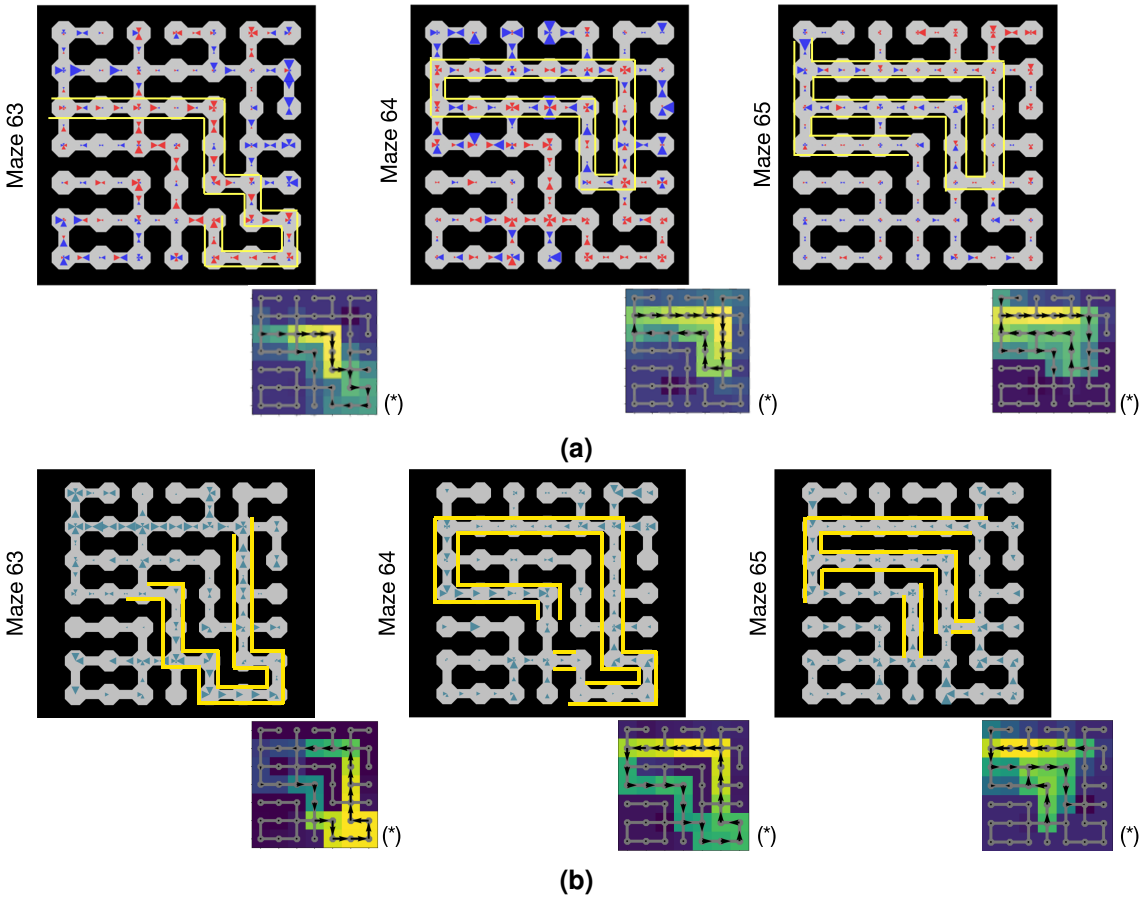


Figure 5.5: Neural factors of hippocampal activity in the maze. Example location-direction pair components resulting from the (a) PCA and (b) NMF of the neural activity recorded in dCA1 for three consecutive maze structures (63, 64, 65). For (a), the red and blue arrows correspond to the negative and positive fractions of the component. (a) and (b) dimensionality reduction of the neural space reveals neural components (highlighted in yellow) resembling behavioural components(*) extracted from mice navigation data (Qin et al.). In the behavioural component maps(*), the arrows depict the direction of movement, with bigger arrow sizes and lighter colours standing for spatial preferences.

shape of loops or partial trajectories around the environment and that, in particular, resemble principal components extracted from the behavioural data of these animals (study by Qin et al., previously mentioned in Section 4.4). Dimensionality reduction of the neural space using NMF yields similar results, with the majority of the components presenting a strong direction component and some of them mirroring the components extracted from behavioural analyses.

Pertinently, the first several encountered neural principal components do not necessarily match the first several principal components found for behaviour. In fact, they are almost always different to each other, empowering this analysis as a means to find the underlying components of the neural patterns observed in the maze, and not just a mere reflection of the behavioural components.

Section 5.4

Discussion

The hippocampus is considered one of the central pieces of the cognitive maps puzzle. However, its study has so far been limited in physical space to simple and deterministic scenarios. While these have been fundamental in the characterisation of single-neuron and population coding (see Section 1.2.3), we can hardly infer how such spatial firing supports real-world behaviours. Therefore, with the goal of better understanding hippocampal representations of complex environments, we implanted mice in dCA1 prior to the performance of goal-directed navigation in a 7x7 grid maze, under a slowly evolving structure paradigm (Figure 2.5).

Following the premise that in dCA1, one can expect to find spatially modulated cells with single fields for specific locations ((O'Keefe and Dostrovsky, 1971) and many others), we were expecting to be able to easily test theoretical predictions against the data collected using our platform. For example, introducing minimal changes to the maze structure that have a global repercussion in terms of the over-

all available shortest paths is predicted to modulate dCA1 neural activity - theories predict shrinkage or disappearance of place fields close to links that have been removed, and elongation of place fields close to newly inserted links (Stachenfeld et al., 2017; Gustafson and Daw, 2011).

Nevertheless, the cells we identified in mouse dCA1 while navigating this maze did not resemble the typical place cell description. While some cells exhibited a unique localised field, most cells presented multiple fields across the entire space (Figure 5.3). And, while this was in line with previously reported responses in wide spaces (Park et al., 2011), the observation of a cell with increasing activity along a specific corridor of the maze caught our attention (Figure 5.3, cell 3) and inspired us to analyse the neural activity from a different perspective.

In Section 5.3.2, we present a series of neural firing maps by location-direction pairs instead of just location. These pairs can be thought of as state-action pairs, where the state is the physical location of the animal in space, and the action is the direction in which they were running when crossing that location. These maps showed our dCA1 cells to be highly directionally selective. The hippocampal cells we recorded exhibit specific directional patterns across space, some of which can be easily interpreted by the human eye - e.g. cells that represent one of the cardinal directions, or cells that represent a given trajectory (Figure 5.4). Although trajectory-like cells have been identified in other structures of the hippocampal formation, such as the retrosplenial cortex (Alexander and Nitz, 2017), these do not entirely resemble the patterns we observe in dCA1. The trajectories we identify are not represented by firing fields in every location across a given path, but at a sub-

set of location along the path, which better aligns with observations of CA1 cells done in a linear spiral-like environment (Nitz, 2011). A recurrent representation of loop structures or other substructures of the maze suggests their importance for successful navigation in the maze. These observations suggest that dCA1 cells do not simply represent space, but instead represent a conjunction of space (nodes and bridges) with movement direction, with specific policies like going inwards towards the centre of the maze, or going counterclockwise around a given loop, with entorhinal input that concedes some periodicity to these representations, and potentially many other feature conjunctions we cannot clearly identify.

Hoping to disambiguate some of these contributions, we searched the underlying factors of these cells' activity by doing a dimensionality reduction of the neural space. They often appear in the form of trajectories and loop structures in the maze. Moreover, some of the factors calculated mirror some of the principal components of behaviour obtained by Qin et al. in the lab. Often, these are central pieces of the maze structure, connecting to multiple other locations otherwise hard to reach within the maze. Note that although these representations are highly correlated with the subjects' behaviour, mice never traverse those structures in their entirety at once. One attractive, although unproven, suggestion is that hippocampal representations are best explained as having localised firing fields in a "behavioural space" of commonly taken routes. As a result, a loop in the maze might serve as a "motorway" to another end of the maze, where the mouse will hop off and enter some other trajectory towards the goal. The hippocampus is then responsible for the representation of conjunctive information of spatial features and

other inputs (that we could not yet dissociate) and for compositionally representing these route maps to plan trajectories to the goal, just like in a lego game you assemble differently shaped pieces to reach the desired outcome, e.g. a lego house (Schwartenbeck et al., 2023).

Although this analysis is supported by the fact that the first components of neural activity are not the same as the first components of behaviour for the same mazes, we must keep into consideration that this analysis can be influenced by the animals' experience in the maze. For the same maze, the one animal will explore a given number of state-spaces (considered the same for both behavioural and neural dimensionality reduction analyses). However, different animals will explore different state-spaces in the corresponding sessions, and because the factors we encounter are a combination of the activity across animal, the lack of visits to a given location will be reflected in the results. Even though we estimate that this does not represent a significant issue for most tested maze structures, as space is very evenly explored across animals, one simple way to control for this limitation would be to simulate a uniform firing rate of an animal in the maze and find the principal components of that population activity, then comparing them to the principal components obtained from the real neural data.

In short, we proposed to unravel the mysteries of complex and dynamic spaces in the mouse hippocampus. Surprisingly, we encountered representations in dCA1 that do not match the classical take on place cell spatial firing. Neural maps of activity in the maze often exhibited multiple fields and highly directional selectivity. Moreover, some of these maps are organised in patterns that can be interpreted

by the eye, such as trajectories or loops around central structures of the maze in a specific direction. While these complicated representations have made the testing of the mentioned neural predictions harder, they inspired us to propose a hippocampal spatial coding mechanism for navigation.

According to it, the HPC compositionally plans paths to rewards by transitioning between these ”routy” representations (ongoing modelling being done in the lab). Although we can correlate the underlying components of neural and behavioural space (Figure 5.5), these correlations require formal quantification. Likewise, additional effort should be employed in the characterisation of these factors, and the quantification and differentiation of the different types of representations found, as well as in the quantification of remapping phenomena. Whilst these are the ongoing and future steps in the analysis of the hippocampal activity recorded, we believe that a simultaneous effort in characterising this behaviour is fundamental for the progress of neural interpretation, and vice-versa.

Of utmost importance is also the discussion of the parameters selected for the directional maps and the PCA and NMF analyses. As mentioned in Section 5.3.2 and Section 5.3.3, we solely consider cells with a minimum of 2000 and maximum of 10000 spikes per session. In sessions of 30 minutes, this corresponds to cells with an average firing rate between 1.1 and 5.5 Hz. Although we had previously qualitatively analysed all cells' the individual firing maps (Figure 5.3) and observed a very reduced number of typical single firing field cells when compared to multiple firing fields ones, this choice can introduce a bias in the type of cells that are being considered in the analysis. On the one hand, we might be excluding a big portion

of classical place cells, as these tend to have baseline firing rates lower than 1.1 Hz. On the other hand, we might be including several interneurons by extending our upper limit to 10000 spikes/session, which have been reported to also play a role in shaping hippocampal spatial information (Wilent and Nitz, 2007). It is thus critical to bear these caveats in mind when interpreting the results presented in this chapter.

6

Online and offline replay of an intricate hippocampal cell coding system

Section 6.1

Introduction

Flexibility is a hallmark of adaptive behaviour, enabling animals and humans to leverage past experiences to maximise outcomes in complex, changing environments (Tolman, 1948b; Sutton and Barto, 1998). In the context of spatial navigation, this flexibility is underpinned by the formation and utilisation of cognitive maps- internal representations of the environment that encode relational information be-

tween space, landmarks, goals, and context, among others (see Section 1.2.3). Crucially, efficient exploration and navigation not solely rely on constructing a clear space representation of the environment, but also on the capacity to update this same map in response to environmental changes, integrating novel information and reinforcing successful navigation strategies (Spiers and Gilbert, 2015; Burgess et al., 2002). Contemporary theories suggest that these processes of map formation, updating, and exploitation occur on multiple timescales, spanning both online periods of active navigation and offline periods of rest or sleep (Foster, 2017; Ólafsdóttir et al., 2018).

Critical for the encoding and retrieval of place and event memories (O'Keefe and Nadel, 1978; Scoville and Milner, 1957), the hippocampus provides a valuable window into these processes of map formation. Notably, besides its relational management of information (Eichenbaum, 1999), the HPC is thought to play a massive role in the flexibility of cognitive spatial maps through an inherent mechanism named replay. Often occurring during SWR (Buzsáki, 1989), replay typically stands for the reactivation of place cell activity in the same sequence of past experiences (Wilson and McNaughton, 1994; Skaggs et al., 1996; Lee and Wilson, 2002), though activity structure can also appear reorganised into new firing patterns (Gupta et al., 2010; Dragoi and Tonegawa, 2011; Ólafsdóttir et al., 2015) or even into disjoint locations (Karlsson and Frank, 2009; Lee and Wilson, 2002). Found both during sleep (offline replay) (Pavlides and Winson, 1989) and during moments of stillness in task performance (online replay) (Kudrimoti et al., 1999), hippocampal replay has been posited to support various functions depending on

the context and content of the events, such as planning (Johnson and Redish, 2007; Pfeiffer and Foster, 2013), learning through credit assignment (Ambrose et al., 2016; Foster and Wilson, 2006; Johnson et al., 2005), memory retrieval (Jadhav et al., 2012) and consolidation (Carr et al., 2011; McClelland et al., 1995), and in forming and maintaining a cognitive map (Gupta et al., 2010; Tolman, 1948b).

In this chapter, we will describe mouse replay results obtained both during sleep (offline) and during the performance of a goal-directed navigation task (online) in an elevated 7x7 grid maze. As described in Figure 2.5, the maze used for this task is a complex structured 2D world that slowly changes configuration across days. By recording hippocampal replay as the animals adapted to the changing maze topology, we sought to characterise how replay supports the updating of cognitive maps, the discovery of novel efficient routes, and the consolidation of successful navigation strategies. We hypothesised that replay would be modulated by changes in the environment, with an increased recapitulation of trajectories that represent the updated maze topology or traverse the new optimal routes. Moreover, given that we are using a goal-directed task in a large state space that precludes the use of habitual strategies, we stand in a privileged position to evaluate if online replay before or immediately after travelling to the goal location is a mechanism for route planning in navigation.

Section 6.2

Methods

6.2.1 Ripple detector

The first step of every replay analysis is to detect the replay events. Hence, we custom-coded two distinct ripple event detector algorithms, one based on the hippocampal LFP signal and the other using the hippocampal spike trains.

The ripple events detected with the spike-based ripple detector were generally more reliable than the events detected using the LFP-based ripple detector. We believe this is due to the presence of a "bug" in the LFP-based ripple detector. Unfortunately, as of the submission of this document, we have not been able to thoroughly re-examine this detector's code, and consequently, nor could we get both detectors to agree on the putative ripple events identified. For these reasons, the results presented in Section 6.3 are relative to the analysis of ripple events selected exclusively using the spike-based ripple detector described in Section 6.2.1.2.

6.2.1.1 LFP-based ripple detector

The LFP-based SWR detector we built was inspired by Kay et al. (2016) and Karlsson and Frank (2009).

The dCA1 LFP was obtained by downsampling the original signal (from all the functional channels, i.e. not dead or damaged) from 30kHz to 2kHz. Subsequent

band-pass filtering between 150-250Hz was applied to the downsampled signal, and the SWR envelope was determined using a Hilbert transform. The envelope was then smoothed with a Gaussian kernel ($\sigma = 4ms$). SWR events were identified as the periods when the signal exceeded the mean by 2 SD for a minimum of 15ms and when these times were detected in at least 10 different channels. SWR periods were defined by including the times immediately before and after the threshold crossing event when the envelope exceeded the mean. Only SWR events occurring when the animal's speed was $<$4cm/s were considered.

6.2.1.2 Spike-based ripple detector

We based our spike-based ripple detector on the detector described by Ólafsdóttir et al. (2016).

Hippocampal cells were considered for this analysis if they fired between 1000 and 10,000 times in 30 minutes sessions, and their activity was smoothed using a Gaussian kernel (σ = 5ms). Moments when neural activity exceeded the mean by 3 SD were used for detection of putative ripple events, and these were considered between the times the activity crossed the mean immediately before and after the threshold crossing moment. Putative ripple events were considered for analysis purposes if their duration was between 15-2000ms, and in periods when the animal's speed was $<$4cm/s.

6.2.2 Spatial decoders

The second step in analyzing replay events is to decode the spike activity recorded during the putative ripple events already detected. In the case of spatial behaviours, as in our project, we ought to decode the animal's location out of its neural activity. For that, we used two different kinds of decoders at different stages of this project - linear and Bayesian decoders - which we will describe next.

6.2.2.1 Linear decoder

We built a linear decoder that allowed for simple decoding of ripple activity online and offline.

For a given session, we calculated the firing map considering locations, or state-actions (i.e. location-direction pairs) of each of the considered neurons for the specific maze session. Often, we considered neurons that fired between 1000 and 10,000 times in a 30 minutes session. Then, for each putative ripple event, we weighted the participation of each neuron involved. Finally, we obtained a new firing map of the ripple activity by calculating the weighted sum of the maze firing maps using those session maps of the neurons involved in the ripple event and the participation weights of each of them (Figure 6.1).

Although this decoder is obtained using straightforward calculations, this decoding technique is limited in the sense that it cannot be used to extract sequences of movements, or trajectories potentially represented in the ripple events. While we can observe the order in which the different neurons fired within the ripple event,

the decoded activity will represent the overall space representation rather than a single location for a specific temporal bin (as seen in Bayesian decoders).

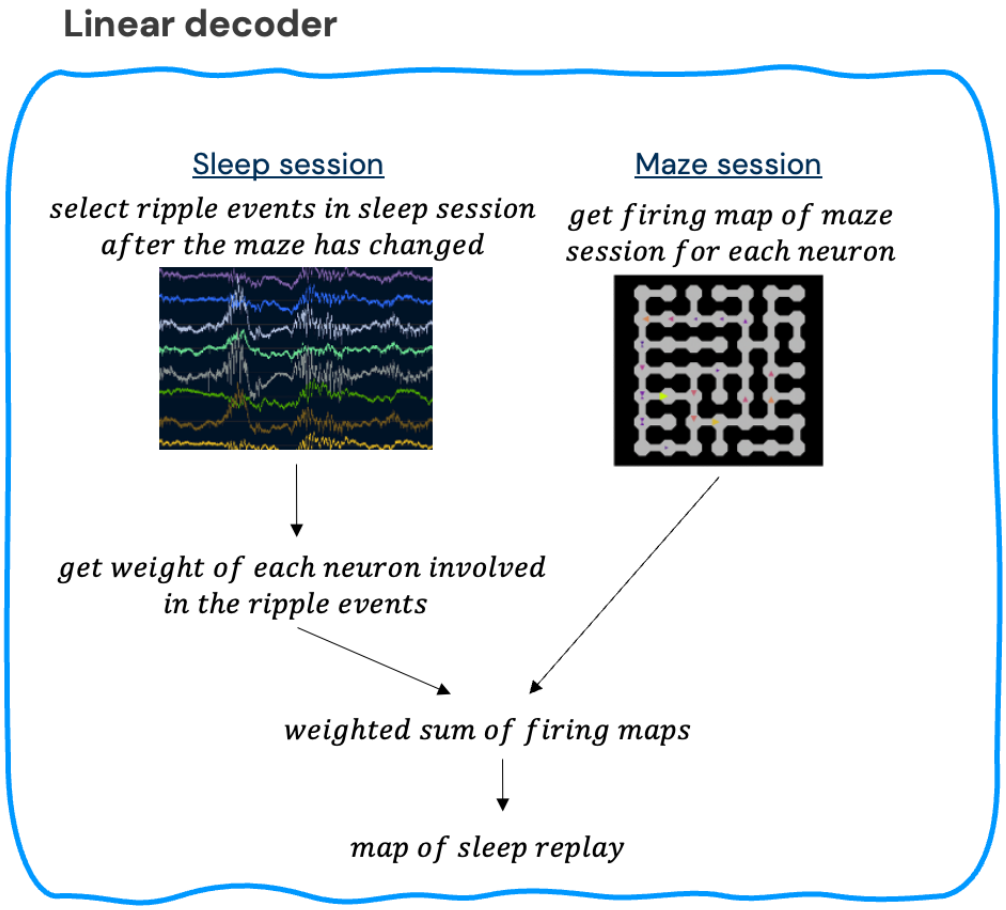


Figure 6.1: Schematic of the linear decoder analysis steps.

6.2.2.2 Bayesian decoder

The built Bayesian decoder was inspired by Ólafsdóttir et al. (2015), considering non-overlapping temporal windows.

For a specific time window (T=200ms), we considered a vector of hippocampal spikes from N neurons as $K = (k_1, ..., k_i, ..., k_N)$, where k_i was the number of spikes fired by the ith neuron. $P(K|x)$ gives us the probability of observing K

spikes in time T given the animal's position x:

$$P(K|x) = \prod_{i=1}^{N} \textit{Poisson}(k_i, Tf_i(x)) = \prod_{i=1}^{N} \frac{(T \times f_i(x))^{k_i}}{k_i!} \times e^{-T \times f_i(x)} \tag{6.1}$$

where x corresponds to the animal's location in the maze and takes one position out of all the nodes and bridges of the maze, or the state-action pair that simultaneously denotes the maze location and the animal's moving direction when crossing that location. The firing rate of the ith neuron at position x is referred to as $fi(x)$ and is obtained from the cell firing maps calculations described in Chapter 5.

The probability of the mouse position given the observed spikes can then be calculated by applying Bayes' rule and assuming a flat prior for position $P(x)$:

$$P(x|K) = \left(\prod_{i=1}^{N} f_i(x)^{k_i} \right) \exp \left(-T \sum_{i=1}^{N} f_i(x) \right) \tag{6.2}$$

This decoding method was trained on 10% of the data for each session, and tested for validity on the remainder 90% of the data. For every T, the decoded position was encountered by taking the peak of the $P(x—K)$ distribution and then compared to the actual location of the animal to obtain the decoder accuracy measure.

Section 6.3

Results

6.3.1 Hippocampal representations allow accurate position decoding in a complex environment

To allow for decoding of ripple activity while running the task on the maze (online) and during rest sessions (offline), we had to confirm that position decoding was possible and accurate using the complex neural representations found in Section 5.3 - dCA1 neurons that do not resemble the typically described one-field place cells and instead exhibit multiple fields scattered across space in a non-random and often highly directionally selective manner.

We used a Bayesian decoder (details in Section 6.2.2.2) to attempt the decoding of location and location-action pairs on the 7x7 maze across all tested structures. In Figure 6.2 we present the confusion matrices for the mouse with the best decoding accuracy values for three different maze sessions at the initial, intermediate and final stages of the experiment.

The average accuracy rate considering decoding of the exact location of the mouse is 19.01$\pm$0.55% (for a chance accuracy between 0.980-0.952%), improving substantially when gradually increasing the tolerance around the real location (Figure 6.3). The perfect decoding mean accuracy for the pairs location-direction

is 17.67±0.46% (for a chance accuracy between 0.238-0.245% for 420 and 408 maximum and minimum number of location-direction pairs available).

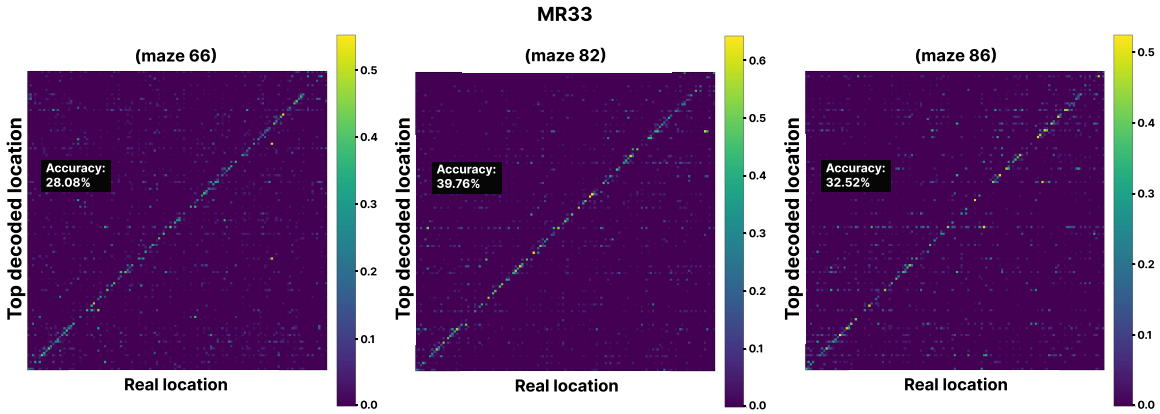


Figure 6.2: Recorded dCA1 neural activity allows decoding of the mouse position in the maze. Confusion matrices of animal decoded location (node or bridge) (y-axis) from neural activity of example mouse (MR33) for three different experienced maze structures (mazes 66, 82, 86); respective accuracy values are depicted for the decoded positions that match the real position of the mouse, with lighter colours indicating higher accuracy values. Chance level between 0.980-0.952%, for a minimum (102) and maximum (105) number of available states experienced. For decoding of location-direction pairs, see Appendix A
.

6.3.2 Online replay analyses

We hypothesise that spontaneous reactivation of hippocampal neurons as the mice run through the maze can facilitate performance in multiple ways. For instance, it could elicit planning of trajectories the subject is about to take (preplay) or replay previously executed trajectories for credit assignment at reward consumption time.

In this section, we analyse hippocampal neurons' reactivation patterns at different stationary times in the maze sessions. For that, we custom-coded two different ripple event detectors (details in Section 6.2.1). For the sake of simplicity, this section will present analyses of ripple activity extracted using the spike-based ripple

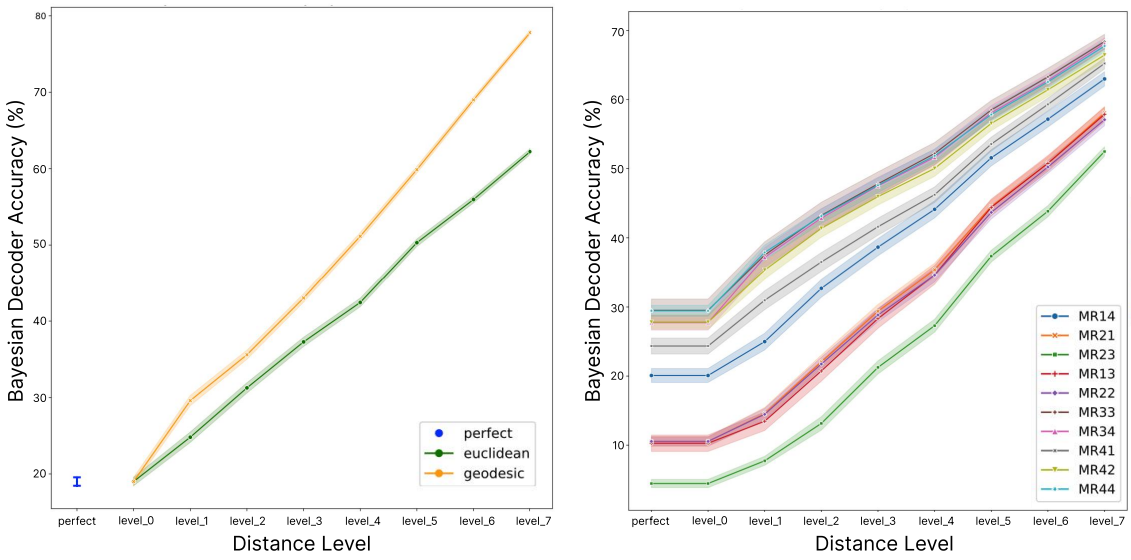
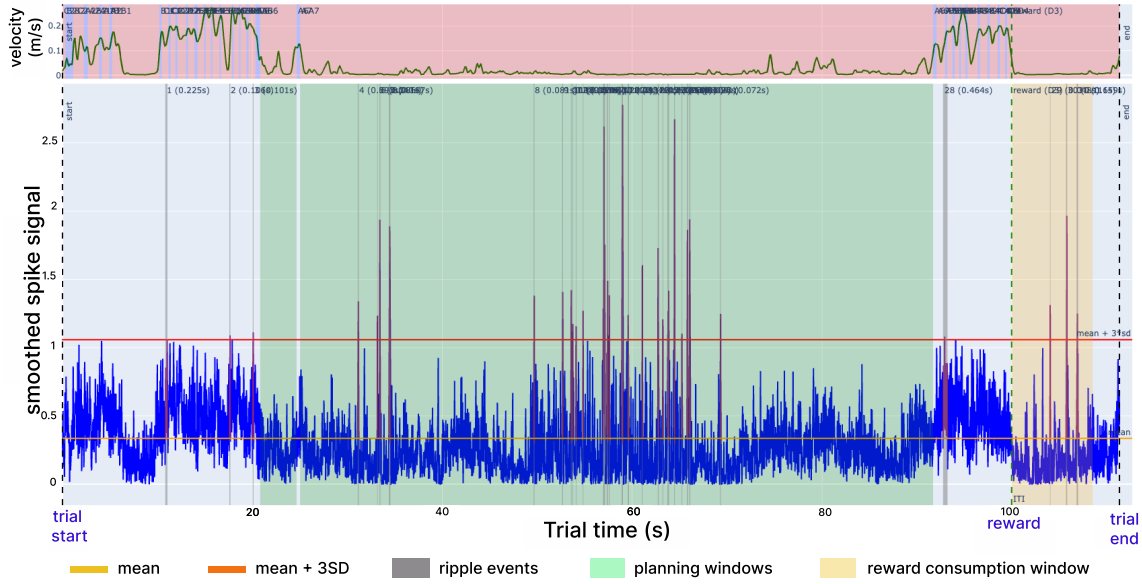


Figure 6.3: Bayesian decoder accuracy. Variation of the mean (a) and mouse (b) Bayesian decoding accuracy considering perfect decoding (blue) and decoding of locations within various levels of Euclidean and geodesic distance from the real location. Each incremented level considers one extra step away in every direction from the real location ($level_0$ stands for zero steps away, i.e. is the same as perfect decoding, $level_1$ stands for one step away from the real location, etc.). Perfect decoding does not coincide with the $level_0$ in the case where we decode pairs location-direction (Appendix A)

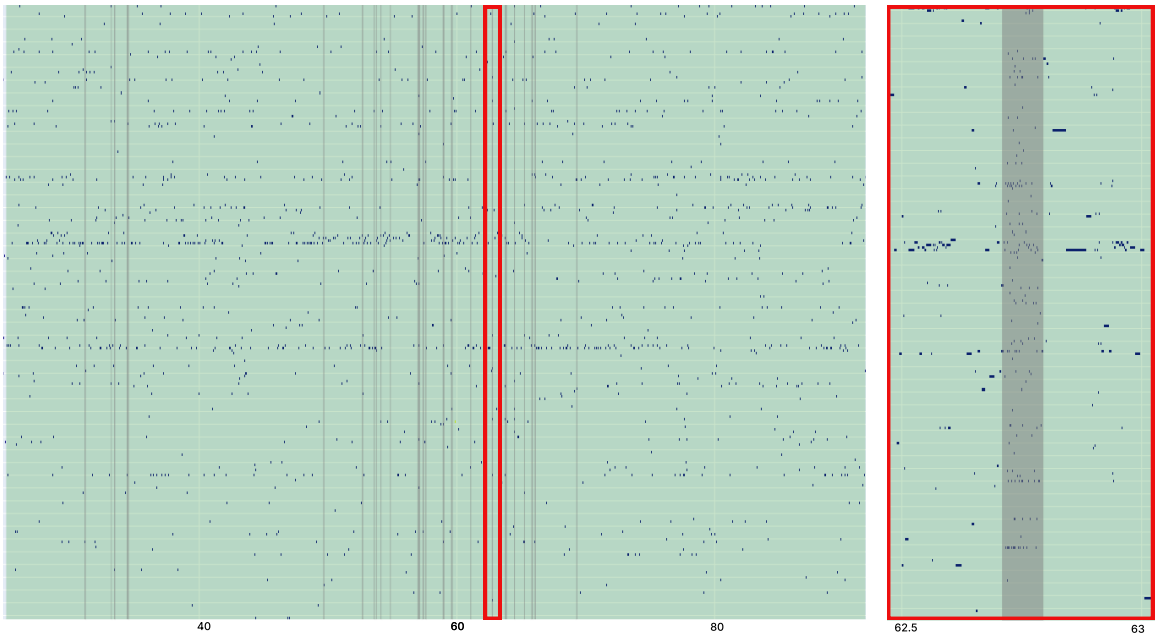
.

detector only (Section 6.2.1.2).

Zooming in at the individual trial level, we verify that, as expected, the number of ripples tends to increase in the periods of stillness immediately before running towards the goal and during reward consumption (Figure 6.4(a)). For investigation of SWR content, we defined these moments as planning and reward consumption windows, respectively, and considered valid the SWR events detected within these periods of stillness for replay analyses (velocity $<$ 4 cm/s).



(a)



(b)

Figure 6.4: Trial ripple activity. (a) Example trial of example mouse and maze structure showing (1) mouse velocity (m/s) (green line) and the passage through nodes (red shade) and bridges (blue shade); (2) the smoothed spike signal trace (dark blue) with the indication of its mean (yellow line) and mean+3SD threshold (orange line) used to detect (3) the putative ripple events using a spike-based ripple detector (Section 6.2.1.2) (purple trace and grey shade); (4) the two planning windows identified immediately before reward collection (green shade) and (5) the reward consumption window (yellow shade). (b) (left) zoom on the last planning window before reward time spike trains and (right) zoomed example of a putative ripple within the planning window period.

6.3.2.1 Hippocampal reactivation at planning times

To investigate the activity of the valid SWR detected, we considered solely the first planning window before the reward time. Potentially, we could consider multiple planning windows before the animal runs to the reward, but often, in the case of multiple windows, the further they are (in order) from the reward, the more likely it is for them to be associated with disengagement from the task and/or periods of roaming before the animal sees the goal cue.

Using a linear decoder (Section 6.2.2.1) and considering location-direction pairs instead of locations, we decoded the ripple activity of the last planning window of each trial, in each session. To that, we subtracted the average decoded map of the respective session's planning windows, obtaining the ripple firing map (normalised decoded SWR activity in maze space) of that planning window. In Figure 6.5, we can observe some examples of planning window ripple maps, alongside the mouse trajectory on the running bout following that window and up to the goal location. While in some cases the ripple activity seems to reflect the path the animal run just before the SWR event (Figure 6.5(a)), it rarely increases in regions the mouse will traverse next (Figure 6.5(b)), and, most often, it decreases (Figure 6.5(a) and Figure 6.5(c)). Importantly, replay activity always appears scattered across the entire space, not being specific to the before and/or after trajectories, but often eliciting pseudo-routes across the structure.

We wanted to assess the role of this planning window replay activity in planning the future steps of the animal at the trial level. To address this, we determined

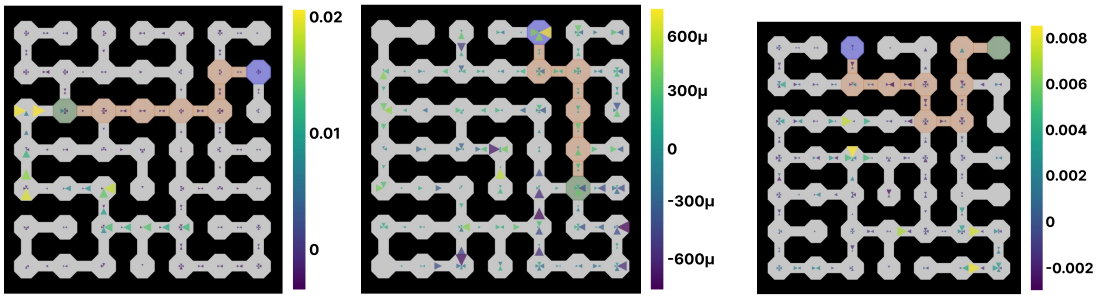


Figure 6.5: Online planning window ripple activity maps. Firing map examples for three different mice in three different maze structures from the activity of ripples detected in the last planning window before reward time using a spike-base ripple detector (Section 6.2.1.2) and a location-direction pair Bayesian decoder (Section 6.2.2.2). The arrows pointing direction the directionality of the firing activity within that location, increasing in size and presenting lighter colours for the preferred direction. The animal trajectory between the planning window and the reward poke is shaded in orange, with the starting location and end location in green and blue, respectively. (left) mouse MR33 in maze 74, trial 1/28, shows increased representation of the trajectory immediately before the planning window of ripple activity considered. (centre) mouse MR14 in maze 67, trial 33/33, shows increased representation within (but not exclusively) the trajectory performed after the planning window of ripple activity considered. (right) mouse MR41 in maze 83, trial 16/67, shows activity suppression for location-pairs within the trajectory subsequent to the planning window of ripple activity considered.

the mean firing rate of planning window SWRs decoded locations that were represented in the mouse subsequent trajectory and not, and, compiling this information across sessions, calculated the difference between these two terms such that a positive value would support a planning trajectory. Looking at Figure 6.6 we conclude that this difference ($pw - after_path$) is significantly negative (p-value: $6.561e^{-13}$), meaning that, on average, the replay activity in the planning window is suppressed for locations that will be visited on the following running bout when compared to all the other locations. Another hypothesis was that the replay activity in this period would reflect the path traversed immediately before it. However, this ($pw - before_path$), too, turned out to be negative (p-value: 4.828^{-48}). We checked

the same analysis looking for over-representation of the current location of the animal ($pw - current_loc$) and of the reward location ($pw - goal_loc$) while they are in the planning window, but these parameters also appear suppressed in comparison to the rest of the environment (p-values: 4.798e^{-94} and 0.0, respectively).

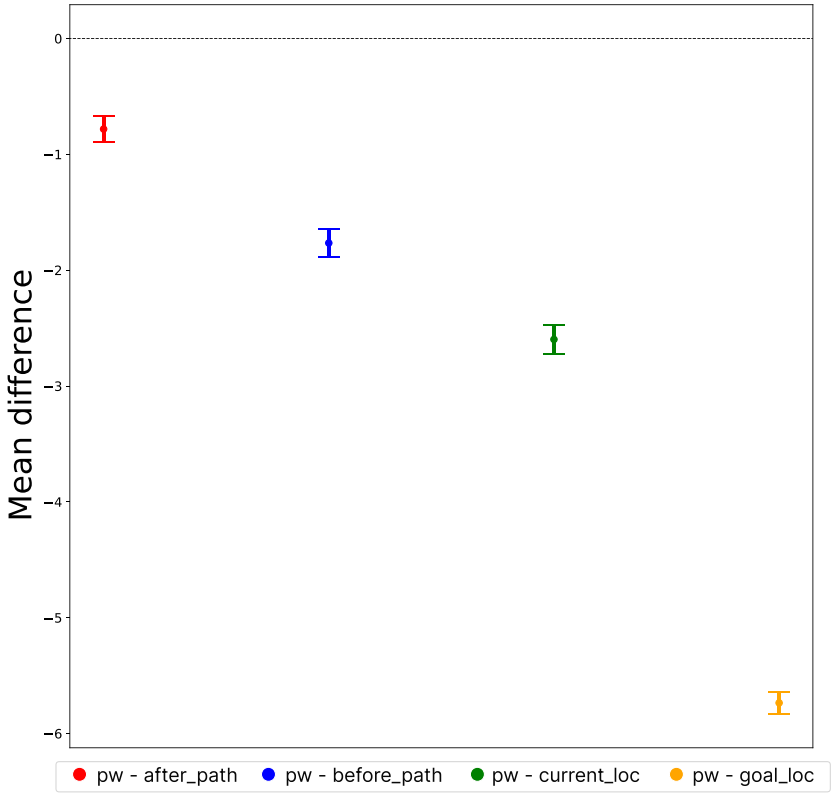


Figure 6.6: Decoding replay representations at planning window times. Analysis was done using a Bayesian decoder on the ripple events identified within the last planning window before the reward (pw). We examined if replay activity was enhanced at the following locations: $pw - after_path$, i.e. path traversed after the planning window and before the reward; $pw - before_path$, i.e. path traversed before the planning window and after the previous planning window; $pw - current_loc$, i.e. mouse location at planning window times; $pw - goal_loc$, i.e. goal location in the trial. The examined conditions were calculated by comparing the mean firing rate of every location-direction pair across animals that meet the criteria with the mean firing rate of all the other location-direction pairs in the maze that do not meet the criteria.

6.3.2.2 Hippocampal reactivation at reward consumption times

Analogously, we analysed the content of SWR events at the reward consumption window. We wanted to investigate if the activity of ripple activity at the reward location was reflecting the trial trajectory up to it ($rw - trial_path$), the trajectory just after the last planning window i.e. the running bout immediately before poking the reward ($rw - before_path$) or the current location of the animal, which in this case coincides with the goal location ($rw - current_loc$). All these conditions seemed to be under-represented when compared to the other locations in the maze (Figure 6.7) (p-values: 3.185e^{-48}, 7.207e^{-54} and 0.0, respectively).

However, we did find SWR properties in the reward consumption window to be modulated by the preceding mouse trajectory. In Figure 6.8, we can observe a positive correlation between the mean firing rate of ripple events and the trial's length of distance travelled by the animal before arriving at the goal.

6.3.3 Offline replay analyses

We also investigated neuronal activity in sleep/rest session SWR events. Particularly, we were interested in understanding how the content of replay in the sleep session immediately after the first sessions of maze change (sleep S5) impacted the animal's performance for the subsequent (same structure) maze session (S2 of the following day). We believed the decoded firing maps would reflect a difference in activity for the areas adjacent to the change, potentially decreasing activity around removed links or paths that led to them, and increasing activity along the

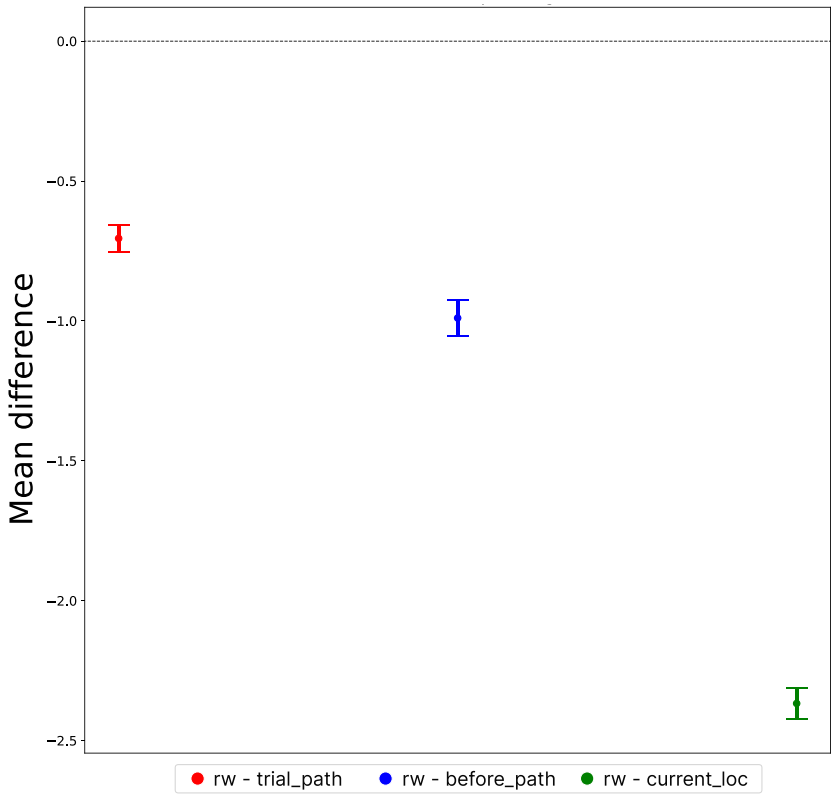


Figure 6.7: Decoding replay representations at reward consumption times. Analysis was done using a Bayesian decoder on the ripple events identified within the reward consumption window (rw). We examined if replay activity was enhanced at the following locations: $rw - trial_path$, i.e. path traversed from the beginning of the trial up to the reward location, $rw - before_path$, i.e. path traversed before poking at the reward location and after the last planning window; $rw - current_loc$, i.e. mouse current location which coincides with the goal location. The examined conditions were calculated by comparing the mean firing rate of every location-direction pair across animals that meet the criteria with the mean firing rate of all the other location-direction pairs in the maze that do not meet the criteria.

newly added links.

Decoding of the ripple activity of sessions S5 was done similarly to the decoding of online ripple activity. We used a spike-based ripple detector (Section 6.2.1.2) and a linear or Bayesian decoder (Section 6.2.2) based on the same day maze sessions to decode the putative ripples of the sleep session (we do not record videos of mice during sleep sessions, so we assume all detected ripples as valid).

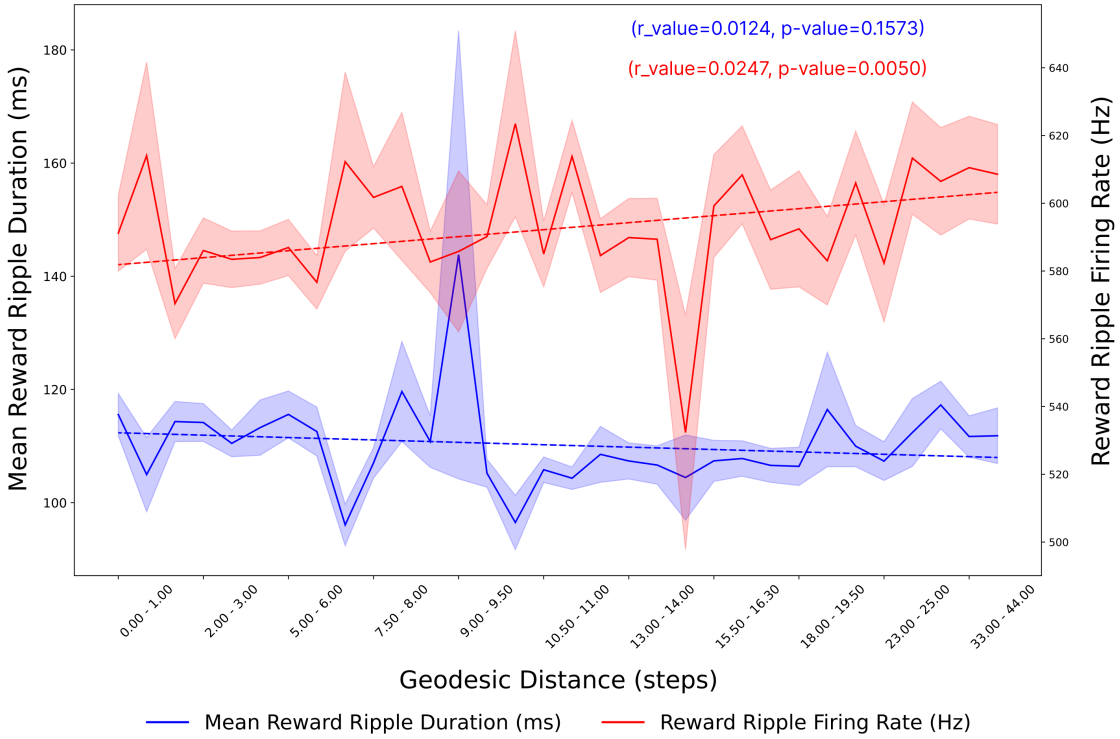


Figure 6.8: Mean SWR firing rate (Hz) at reward consumption increases with the length of the trial's trajectory. Linear correlation between the mouse traveled distance in the trial before getting to the goal location and the mean SWR firing rate or duration in the reward consumption window. Mean reward ripple duration: slope=-0.0124, p-value:0.1573; Mean reward ripple firing rate (Hz): slope=0.0247, p-value=0.0050. Distance travelled discretised into equal-sized quantiles of data (2%).

We did the same for session S3 and then plotted the difference between S5 and S3 (sleep session before the maze has changed) ripple activity across the maze (Figure 6.9). Here, we present some examples of mice sleep session firing maps after the maze structure has changed. Additionally, we present the corresponding difference in maze normalised location occupancy (in terms of the number of visits paid to each location) before and after the analysed sleep session has occurred (Figure 6.9, centre column).

By looking at the firing maps in Figure 6.9 (left and right columns), we can at times observe similarities between the variation in ripple firing activity and the

variation of location occupancy in the subsequent maze session. Nevertheless, nuances should be taken into consideration when interpreting the corresponding sleep ripple maps. The calculation of these maps cannot account for all the information and simultaneously avoid possible confounds. If we use the neuronal firing maps of mazes S2 and S4 for decoding ripples of sleep sessions S3 and S5, respectively, we are providing the corresponding maze neural activity to the subsequent sleep ripples; yet, we are introducing a comparison between maps generated with distinct behavioural bases (from mazes S2 and S4) - when we know the firing activity is strongly influenced by behaviour (Section 5.3). Contrarily, when we consider the same maze neuronal firing maps (S2 or S4) to decode both sleep sessions whilst minimising these behavioural differences, we are constructing one of the sleep ripple maps with the activity of a maze that did not precede it.

We believe that quantitative assessment of the correlations between activity and occupancy will be informative regarding the influences of sleep ripple activity on the consolidation of the changed-link information and consequent modulation of trajectories in the maze. To investigate this, we calculated a combination of correlation coefficients between the neural ripple activity of the sleep sessions of interest and the maze occupancy of the maze sessions of interest.

Decoding of location or location-direction pairs was done using the neural maps from sessions D1S2 or D1S4 (i.e. maze sessions 2 and 4 of day 1, respectively). Correlations were drawn between the ripple activity in D1S5 (i.e. the sleep session after the maze where the change was introduced in D1S4) or D1S5-D1S3 and occupancies in D1S4, D1S4-D1S2, D2S2, D2S2-D1S4; and also between ripple

activity of D1S3 and occupancies of D1S2, D1S4, D1S4-D1S2. For each combination, the correlations were assessed overall (i.e. considering all sessions) or grouped by mouse or maze. When justifiable, the correlations between absolute values of ripple and/or occupancy changes were also evaluated. When considering a change in occupancy, this refers to a change in relative occupancy. Lastly, we investigated whether the existing correlations could depend on the type of change

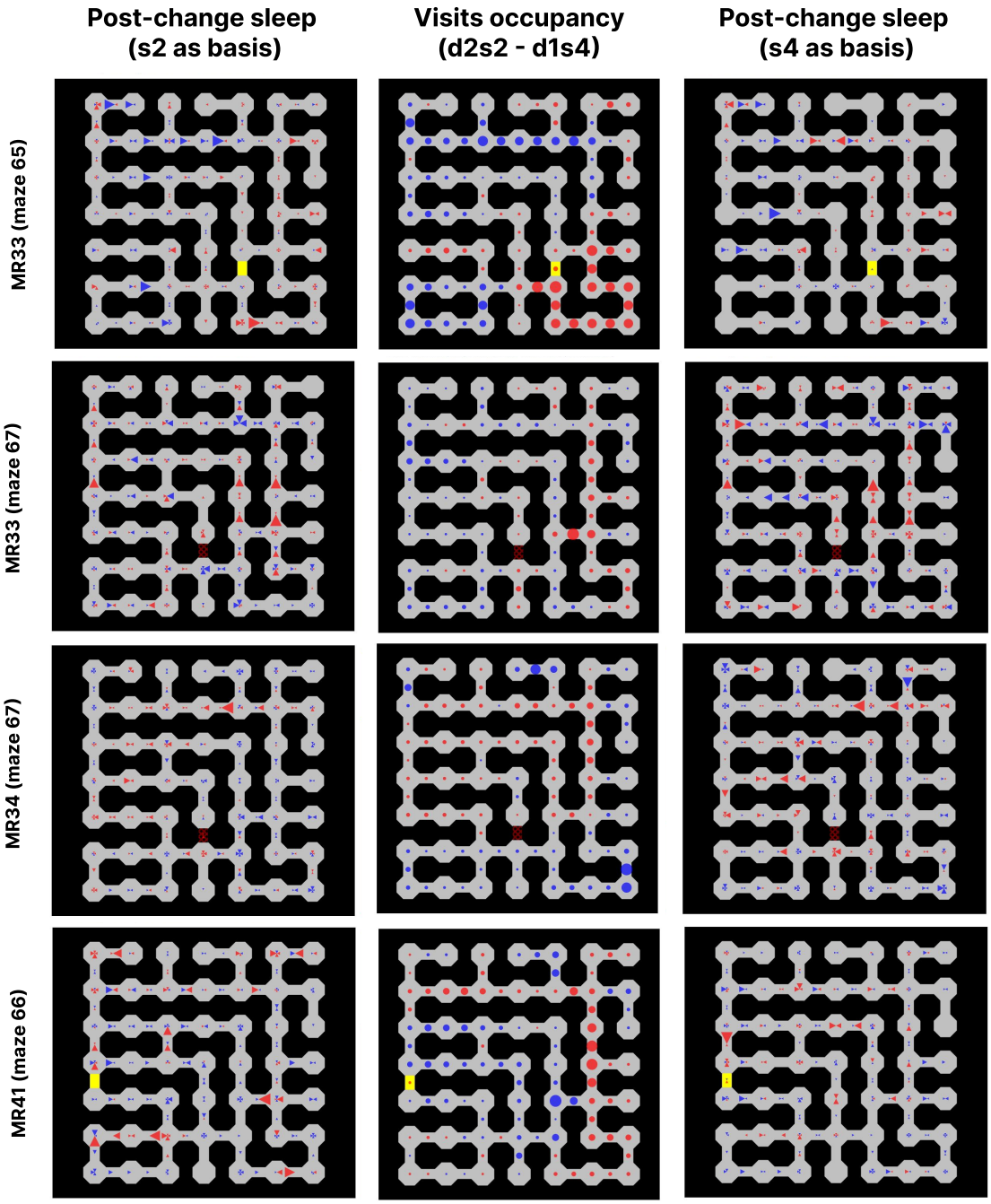


Figure 6.9: Post-change sleep session decoded ripple firing maps. Examples of post-change sleep sessions representing ripple activity decoded using the previous maze firing maps as the basis for linearly decoding (day 1, session 2, i.e. the maze structure prior to the one shown in each example) (left) or the firing maps of the maze immediately before the sleep session (day 1, session 4, i.e. same structure as the one shown in each example) (right). The variation of occupancy (in visits) from the maze session before the sleep and the maze session after the sleep (same structure in both) is depicted by location (centre). For all graphs, blue and red indicate an increase and decrease in activity or occupancy, respectively. The changed bridge is indicated in yellow when it has been added, or in crossed red when it has been removed. The same analysis was performed using a Bayesian decoder.

inflicted in the current or previous maze structure - i.e., added link vs removed link and previously added vs previously removed link.

As suggested by the analysis summarised in Figure 6.9, the difference in ripple activity detected in sleep immediately after the maze has changed (D1S5 - D1S3) did show significant correlations with the difference in maze occupancy after the sleep (D2S2-D1S4) across multiple conditions. These correlations were stronger when decoding location (Appendix A - Table A.1) as opposed to location-direction (Appendix A - Table A.2), and when considering the absolute values of occupancy change and positive ripple differences.

When considering location only, we can find a positive correlation between the sleep ripple activity variation and the change in occupancy of the maze the following day after a link has been added to the structure (t-test = 2.2609, p-value = 0.0264), but also after a link has been removed from the structure on the previous day (t-test = 2.0087, p-value = 0.0479). However, when grouping by mouse or maze, the positive correlation only holds when a link is added on the current day (by maze: t-test = 2.8186, p-value = 0.0155; by mouse: t-test = 2.4587, p-value = 0.0362). When we consider the positive ripple activity differences and the absolute value of the relative occupancy changes, every considered condition shows a significant correlation, but a negative one (Appendix A - Table A.1, absolute).

If we consider the location-direction decoding, correlations exhibit the same overall tendencies as described for location decoding, albeit less significant (Appendix A - Table A.2). Encouragingly, when using the post-change sleep ripple activity (D1S5) instead of the change in ripple activity from the previous sleep (D1S5-

D1S3), no correlation between it and the occupancy values was found (Appendix A - Table A.3). Likewise, neural ripple activity decoded from D1S3 did not correlate with any of the above-enumerated maze occupancies (Appendix A - Table A.4).

The fact that new content in sleep replay after the mouse has experienced an added link positively correlates with posterior navigation changes in the animal's behaviour aligns with the belief that offline replay is involved in learning and consolidation. However, a negative correlation between positive replay differences and the absolute value of occupancy change implies that the most replayed regions of space will be the ones where occupancy will change the most, either by increasing or reducing the number of visits.

We proceeded to examine the general effect of offline post-change ripple activity on the performance of the animal in the following maze session (day 2, session 2), which has the same structure. For that, we defined a simple maze navigation performance score measure, which corresponds to the fraction of optimal and non-optimal choices in the maze following a geodesic strategy. At every decision step, a choice is considered optimal when it leads the animal closer to the goal in a geodesic fashion $(+1)$ and non-optimal when otherwise. These scores are normalised by the number of total decisions in a session. We then compare these results for all the sessions with the total number of putative SWRs detected in the post-change sleep session. An increasing number of ripples in the post-change sleep session seems to lead to a better performance in the subsequent maze session (Figure 6.10).

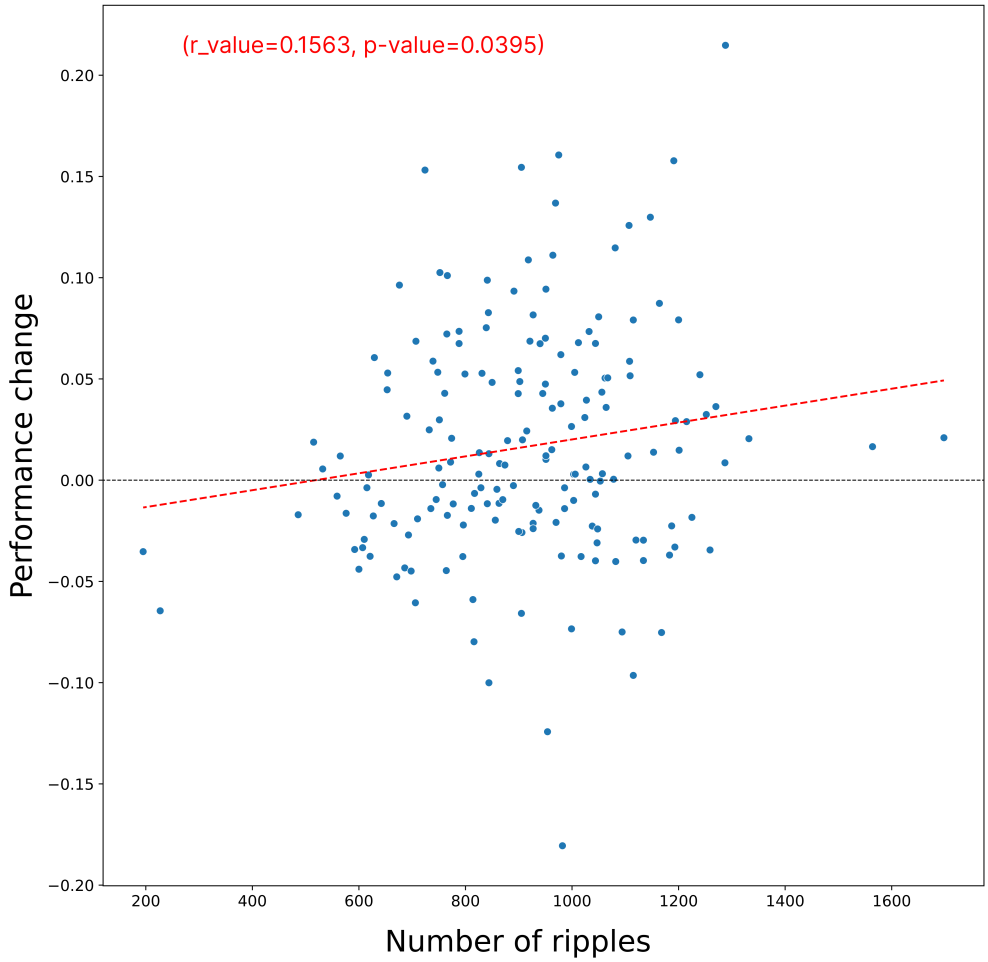


Figure 6.10: Increase in offline SWRs promotes a better performance of goal-directed navigation in a previously experienced complex maze structure. Positive correlation between the difference in performance between the second and the first experiences of the same maze structure and the total number of putative ripples detected in the sleep session between the first and second maze experiences.($r - value : 0.1563$, $p - value : 0.0395$)

Section 6.4

Discussion

Hippocampal replay is perhaps the most intriguing and special mechanism found in the brain (from a cognition perspective). It has been detected in various forms

- online, offline, in reverse and forward orders -, and is believed to support the formation and adaptation of cognitive maps in several ways - through memory consolidation, credit assignment, planning. In the 7x7 grid maze, we are in a privileged position to explore some of these questions due to its structure and task designs. We expected to understand replay involvement in on-the-fly policy updates during the task and in the learning and consolidation in memory of the space structure during post-maze sleep sessions.

We have observed in Chapter 5 that mouse hippocampal cells recorded from dCA1 often present multiple fields across the 7x7 grid maze in a conjunctive structured manner. Regardless, we can accurately decode the animal's location (and pair location-direction) in this environment using a Bayesian decoder (Figure 6.2 and Figure 6.3), thus enabling us to investigate the content of SWR events activity.

Considering SWRs that occurred during task performance within the planning window period, decoded activity appeared scattered across the available states (location-direction pairs) of the maze (Figure 6.5). Moreover, a quantitative analysis of the preferential preplay of future trajectories in this window challenges the view of online ripple events as a mechanism to support planning of future trajectories (Gupta et al., 2010; Pfeiffer and Foster, 2013; Ólafsdóttir et al., 2015) (Figure 6.6). Equally, it does not sustain the replay of the previous trajectories, nor of goal or subject current locations above other maze states. Similar analysis of the content of SWRs at the reward consumption window fails to corroborate the credit assignment and online memory consolidation views of replay at outcome times (Lee and Wilson, 2002; Foster and Wilson, 2006), due to a significant decrease in

replay of the trajectories leading to the goal location.

As much as this last result may seem surprising at first, it can make sense when performing this task in such a large and non-deterministic space like this. The fact that reward allocation is random in each trial, devalues integration of trajectories to goal, as it is highly unlikely that a trial will ever repeat with the same start and end states and, thus, one does not need to encode such information. When considering SWRs in the planning period, the fact that these do not encode planning of future visited states may seem discouraging. However, if we believe the basis of neural activity of these cells, as well as their firing maps, is organised in the shape of relevant routes and/or other policies for efficient navigation (Chapter 5), we should not, for instance, expect the replay of specific trajectories, but rather of the best policy composition for solving the trial. Such interpretations and speculations are to be considered not without forgetting that the directional firing maps and dimensionality reduction analyses have both been performed using a specific and, perhaps suboptimal, group of cells - those with 2000-10000 spikes/session, as discussed in Section 5.4. These chosen parameters are particularly relevant in the case of ripple activity analysis as, by excluding a significant portion of place cells from the analysis (lower firing cells), the spatial tuning of the obtained maps might be hindered, consequently disrupting the decoding of location during replay, and so leading to the more scattered ripple activity observed in both the planning and reward consumption windows.

We should additionally test for the over-representation in ripples of locations that were visited longer ago (Gupta et al., 2010; Carey et al., 2019), or perhaps,

a composition of the most stable states in the environment (i.e. subregions of the maze that have not suffered any alterations to that point). Exceptionally, replay of future trajectories could be justified in situations when mice encounter missing links and need to replan. Unfortunately, we did not have the time to quantitatively address these hypotheses but intend to do so in the near future.

Importantly, we should mention that the fact that all these tested mean ripple activity differences have resulted in negative values for all the assessed conditions might be the reflection of a problem with the analysis code instead of a real result. Obtaining mean values with p-values of 0.0 is particularly odd and reinforces our suspicion. We will re-examine this analysis code more carefully in order to clarify this concern.

A final analysis of online SWR events within the reward consumption window revealed the existence of a positive correlation between the mean ripple duration and the Euclidean distance to goal.

When looking at sleep, we believed that this would be involved in supporting mouse navigation performance in the maze, as lesion studies and disruption of sleep ripple activity have proven to impair spatial orientation performance (Scoville and Milner, 1957; Morris et al., 1982; Girardeau et al., 2009; Ego-Stengel and Wilson, 2010). To test that, we analysed the effect that the presence of more or fewer SWRs has on the animals' subsequent performance. The discovery of a significantly positive correlation between the number of ripples per sleep session and the performance of the animal (according to a geodesic strategy) (Figure 6.10) goes in line with the theory that offline replay plays a key role in the learning and

consolidation of cognitive maps (Pfeiffer and Foster, 2013). This relationship can be investigated further to assess the importance of offline replay in mice that generally have better or worse performances (Eldar et al., 2020). Additionally, we can examine an alternative measure of performance that better aligns with the navigation strategy we speculate mice are adopting. As discussed in Chapter 4 and Chapter 5, animals navigation is likely to be compositionally constructed out of a common pool of routes, and thus, perhaps a better performance measure would not simply be associated with a geodesic strategy, but with the composition of a fewer number of routes (and consequently activation of fewer neural basis).

However, we do not yet know how exactly this enhancement of performance is achieving that. We hope ongoing quantitative characterisation of the content of offline ripple events and its correlation with the specific changes in behaviour around the maze will help us answer this question (Figure 6.9). We hypothesise that ripple activity will correlate with the most stable locations in the maze to create a stable submap of the environment that would support efficient navigation (Pfeiffer and Foster, 2013). Yet, the newly changed-link areas, or new routes resulting from the change might also be replayed to help consolidate new information (Cheng and Frank, 2008).

For decades, hippocampal replay has been studied in the context of spatial tasks in which it has been credited with a role in the formation of cognitive maps and its flexibility. Nonetheless, such results have mostly been extracted out of simple environments which render simple behavioural routes, thus facilitating sequence matching of replayed activity (Ólafsdóttir et al., 2018). Our task takes place

in a structured and changing environment, which produces complex behaviour (Chapter 4) and atypical hippocampal representations (Chapter 5). The replay activity we found both online and offline was also intricate and complicated to characterise. The strongest result we provide in this chapter is that online replay at the planning period is not a mechanism for planning of future trajectories, although we do not yet know how to interpret the negative correlation of planning SWR activity with future and previous trajectories. Likewise, replay cannot be seen as straightforward credit assignment of previous actions at reward consumption times. The data suggests the possibility of more complex interactions with the representation bases of neural activity and behaviour. Additionally, sleep replay seems to support better performance in subsequent mazes with the same structure, aligning with theories that pose offline replay as fundamental for learning and memory consolidation.

7

Conclusion

Summary

Despite the extensive growth in knowledge in the field of cognitive maps and spatial navigation, we still lack a unifying framework that can reconcile the neural and mathematical components of complex real-world scenarios. This is evident in the disconnect between formal mathematical descriptions of internal models and the precise neuronal evidence of simple space representations, which cannot fully explain intricate flexible behaviours.

Interpreting navigation as a spatial version of sequential decision-making and spatial cognitive maps as a particular case of relational memory space, we sought to use goal-directed navigation to investigate the representation and update of structured and dynamic environments and thus to help illuminate both sides of this problem. We implemented a novel experimental platform that allows the testing of

existing models against neural data collected during the performance of a complicated spatial behaviour - the 7x7 grid maze. Comprising a total of 49 potentially rewarding states, this large structure can be rearranged in specific configurations according to the predictions to test.

We hypothesised that, to efficiently navigate such a space, mice must rely on their internal map of the environment. Hence, the optimisation of the setup's structure (Chapter 3) was performed such that it could differentiate the utilisation of a MF or a MB RL decision-making strategy by agents that would goal-directedly navigate the maze. Moreover, the structure of space was not fixed, with a small local change being implemented every day, albeit at the location with the highest calculated global repercussion throughout the entire environment.

Naturally, the first order of interrogations was related to the mouse behaviour in such a paradigm. What kind of strategy are mice using to goal-directedly navigate a complex space? And how fast can they adapt to the spatial changes introduced? Chapter 4 described behavioural signatures of mice goal-directed navigation in the 7x7 grid maze following the structural changes in the chosen sequence of structures optimised. To probe the hypothesised need for an internal model of the world to efficiently solve the task, behavioural data was collected and quantitatively analysed to determine the preferred strategy of mice in the maze, and to assess the amount of experience needed to incorporate the spatial alterations.

The second order of questions was related to how the brain computes such behavioural flexibility in this experiment. What are the underlying neuronal representations that support such efficient complex behaviour? How are they formed?

And how are they adapted and maintained in the face of change? As the HPC is known to play a key role in the formation and representation of cognitive maps, we addressed this question by recording single-cell activity from hippocampal (dCA1) neurons while performing and in the rest periods preceding and following that performance. Inspired by existing theories of how the HPC represents space (Stachenfeld et al., 2017), we expected that, during task performance, the maze would be represented by place cells with potentially multiple fields due to its big size, and that these fields would appear elongated along the newly formed pathways, and shrink or disappear at the locations where paths had been eliminated. Additionally, we wanted to disambiguate the multiple roles attributed to hippocampal replay - *online* and *offline* -, expecting to find evidence of planning for policy correction in the maze, and memory consolidation of new maze features during sleep. These matters are explored in Chapter 5 and Chapter 6, through various analyses of the neuronal data.

The 7x7 grid maze as a platform for quantitative prediction testing

As mentioned in Chapter 2, the story of the grid maze starts in Lisbon, at the Champalimaud Foundation, around a decade ago. However, in the course of this project, we have implemented its most reliable and controlled version so far. Its intrinsic grid design allows for the formulation of smart questions and quantifiable

predictions of behavioural and neuronal responses while, in parallel, delivering extremely rich datasets that can be explored to inform better models and theories.

This platform is not restricted to the experimental setup. It is also now comprised of a structured pre-processing pipeline that enables same-day investigation of the signals collected, which can be crucial to speed up the scientific quest but also to inform experimental decisions in subsequent days; as well as a series of graphical interfaces that permit real-time filtering and visualisation of neuronal data for some of the analyses mentioned in this thesis.

With new replicas now operational in London, we are pleased to have established such a paradigm and predict it will continue being used (at least!) within the Akam and Behrens labs for the foreseeable future.

Mice rapidly adjust navigation strategy to changes in the environment

The first step to understanding the representations associated with any kind of behaviour is to understand the behaviour itself. In a goal-directed spatial navigation scenario, we draw from the RL framework and its approximation to the dichotomy between habitual and goal-directed behaviour in the form of model-free and model-based strategies (Sutton and Barto, 1998; Daw et al., 2005) to, in turn, approximate them to vector and shortest path navigation, respectively. After navigating the maze for randomly allocated rewards, the animals' underlying strategy

was revealed to be a mixture of both these strategies, with increasing use of the internal knowledge of the environment over time when compared to simply following the light to reward (Chapter 4). Furthermore, mice successfully incorporate the information of one-link changes in the environment within the first session of experience. This was essential to guide the schedule of the recording's experiments and, more relevantly, corroborated the existence of a cognitive map for navigation in the mouse brain.

dCA1 representations as states of a compositional system for navigation

Hippocampal cell representations are less straightforwardly interpretable with the increase of the environmental complexity (Chapter 5). While we would have hypothesised them to exhibit multiple firing fields given the dimension of the space to explore, we did not predict that some of these neurons would have such a strong directional dependence. Nor did we expect "place cells" to encode in an organised fashion along certain routes, partially (more commonly) or in its entirety (more rarely). These representations seem to reflect a conjunction of spatial and internal features and policies - location and type of location (platform or bridge), the direction of movement, particular movements (e.g. turning), and entire routes composed of commonly used sub-trajectories in the maze. These findings have inspired a compositional interpretation of the animal's navigation where these routes

function like a motorway - a state that one can enter and leave at many locations of the provided space -, and the trajectory to a specific goal is composed of multiple transitions between the different route representations available (or motorways) in the HPC. This idea led to a collaboration with another PhD student in the lab who took the lead in further investigating the animals' behaviour, finding the essential routes that shape their navigation beyond a simple vector/shortest path navigation trade-off. Encouragingly for our hypothesis, the bases of neural space we find in this thesis often mirror the bases of behavioural space.

Online replay in HPC does not reflect planning of future trajectories

One of the most intriguing types of hippocampal representations is replay. Or all forms of it. Pointed as a substrate for memory consolidation, credit assignment (and subsequent policy adjustments) and planning of future or imagined useful future states to visit, the role of replay in the support of cognitive maps is still not fully understood. Here, we tested ripple events for the ideas of planning and credit assignment at times of quiescent activity in the maze before running to the goal location and at reward consumption, respectively. While we could not find evidence for planning future trajectories or consolidation of previous steps, a positive correlation between the ripple power at reward consumption and the length of the path traversed to arrive at the reward location suggests a possible correlation between

the encoded ripple activity and/or content and the trajectories taken to goal. We intend to further explore this question.

Increase in SWRs during sleep facilitates performance of goal-directed navigation on a complex maze

While we could not yet assess the correlation between the decoded sleep maze maps and the route changes in the mouse behaviour, we assessed the relation between ripple features, i.e. number of events per sleep session, and the repercussions it has in mouse posterior performance in the same maze structure. Evidence suggests that offline replay events support learning and experience encoding in the brain, which reflects in performance improvement. We intend to also test for other ripple properties such as duration, as it has been observed to also boost memory of rodents in spatial contexts (Fernández-Ruiz et al., 2019).

Open questions and future directions

This project aimed to contribute to a vast body of literature that addresses the research of cognitive maps through spatial navigation. While we hope our contribution is relevant to the understanding of hippocampal representations, there are many open questions left unanswered in this thesis.

First, it would be valuable to quantitatively characterise the hippocampal cell population recorded. We have observed a minority of single-field cells and a sig-

nificant proportion of multi-field directionally selective cells that seem to integrate various levels of relevant information, but we have not yet quantified these groups. Additionally, we tried parsing the information delivered in these complex representations using dimensionality reduction analyses, uncovering factors resembling behavioural routes components but simultaneously uncovering hardly interpretable patterns. This could also not be solved using a variational auto-encoder. Developing an unbiased probabilistic method capable of highlighting the distinct components of this mixed selectivity at the single neuron level (a good start could be gathering inspiration from Hardcastle et al. (2017) and, at the population level from Pillow et al. (2008)) ought to be done to avoid constrained classification of cells in the HPC to pre-defined assumptions humans can imagine by eye.

Second, the evaluation of response to changes at the neuronal level has not been thoroughly defined in terms of hippocampal remapping. A major drawback in our recordings approach is the impossibility of tracking the same cells across the entire experiment (as a 2-photon microscopy technique would allow). Nonetheless, we can consistently compare neural responses within the same day and attempt across-day pattern comparison by re-sorting the spike data across multiple days, or using unit matching software (Van Beest et al., 2024).

Third, we have not managed to establish a putative role of online and offline replay. Besides fixing the LFP-based decoder and comparing detected events for this and the spike-based approach, replay analysis should be refined to uncover the relevant features depicted using this mechanism. It is essential to analyse the ripple activity content by considering the temporal axis along with the overall structure-

specific activity we investigated. Testing the classifier presented in Denovellis et al. (2021) might provide a smoother decoding over time. Moreover, we will investigate replay modulation under the modes of the common behavioural strategy evaluated (optimisation of the number of steps to goal) and concurrently through the glasses of a compositional strategy that combines routes more efficiently (i.e. through execution of fewer transitions, combination of fewer routes), as perhaps these allow for the accommodation of behavioural peculiarities that mice exhibit when running in a grid world.

Additionally, the lack of manual cell curation and the choice of cells to analyse based on a pre-defined number of spikes/session that could have excluded relevant cells for the performance of the task are two simple factors that could be tweaked in future analysis to improve the quality of the results.

Experimental considerations should also be made to better answer these questions. For instance, some of the best characteristics of this task to avoid habitual behaviour and prediction of rewarded or changed locations were the same to constrain model testing due to the scarcity of data points in shared conditions - this applied to behaviour and neuronal analysis. Likewise, the breadth of states-spaces a mouse can visit in one session together with the randomness of reward location and a fixed session duration made impossible for animals to visit the entirety of space in every session, and in turn this has severely underpowered our ability to perform certain quantitative analysis. To understand the contributions of EC inputs to the hippocampal representations of space, a simple suggestion would be to simultaneously record from these two regions in this task. We expect the EC

to encode the low-dimensional structure of this task, however we do not yet know what that looks like in a structured environment like this maze.

Modifications to the setup to allow automated movement of the bridges would allow the exploration of another timescale of remapping and the comparison of the role of replay in a slow-changing space vs a fast-changing one. We predict that, as seen in humans (Eldar et al., 2020), successful policy updates will be guided mainly by online ripple events, and that a failure to react as flexibly on the fly should be compensated with an increased number of offline events.

Lastly, finding a better way to test for a compositional theory of spatial navigation can be done in this setup. In an example task, the mouse is presented with multiple rewarded locations at once, knowing that in each trial it should collect all rewards in (1) a specific order or (2) in any desired order. This would allow the observation of deliberate composition of available routes when constrained differently by the policy of the environment, besides structure. On another note, we have also anecdotal evidence that mice can perform a memory version of the maze, which can, together with closed-loop ephys and stimulation systems, provide a good platform to asses replay causality in performance.

Bibliography

Aggleton, J. P., O'Mara, S. M., Vann, S. D., Wright, N. F., Tsanov, M., and Erichsen, J. T. Hippocampal–anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *European Journal of Neuroscience*, 31(12):2292–2307, June 2010. ISSN 0953-816X, 1460-9568. doi: 10.1111/j.1460-9568.2010.07251.x.

Akam, T., Costa, R., and Dayan, P. Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task. *PLOS Computational Biology*, 11(12):e1004648, Dec. 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004648.

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R. F., Dayan, P., and Costa, R. M. The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron*, 109(1):149–163.e7, Jan. 2021. ISSN 10974199. doi: 10.1016/j.neuron.2020.10.013. Publisher: Cell Press.

Alexander, A. S. and Nitz, D. A. Spatially Periodic Activation Patterns of Retrosplenial Cortex Encode Route Sub-spaces and Distance Traveled. *Current Biology*,

27(11):1551–1560.e4, June 2017. ISSN 09609822. doi: 10.1016/j.cub.2017.04.036.

Amaral, D. and Witter, M. The three-dimensional organization of the hippocampal formation: A review of anatomical data. *Neuroscience*, 31(3):571–591, Jan. 1989. ISSN 03064522. doi: 10.1016/0306-4522(89)90424-7.

Amaral, D. G., Scharfman, H. E., and Lavenex, P. The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies). In *Progress in Brain Research*, volume 163, pages 3–790. Elsevier, 2007. ISBN 978-0-444-53015-8. doi: 10.1016/S0079-6123(07)63001-5.

Ambrose, R. E., Pfeiffer, B. E., and Foster, D. J. Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. *Neuron*, 91(5):1124–1136, Sept. 2016. ISSN 10974199. doi: 10.1016/j.neuron.2016.07.047. Publisher: Cell Press.

Andersen, Morris, R. G. M., Amaral, D., Bliss, T., and O'Keefe, J. *The Hippocampus Book*. Oxford University Press, Oxford, 2007.

Anderson, J. S., Carandini, M., and Ferster, D. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *Journal of neurophysiology*, 84(2):909–26, Aug. 2000. ISSN 0022-3077. ISBN: 1033310344.

Aronov, D., Nevers, R., and Tank, D. W. Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. *Nature*, 543(7647):719–722, Mar. 2017. ISSN 14764687. doi: 10.1038/nature21692. Publisher: Nature Publishing Group.

Aston-Jones, G. and Cohen, J. D. AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal

Performance. *Annual Review of Neuroscience*, 28(1):403–450, July 2005. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev.neuro.28.061604.135709.

Bakermans, J. J., Warren, J., Whittington, J. C., and Behrens, T. E. Constructing future behaviour in the hippocampal formation through composition and replay. *bioRxiv*, Apr. 2023. doi: 10.1101/2023.04.07.536053.

Balleine, B. W. and Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5):407–419, Apr. 1998. ISSN 00283908. doi: 10.1016/S0028-3908(98)00033-1.

Balleine, B. W. and O'Doherty, J. P. Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology*, 35(1):48–69, Jan. 2010. ISSN 0893-133X, 1740-634X. doi: 10.1038/npp.2009.131.

Bannerman, D. M., Good, M. A., Yee, B. K., Heupel, M. J., Iversen, S. D., and Rawlins, J. N. Double dissociation of function within the hippocampus: A comparison of dorsal, ventral, and complete hippocampal cytotoxic lesions. *Behavioral Neuroscience*, 113(6):1170–1188, 1999. ISSN 07357044. doi: 10.1037/0735-7044.113.6.1170.

Barnes, C. A. Memory deficits associated with senescence: A neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, 93(1):74–104, 1979. ISSN 0021-9940. doi: 10.1037/h0077579.

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O'Keefe, J., Jeffery, K., and Burgess, N. The Boundary Vector Cell Model of Place Cell Firing and Spatial Memory. *Reviews in the Neurosciences*, 17(1-2), Jan. 2006. ISSN 2191-0200,

0334-1763. doi: 10.1515/REVNEURO.2006.17.1-2.71.

Barry, C., Hayman, R., Burgess, N., and Jeffery, K. J. Experience-dependent rescaling of entorhinal grids. *Nature Neuroscience*, 10(6):682–684, June 2007. ISSN 10976256. doi: 10.1038/nn1905.

Barry, C., Lin, L., Keefe, J. O., and Burgess, N. Grid cell fi ring patterns signal environmental novelty by expansion. 109(43):17687–17692, 2012. doi: 10.1073/pnas.1209918109.

Barto, A. and Dayan, P. Improving Generalization for Temporal Difference Learning: The Successor Representation. Technical report, 1995.

Bartolo, R. and Averbeck, B. B. Prefrontal Cortex Predicts State Switches during Reversal Learning. *Neuron*, 106(6):1044–1054.e4, June 2020. ISSN 08966273. doi: 10.1016/j.neuron.2020.03.024.

Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., and Kurth-Nelson, Z. What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, 100(2):490–509, Oct. 2018. ISSN 10974199. doi: 10.1016/j.neuron.2018.10.002. Publisher: Cell Press.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10 (9):1214–1221, Sept. 2007. ISSN 1097-6256. doi: 10.1038/nn1954. arXiv: NIHMS150003 ISBN: 1097-6256 (Print)\n1097-6256 (Linking).

Bicanski, A. and Burgess, N. A neural-level model of spatial memory and imagery. *eLife*, 7:e33752, Sept. 2018. ISSN 2050-084X. doi: 10.7554/eLife.33752.

Blodgett, H. *The effect of the introduction of reward upon the maze performance*

of rats, 4(8):113–134, 1929.

Bostock, E., Muller, R. U., and Kubie, J. L. Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 1(2):193–205, Apr. 1991. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.450010207.

Bradfield, L. A., Leung, B. K., Boldt, S., Liang, S., and Balleine, B. W. Goal-directed actions transiently depend on dorsal hippocampus. *Nature Neuroscience*, 23 (10):1194–1197, Oct. 2020. ISSN 15461726. doi: 10.1038/s41593-020-0693-8. Publisher: Nature Research.

Brogden, W. J. Sensory pre-conditioning. *Journal of Experimental Psychology*, 25 (4):323–332, Sept. 1939. ISSN 0022-1015. doi: 10.1037/h0058944.

Bunsey, M. and Eichenbaum, H. Conservation of hippocampal memory function in rats and humans. *Nature*, 379(6562):255–257, Jan. 1996. ISSN 0028-0836, 1476-4687. doi: 10.1038/379255a0.

Burak, Y. and Fiete, I. R. Accurate Path Integration in Continuous Attractor Network Models of Grid Cells. *PLoS Computational Biology*, 5(2):e1000291, Feb. 2009. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000291.

Burgess, N. Grid cells and theta as oscillatory interference: Theory and predictions. *Hippocampus*, 18(12):1157–1174, Dec. 2008. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20518.

Burgess, N., Recce, M., and O'Keefe, J. A model of hippocampal function. *Neural Networks*, 7(6-7):1065–1081, Jan. 1994. ISSN 08936080. doi: 10.1016/S0893-6080(05)80159-5.

Burgess, N., Maguire, E. A., and O'Keefe, J. The human hippocampus and spatial

and episodic memory. *Neuron*, 35(4):625–641, Aug. 2002. ISSN 0896-6273. doi: 10.1016/s0896-6273(02)00830-9.

Burwell, R. D. The Parahippocampal Region: Corticocortical Connectivity. *Annals of the New York Academy of Sciences*, 911(1):25–42, June 2000. ISSN 0077-8923, 1749-6632. doi: 10.1111/j.1749-6632.2000.tb06717.x.

Bush, D., Barry, C., Manson, D., and Burgess, N. Using Grid Cells for Navigation. *Neuron*, 87(3):507–520, Aug. 2015. ISSN 10974199. doi: 10.1016/j.neuron.2015.07.006. Publisher: Cell Press.

Buzsáki, G. Two-stage model of memory trace formation: A role for "noisy" brain states. *Neuroscience*, 31(3):551–570, Jan. 1989. ISSN 03064522. doi: 10.1016/0306-4522(89)90423-5.

Buzsáki, G. Theta Oscillations in the Hippocampus. *Neuron*, 33(3):325–340, Jan. 2002. ISSN 08966273. doi: 10.1016/S0896-6273(02)00586-X.

Byrne, P., Becker, S., and Burgess, N. Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114(2):340–375, Apr. 2007. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.114.2.340.

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., Flores, S. E., Kim, I., Sano, Y., Zhou, M., Baumgaertel, K., Lavi, A., Kamata, M., Tuszynski, M., Mayford, M., Golshani, P., and Silva, A. J. A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, 534(7605):115–118, June 2016. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature17955.

Carey, A. A., Tanaka, Y., and Meer, M. A. A. V. D. content away from the preferred outcome. *Nature Neuroscience*, 22(September), 2019. ISSN 1546-1726. doi: 10.1038/s41593-019-0464-6. Publisher: Springer US.

Carr, M. F., Jadhav, S. P., and Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, 14(2):147–153, Feb. 2011. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.2732.

Chen, G., King, J. A., Burgess, N., and O'Keefe, J. How vision and movement combine in the hippocampal place code. *Proceedings of the National Academy of Sciences*, 110(1):378–383, Jan. 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1215834110.

Chen, X., Vieweg, P., and Wolbers, T. Computing distance information from landmarks and self-motion cues - Differential contributions of anterior-lateral vs. posterior-medial entorhinal cortex in humans. *NeuroImage*, 202:116074, Nov. 2019. ISSN 10538119. doi: 10.1016/j.neuroimage.2019.116074.

Cheng, S. and Frank, L. M. New Experiences Enhance Coordinated Neural Activity in the Hippocampus. *Neuron*, 57(2):303–313, Jan. 2008. ISSN 0896-6273. doi: 10.1016/J.NEURON.2007.11.035. Publisher: Cell Press.

Clark, R. E. and Squire, L. R. Similarity in form and function of the hippocampus in rodents, monkeys, and humans. *Proceedings of the National Academy of Sciences*, 110(supplement_2):10365–10370, June 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1301225110.

Cohen, N. J. and Eichenbaum, H. *Memory, amnesia, and the hippocampal system*.

A Bradford book. MIT Press, Cambridge, Mass., 2. print edition, 1994. ISBN 978-0-262-53132-0 978-0-262-03203-2.

Cohen, N. J. and Squire, L. R. Preserved Learning and Retention of Pattern-Analyzing Skill in Amnesia: Dissociation of Knowing How and Knowing That. *Science*, 210(4466):207–210, Oct. 1980. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.7414331.

Conrad, C. D., Galea, L. A. M., Kuroda, Y., and McEwen, B. S. Chronic stress impairs rat spatial memory on the Y maze, and this effect is blocked by tianeptine treatment. *Behavioral Neuroscience*, 110(6):1321–1334, 1996. ISSN 1939-0084, 0735-7044. doi: 10.1037/0735-7044.110.6.1321.

Constantinescu, A. O., OReilly, J. X., and Behrens, T. E. J. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, June 2016. ISSN 0036-8075. doi: 10.1126/science.aaf0941.

Costa, V., Dal Monte, O., Lucas, D., Murray, E., and Averbeck, B. Amygdala and Ventral Striatum Make Distinct Contributions to Reinforcement Learning. *Neuron*, 92(2):505–517, Oct. 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.09.025.

Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. *European Journal of Neuroscience*, 26(3):704–716, Aug. 2007. ISSN 0953-816X, 1460-9568. doi: 10.1111/j.1460-9568.2007.05684.x.

Cunningham, J. P. and Yu, B. M. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17(11):1500–1509, Oct. 2014. ISSN

15461726. doi: 10.1038/nn.3776. Publisher: Nature Publishing Group.

Cushman, J. D., Aharoni, D. B., Willers, B., Ravassard, P., Kees, A., Vuong, C., Popeney, B., Arisaka, K., and Mehta, M. R. Multisensory Control of Multimodal Behavior: Do the Legs Know What the Tongue Is Doing? *PLoS ONE*, 8(11): e80465, Nov. 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0080465.

Darwin, C. *On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life.* John Murray, London, 1859.

Darwin, C. *The descent of man, and selection in relation to sex*. John Murray, London, 1871.

Davidson, T. J., Kloosterman, F., and Wilson, M. A. Hippocampal Replay of Extended Experience. *Neuron*, 63(4):497–507, Aug. 2009. ISSN 08966273. doi: 10.1016/j.neuron.2009.07.027.

Daw, N. D., Niv, Y., and Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12):1704–1711, Dec. 2005. ISSN 10976256. doi: 10.1038/nn1560.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. Article Model-Based Influences on Humans ' Choices and Striatal Prediction Errors. *Neuron*, 69(6):1204–1215, 2011. ISSN 0896-6273. doi: 10.1016/j.neuron.2011.02.027. Publisher: Elsevier Inc.

Dayan, P. Improving Generalization for Temporal Difference Learning: The Successor Representation. *Neural Computation*, 5(4):613–624, 1993. ISSN 0899-7667. doi: 10.1162/neco.1993.5.4.613.

Deacon, R. M. J., Bannerman, D. M., Kirby, B. P., Croucher, A., and Rawlins, J.

N. P. Effects of cytotoxic hippocampal lesions in mice on a cognitive test battery. 133:57–68, 2002.

Deacon, T. W., Eichenbaum, H., Rosenberg, P., and Eckmann, K. W. Afferent connections of the perirhinal cortex in the rat. *The Journal of Comparative Neurology*, 220(2):168–190, Oct. 1983. ISSN 0021-9967. doi: 10.1002/cne.902200205. ISBN: 0021-9967 (Print)\r0021-9967 (Linking).

Deb, K. *Multi-objective optimisation using evolutionary algorithms*. Willey, Chischester, UK, 2001.

Denovellis, E. L., Gillespie, A. K., Coulter, M. E., Sosa, M., Chung, J. E., Eden, U. T., and Frank, L. M. Hippocampal replay of experience at real-world speeds. *eLife*, 10:e64505, Sept. 2021. ISSN 2050-084X. doi: 10.7554/eLife.64505.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-b., and Moser, E. I. Fragmentation of grid cell maps in a multicompartment environment. *Nature Neuroscience*, 12(10):1325–1332, 2009. ISSN 1097-6256. doi: 10.1038/nn.2396. Publisher: Nature Publishing Group.

Deshmukh, S. S. and Knierim, J. J. Representation of Non-Spatial and Spatial Information in the Lateral Entorhinal Cortex. *Frontiers in Behavioral Neuroscience*, 5, 2011. ISSN 1662-5153. doi: 10.3389/fnbeh.2011.00069.

Deuker, L., Bellmund, J. L., Navarro Schröder, T., and Doeller, C. F. An event map of memory space in the hippocampus. *eLife*, 5:e16534, Oct. 2016. ISSN 2050-084X. doi: 10.7554/eLife.16534.

Diba, K. and Buzsáki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, 10(10):1241–1242, Oct. 2007. ISSN

10976256. doi: 10.1038/nn1961.

Dickinson, A. Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 308(1135):67–78, Feb. 1985. ISSN 0080-4622, 2054-0280. doi: 10.1098/rstb.1985.0010.

Dickinson, A., Nicholas, D. J., and Adams, C. D. The Effect of the Instrumental Training Contingency on Susceptibility to Reinforcer Devaluation. *The Quarterly Journal of Experimental Psychology Section B*, 35(1b):35–51, Feb. 1983. ISSN 0272-4995, 1464-1321. doi: 10.1080/14640748308400912.

Doeller, C. F., Barry, C., and Burgess, N. Evidence for grid cells in a human memory network. *Nature*, 463(7281):657–661, Feb. 2010. ISSN 0028-0836. doi: 10.1038/nature08704. Publisher: Nature Publishing Group.

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., and Daw, N. D. Model-based choices involve prospective neural activity. 18(5), 2015. doi: 10.1038/nn.3981.

Dolleman-van Der Weel, M. J., Griffin, A. L., Ito, H. T., Shapiro, M. L., Witter, M. P., Vertes, R. P., and Allen, T. A. The nucleus reuniens of the thalamus sits at the nexus of a hippocampus and medial prefrontal cortex circuit enabling memory and behavior. *Learning & Memory*, 26(7):191–205, July 2019. ISSN 1549-5485. doi: 10.1101/lm.048389.118.

Dordek, Y., Soudry, D., Meir, R., and Derdikman, D. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *eLife*, 5(MARCH2016):1–36, 2016. ISSN 2050084X. doi: 10.7554/eLife.10094.

Douglas, R. J. The hippocampus and behavior. *Psychological Bulletin*, 67(6):416–

442, 1967. ISSN 1939-1455, 0033-2909. doi: 10.1037/h0024599.

Doya, K. Modulators of decision making. *Nature Neuroscience*, 11(4):410–416, Apr. 2008. ISSN 1097-6256. doi: 10.1038/nn2077. ISBN: 1097-6256 (Print)\r1097-6256 (Linking).

Dragoi, G. and Buzsáki, G. Temporal Encoding of Place Sequences by Hippocampal Cell Assemblies. *Neuron*, 50(1):145–157, Apr. 2006. ISSN 08966273. doi: 10.1016/j.neuron.2006.02.023.

Dragoi, G. and Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*, 469(7330):397–401, Jan. 2011. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature09633.

Dupret, D., Revest, J.-M., Koehl, M., Ichas, F., De Giorgi, F., Costet, P., Abrous, D. N., and Piazza, P. V. Spatial Relational Memory Requires Hippocampal Adult Neurogenesis. *PLoS ONE*, 3(4):e1959, Apr. 2008. ISSN 1932-6203. doi: 10.1371/journal.pone.0001959.

Dupret, D., O'Neill, J., Pleydell-Bouverie, B., and Csicsvari, J. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature Neuroscience*, 13(8):995–1002, Aug. 2010. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.2599.

Dusek, J. A. and Eichenbaum, H. The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America*, 94(13):7109–7114, June 1997. ISSN 0027-8424. doi: 10.1073/pnas.94.13.7109.

Ego-Stengel, V. and Wilson, M. A. Disruption of ripple-associated hippocampal

activity during rest impairs spatial learning in the rat. *Hippocampus*, 20(1):1–10, Jan. 2010. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20707.

Eichenbaum, H. The hippocampus and mechanisms of declarative memory. *Behavioural Brain Research*, 103(2):123–133, Sept. 1999. ISSN 01664328. doi: 10.1016/S0166-4328(99)00044-3.

Eichenbaum, H. and Cohen, N. J. Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron*, 83(4):764–770, Aug. 2014. ISSN 1097-4199. doi: 10.1016/j.neuron.2014.07.032.

Eichenbaum, H., Yonelinas, A., and Ranganath, C. The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, 30(1):123–152, July 2007. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev.neuro.30.051606.094328.

Ekstrom, A., Meltzer, J., McNaughton, B., and Barnes, C. NMDA Receptor Antagonism Blocks Experience-Dependent Expansion of Hippocampal "Place Fields". *Neuron*, 31(4):631–638, Aug. 2001. ISSN 08966273. doi: 10.1016/S0896-6273(01)00401-9.

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., and Fried, I. Cellular networks underlying human spatial navigation. *Nature*, 425(6954):184–188, Sept. 2003. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature01964.

Eldar, E., Lièvre, G., Dayan, P., and Dolan, R. J. The roles of online and offline replay in planning. *eLife*, 9:e56911, June 2020. ISSN 2050-084X. doi: 10.7554/eLife.56911.

Eliav, T., Maimon, S. R., Aljadeff, J., Tsodyks, M., Ginosar, G., Las, L., and Ulanovsky, N. Multiscale representation of very large environments in the hippocampus of flying bats. *Science*, 372(6545):eabg4020, May 2021. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.abg4020.

Epstein, R. A. and Vass, L. K. Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20120533, Feb. 2014. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2012.0533.

Erdem, U. M. and Hasselmo, M. E. A biologically inspired hierarchical goal directed navigation model. *Journal of Physiology-Paris*, 108(1):28–37, Feb. 2014. ISSN 09284257. doi: 10.1016/j.jphysparis.2013.07.002.

Euston, D., Gruber, A., and McNaughton, B. The Role of Medial Prefrontal Cortex in Memory and Decision Making. *Neuron*, 76(6):1057–1070, Dec. 2012. ISSN 08966273. doi: 10.1016/j.neuron.2012.12.002.

Euston, D. R., Tatsuno, M., and McNaughton, B. L. Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex During Sleep. *Science*, 318 (5853):1147–1150, Nov. 2007. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1148979.

Fenton, A. A., Kao, H.-Y., Neymotin, S. A., Olypher, A., Vayntrub, Y., Lytton, W. W., and Ludvig, N. Unmasking the CA1 Ensemble Place Code by Exposures to Small and Large Environments: More Place Cells and Multiple, Irregularly Arranged, and Expanded Place Fields in the Larger Space. *The Journal of Neuroscience*, 28(44):11250–11262, Oct. 2008. ISSN 0270-6474, 1529-2401. doi:

10.1523/JNEUROSCI.2862-08.2008.

Fernández-Ruiz, A., Oliva, A., Fermino De Oliveira, E., Rocha-Almeida, F., Tingley, D., and Buzsáki, G. Long-duration hippocampal sharp wave ripples improve memory. *Science*, 364(6445):1082–1086, June 2019. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aax0758.

Foster, D. J. Replay Comes of Age. *Annual Review of Neuroscience*, 40 (1):581–602, July 2017. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev-neuro-072116-031538.

Foster, D. J. and Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683, Mar. 2006. ISSN 14764687. doi: 10.1038/nature04587. Publisher: Nature Publishing Group.

Fox, G. B., Fan, L., LeVASSEUR, R. A., and Faden, A. I. Effect of Traumatic Brain Injury on Mouse Spatial and Nonspatial Learning in the Barnes Circular Maze. *Journal of Neurotrauma*, 15(12):1037–1046, Dec. 1998. ISSN 0897-7151, 1557-9042. doi: 10.1089/neu.1998.15.1037.

Frank, L. M., Stanley, G. B., and Brown, E. N. Hippocampal Plasticity across Multiple Days of Exposure to Novel Environments. *The Journal of Neuroscience*, 24(35):7681–7689, Sept. 2004. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1958-04.2004.

Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., and Moser, E. I. Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, 446(7132):190–194, Mar. 2007. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature05601.

Gaffan, D. Scene-Specific Memory for Objects: A Model of Episodic Memory Impairment in Monkeys with Fornix Transection. *Journal of Cognitive Neuroscience*, 6(4):305–320, July 1994. ISSN 0898-929X, 1530-8898. doi: 10.1162/jocn.1994.6.4.305.

Gallagher, M., Burwell, R., and Burchinal, M. R. Severity of spatial learning impairment in aging: Development of a learning index for performance in the Morris water maze. *Behavioral Neuroscience*, 107(4):618–626, 1993. ISSN 1939-0084, 0735-7044. doi: 10.1037/0735-7044.107.4.618.

Gardner, R. S., Uttaro, M. R., Fleming, S. E., Suarez, D. F., Ascoli, G. A., and Dumas, T. C. A secondary working memory challenge preserves primary place strategies despite overtraining. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 20(11):648–656, Oct. 2013. ISSN 1549-5485. doi: 10.1101/lm.031336.113.

Gauthier, J. L. and Tank, D. W. Article A Dedicated Population for Reward Coding in the Article A Dedicated Population for Reward Coding in the Hippocampus. *Neuron*, 99(1):179–193.e7, 2018. ISSN 0896-6273. doi: 10.1016/j.neuron.2018.06.008. Publisher: Elsevier Inc.

Geerts, J. P., Gershman, S. J., Burgess, N., and Stachenfeld, K. L. A probabilistic successor representation for context-dependent learning. *Psychological Review*, 131(2):578–597, Mar. 2024. ISSN 1939-1471, 0033-295X. doi: 10.1037/rev0000414.

Gershman, S. J., Radulescu, A., Norman, K. A., and Niv, Y. Statistical Computations Underlying the Dynamics of Memory Updating. *PLoS Computational Biology*, 10(11):e1003939, Nov. 2014. ISSN 1553-7358. doi: 10.1371/journal.pcbi.

1003939.

Gershman, X. S. J. The Successor Representation : Its Computational Logic and Neural Substrates. 38(33):7193–7200, 2018. doi: 10.1523/JNEUROSCI.0151-18.2018.

Geva-Sagiv, M., Las, L., Yovel, Y., and Ulanovsky, N. Spatial cognition in bats and rats: from sensory acquisition to multiscale maps and navigation. *Nature Reviews Neuroscience*, 16(2):94–108, Feb. 2015. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn3888.

Gil, M., Ancau, M., Schlesiger, M. I., Neitz, A., Allen, K., De Marco, R. J., and Monyer, H. Impaired path integration in mice with disrupted grid cell firing. *Nature Neuroscience*, 21(1):81–91, Jan. 2018. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-017-0039-3.

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., and Zugaro, M. B. Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, 12(10):1222–1223, Oct. 2009. ISSN 1097-6256. doi: 10.1038/nn.2384. Publisher: Nature Publishing Group.

Goshen, I., Brodsky, M., Prakash, R., Wallace, J., Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. Dynamics of Retrieval Strategies for Remote Memories. *Cell*, 147(3):678–689, Oct. 2011. ISSN 00928674. doi: 10.1016/j.cell.2011.09.033.

Greene, E. and Stauff, C. Behavioral role of hippocampal connections. *Experimental Neurology*, 45(1):141–160, Oct. 1974. ISSN 00144886. doi: 10.1016/0014-4886(74)90107-1.

Grieves, R. M. and Jeffery, K. J. The representation of space in the brain. *Be-*

havioural Processes, 135:113–131, Feb. 2017. ISSN 18728308. doi: 10.1016/j.beproc.2016.12.012. Publisher: Elsevier B.V.

Gupta, A. S., Van Der Meer, M. A., Touretzky, D. S., and Redish, A. D. Hippocampal Replay Is Not a Simple Function of Experience. *Neuron*, 65(5):695–705, Mar. 2010. ISSN 08966273. doi: 10.1016/j.neuron.2010.01.034.

Gustafson, N. J. and Daw, N. D. Grid Cells , Place Cells , and Geodesic Generalization for Spatial Reinforcement Learning. 7(10), 2011. doi: 10.1371/journal.pcbi.1002235.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., and Moser, E. I. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005. ISSN 00280836. doi: 10.1038/nature03721.

Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. The Role of the Ventromedial Prefrontal Cortex in Abstract State-Based Inference during Decision Making in Humans. *Journal of Neuroscience*, 26(32):8360–8367, Aug. 2006. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.1010-06.2006. ISBN: 1529-2401 (Electronic)\r0270-6474 (Linking).

Hardcastle, K., Maheswaranathan, N., Ganguli, S., Giocomo, L. M., Hardcastle, K., Maheswaranathan, N., Ganguli, S., and Giocomo, L. M. Article A Multiplexed , Heterogeneous , and Adaptive Code for Navigation in Medial Entorhinal Cortex Article A Multiplexed , Heterogeneous , and Adaptive Code for Navigation in Medial Entorhinal Cortex. *Neuron*, 94(2):375–387.e7, 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2017.03.025. Publisher: Elsevier Inc.

Harrison, F. E., Reiserer, R. S., Tomarken, A. J., and Mcdonald, M. P. Spatial and

nonspatial escape strategies in the Barnes maze. pages 809–819, 2006. doi: 10.1101/lm.334306.tors.

Hartley, T., Burgess, N., Lever, C., Cacucci, F., and O'Keefe, J. Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus*, 10(4): 369–379, 2000. ISSN 1050-9631, 1098-1063. doi: 10.1002/1098-1063(2000)10:4⟨369::AID-HIPO3⟩3.0.CO;2-0.

Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, 104(5):1726–1731, Jan. 2007. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0610561104.

Hasselmo, M. E. Grid cell mechanisms and function: Contributions of entorhinal persistent spiking and phase resetting. *Hippocampus*, 18(12):1213–1229, Dec. 2008. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20512.

Hasselmo, M. E. and Stern, C. E. Theta rhythm and the encoding and retrieval of space and time. *NeuroImage*, 85:656–666, Jan. 2014. ISSN 10538119. doi: 10.1016/j.neuroimage.2013.06.022.

Hasz, B. M. and Redish, A. D. Deliberation and Procedural Automation on a Two-Step Task for Rats. *Frontiers in Integrative Neuroscience*, 12:30, Aug. 2018. ISSN 1662-5145. doi: 10.3389/fnint.2018.00030.

Hinman, J. R., Brandon, M. P., Climer, J. R., Chapman, G. W., and Hasselmo, M. E. Multiple Running Speed Signals in Medial Entorhinal Cortex. *Neuron*, 91 (3):666–679, Aug. 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.06.027.

Hock, B. J. and Bunsey, M. D. Differential Effects of Dorsal and Ventral Hippocam-

pal Lesions. *The Journal of Neuroscience*, 18(17):7027–7032, Sept. 1998. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.18-17-07027.1998.

Holland, J. H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Complex adaptive systems. MIT Press, Cambridge, Mass, 1st mit press ed edition, 1992. ISBN 978-0-262-08213-6 978-0-262-58111-0.

Hoover, W. B. and Vertes, R. P. Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain Structure and Function*, 212(2):149–179, Sept. 2007. ISSN 1863-2653, 1863-2661. doi: 10.1007/s00429-007-0150-4.

Howard, L., Javadi, A., Yu, Y., Mill, R., Morrison, L., Knight, R., Loftus, M., Staskute, L., and Spiers, H. The Hippocampus and Entorhinal Cortex Encode the Path and Euclidean Distances to Goals during Navigation. *Current Biology*, 24(12):1331–1340, June 2014. ISSN 09609822. doi: 10.1016/j.cub.2014.05.001.

Høydal, Ø. A., Skytøen, E. R., Andersson, S. O., Moser, M.-B., and Moser, E. I. Object-vector coding in the medial entorhinal cortex. *Nature*, 568(7752):400–404, Apr. 2019. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-019-1077-7.

Hölscher, C. Time, Space and Hippocampal Functions. *Reviews in the Neurosciences*, 14(3), Jan. 2003. ISSN 2191-0200, 0334-1763. doi: 10.1515/REVNEURO.2003.14.3.253.

Isaacson, R. L. *The Limbic System*. Springer US, Boston, MA, 1982. ISBN 978-1-4757-6757-5 978-1-4757-6755-1. doi: 10.1007/978-1-4757-6755-1.

Jacobs, J. Hippocampal theta oscillations are slower in humans than in rodents:

implications for models of spatial navigation and memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20130304, Feb. 2014. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2013.0304.

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X. X., Suthana, N., Sperling, M. R., Sharan, A. D., Fried, I., and Kahana, M. J. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*, 16(9):1188–1190, Sept. 2013. ISSN 10976256. doi: 10.1038/nn.3466.

Jadhav, S., Rothschild, G., Roumis, D., and Frank, L. Coordinated Excitation and Inhibition of Prefrontal Ensembles during Awake Hippocampal Sharp-Wave Ripple Events. *Neuron*, 90(1):113–127, Apr. 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.02.010.

Jadhav, S. P., Kemere, C., German, P. W., and Frank, L. M. Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory. *Science*, 336(6087):1454–1458, June 2012. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1217230.

Jankowski, M. M., Ronnqvist, K. C., Tsanov, M., Vann, S. D., Wright, N. F., Erichsen, J. T., Aggleton, J. P., and O'Mara, S. M. The anterior thalamus provides a subcortical circuit supporting memory and spatial navigation. *Frontiers in Systems Neuroscience*, 7, 2013. ISSN 1662-5137. doi: 10.3389/fnsys.2013.00045.

Jepma, M. and Nieuwenhuis, S. Pupil Diameter Predicts Changes in the Exploration–Exploitation Trade-off: Evidence for the Adaptive Gain Theory. *Journal of Cognitive Neuroscience*, 23(7):1587–1596, July 2011. ISSN 0898-929X, 1530-8898. doi: 10.1162/jocn.2010.21548.

Johnson, A. and Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45): 12176–12189, Nov. 2007. ISSN 02706474. doi: 10.1523/JNEUROSCI.3761-07.2007.

Johnson, A., Seeland, K., and Redish, A. D. Reconstruction of the postsubiculum head direction signal from neural ensembles. *Hippocampus*, 15(1):86–96, 2005. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20033.

Johnson, C. T., Olton, D. S., Gage, F. H., and Jenko, P. G. Damage to hippocampus and hippocampal connections: Effects on DRL and spontaneous alternation. *Journal of Comparative and Physiological Psychology*, 91(3):508–522, 1977. ISSN 0021-9940. doi: 10.1037/h0077346.

Karlsson, M. P. and Frank, L. M. Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, 12(7):913–918, July 2009. ISSN 1546-1726. doi: 10.1038/nn.2344.

Kay, K., Sosa, M., Chung, J. E., Karlsson, M. P., Larkin, M. C., and Frank, L. M. A hippocampal network for spatial coding during immobility and sleep. *Nature*, 531(7593):185–190, Mar. 2016. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature17144.

Kay, K., Chung, J. E., Sosa, M., Schor, J. S., Karlsson, M. P., Larkin, M. C., Liu, D. F., and Frank, L. M. Constant Sub-second Cycling between Representations of Possible Futures in the Hippocampus. *Cell*, 180(3):552–567.e25, Feb. 2020. ISSN 10974172. doi: 10.1016/j.cell.2020.01.014. Publisher: Cell Press.

Keinath, A. T., Mosser, C.-A., and Brandon, M. P. The representation of context in

mouse hippocampus is preserved despite neural drift. *Nature Communications*, 13(1):2415, May 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-30198-7.

Kentros, C. G., Agnihotri, N. T., Streater, S., Hawkins, R. D., and Kandel, E. R. Increased Attention to Spatial Context Increases Both Place Field Stability and Spatial Memory. *Neuron*, 42(2):283–295, Apr. 2004. ISSN 08966273. doi: 10.1016/S0896-6273(04)00192-8.

Kim, D. and Park, H. Nonnegative matrix factorization based on alternating nonnegativity-constrained least squares and the active set method. In *Matrix Analysis and Aplications*, volume 23, pages 713–730. 2 edition, 2008.

Kimble, D. P. THE EFFECTS OF BILATERAL HIPPOCAMPAL LESIONS IN RATS 1. Technical report, 1963. Publication Title: Journal of Comparative and Physiological Psychology Volume: 56 Issue: 2.

Kimble, D. P. Hippocampus and internal inhibition. *Psychological Bulletin*, 70(5): 285–295, 1968. ISSN 1939-1455, 0033-2909. doi: 10.1037/h0026470.

Knierim, J., Kudrimoti, H., and McNaughton, B. Place cells, head direction cells, and the learning of landmark stability. *The Journal of Neuroscience*, 15(3):1648–1659, Mar. 1995. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.15-03-01648.1995.

Knierim, J. J., Lee, I., and Hargreaves, E. L. Hippocampal place cells: Parallel input streams, subregional processing, and implications for episodic memory. *Hippocampus*, 16(9):755–764, Sept. 2006. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20203.

Knudsen, E. B. and Wallis, J. D. Hippocampal neurons construct a map of an ab-

stract value space. *Cell*, 184(18):4640–4650.e10, Sept. 2021. ISSN 00928674. doi: 10.1016/j.cell.2021.07.010.

Koenig, S. and Simmons, R. G. The effect of representation and knowledge on goal-directed exploration with reinforcement-learning algorithms. *Machine Learning*, 22(1-3):227–250, 1996. ISSN 0885-6125, 1573-0565. doi: 10.1007/BF00114729.

Koster, R., Chadwick, M. J., Chen, Y., Berron, D., Banino, A., Düzel, E., Hassabis, D., and Kumaran, D. Big-Loop Recurrence within the Hippocampal System Supports Integration of Information across Episodes. *Neuron*, 99(6):1342–1354.e6, Sept. 2018. ISSN 08966273. doi: 10.1016/j.neuron.2018.08.009.

Kraus, B., Robinson, R., White, J., Eichenbaum, H., and Hasselmo, M. Hippocampal "Time Cells": Time versus Path Integration. *Neuron*, 78(6):1090–1101, June 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.04.015.

Krieckhaus, E. E. and Wolf, G. Acquisition of sodium by rats: Interaction of innate mechanisms and latent learning. *Journal of Comparative and Physiological Psychology*, 65(2):197–201, 1968. ISSN 0021-9940. doi: 10.1037/h0025547.

Kropff, E., Carmichael, J. E., Moser, M.-B., and Moser, E. I. Speed cells in the medial entorhinal cortex. *Nature*, 523(7561):419–424, July 2015. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature14622.

Krupic, J., Burgess, N., and O'Keefe, J. Neural Representations of Location Composed of Spatially Periodic Bands. *Science*, 337(6096):853–857, Aug. 2012. ISSN 0036-8075. doi: 10.1126/science.1222403.

Krupic, J., Bauza, M., Burton, S., and O'Keefe, J. Framing the grid: effect of

boundaries on grid cells and navigation. *The Journal of Physiology*, 594(22): 6489–6499, Nov. 2016. ISSN 0022-3751, 1469-7793. doi: 10.1113/JP270607.

Kubie, J. L. and Fenton, A. A. Linear Look-Ahead in Conjunctive Cells: An Entorhinal Mechanism for Vector-Based Navigation. *Frontiers in Neural Circuits*, 6, 2012. ISSN 1662-5110. doi: 10.3389/fncir.2012.00020.

Kudrimoti, H. S., Barnes, C. A., and McNaughton, B. L. Reactivation of Hippocampal Cell Assemblies: Effects of Behavioral State, Experience, and EEG Dynamics. *The Journal of Neuroscience*, 19(10):4090–4101, May 1999. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.19-10-04090.1999.

Kumaran, D. and McClelland, J. L. Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, 119(3):573–616, July 2012. ISSN 1939-1471, 0033-295X. doi: 10.1037/a0028681.

Kumaran, D., Summerfield, J. J., Hassabis, D., and Maguire, E. A. Tracking the Emergence of Conceptual Knowledge during Human Decision Making. *Neuron*, 63(6):889–901, Sept. 2009. ISSN 08966273. doi: 10.1016/j.neuron.2009.07.030.

Kurth-nelson, Z., Economides, M., Dolan, R. J., Dayan, P., Kurth-nelson, Z., Economides, M., Dolan, R. J., and Dayan, P. Fast Sequences of Non-spatial State Representations in Humans Article Fast Sequences of Non-spatial State Representations in Humans. *Neuron*, 91(1):194–204, 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.05.028. Publisher: The Authors.

Köhler, W. *The mentality of apes*. 1917.

Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., Moser, E. I., and Moser, M.-B. Development of the Spatial Representation System in the Rat. *Science*, 328(5985):1576–1580, June 2010. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1188210.

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., and Pennartz, C. M. A. Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25):6372–6382, June 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1054-08.2008.

Lashley, K. In search of the engram. (In Society for Experimental Biology, Physiological mechanisms in animal behavior.):454–482, 1950.

Lavenex, P. and Amaral, D. G. Hippocampal-neocortical interaction: A hierarchy of associativity. *Hippocampus*, 10(4):420–430, 2000. ISSN 1050-9631, 1098-1063. doi: 10.1002/1098-1063(2000)10:4⟨420::AID-HIPO8⟩3.0.CO;2-5.

Lee, A. K. and Wilson, M. A. Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep. *Neuron*, 36(6):1183–1194, Dec. 2002. ISSN 08966273. doi: 10.1016/S0896-6273(02)01096-6.

Lee, S., Shimojo, S., and O'Doherty, J. Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron*, 81(3):687–699, Feb. 2014. ISSN 08966273. doi: 10.1016/j.neuron.2013.11.028.

Leutgeb, J. K., Leutgeb, S., Treves, A., Meyer, R., Barnes, C. A., McNaughton, B. L., Moser, M. B., and Moser, E. I. Progressive transformation of hippocampal neuronal representations in "morphed" environments. *Neuron*, 48(2):345–358,

2005. ISSN 08966273. doi: 10.1016/j.neuron.2005.09.007.

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., and Moser, E. I. Pattern Separation in the Dentate Gyrus and CA3 of the Hippocampus. *Science*, 315(5814):961–966, Feb. 2007. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1135801.

Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M.-B., and Moser, E. I. Distinct Ensemble Codes in Hippocampal Areas CA3 and CA1. *Science*, 305(5688): 1295–1298, Aug. 2004. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1100265.

Lever, C., Wills, T., Cacucci, F., Burgess, N., and O'Keefe, J. Long-term plasticity in hippocampal place-cell representation of environmental geometry. *Nature*, 416 (6876):90–94, Mar. 2002. ISSN 0028-0836, 1476-4687. doi: 10.1038/416090a.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., and Burgess, N. Boundary Vector Cells in the Subiculum of the Hippocampal Formation. *The Journal of Neuroscience*, 29(31):9771–9777, Aug. 2009. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1319-09.2009.

Lopes, G., Bonacchi, N., FrazÃ£o, J., Neto, J. P., Atallah, B. V., Soares, S., Moreira, L., Matias, S., Itskov, P. M., Correia, P. A., Medina, R. E., Calcaterra, L., Dreosti, E., Paton, J. J., and Kampff, A. R. Bonsai: an event-based framework for processing and controlling data streams. *Frontiers in Neuroinformatics*, 9, Apr. 2015. ISSN 1662-5196. doi: 10.3389/fninf.2015.00007.

Louie, K. and Wilson, M. A. Temporally Structured Replay of Awake Hippocampal Ensemble Activity during Rapid Eye Movement Sleep. *Neuron*, 29(1):145–156, Jan. 2001. ISSN 08966273. doi: 10.1016/S0896-6273(01)00186-6.

Lubbock, S. J. *On the senses, instincts, and intelligence of animals, with special reference to insects*. Trench & Company, London, k. paul edition, 1889.

MacDonald, C. J., Carrow, S., Place, R., and Eichenbaum, H. Distinct Hippocampal Time Cell Sequences Represent Odor Memories in Immobilized Rats. *The Journal of Neuroscience*, 33(36):14607–14616, Sept. 2013. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1537-13.2013.

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., and Leutgeb, J. K. Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, 109(47):19462–19467, Nov. 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1214107109.

Markus, E., Qin, Y., Leonard, B., Skaggs, W., McNaughton, B., and Barnes, C. Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *The Journal of Neuroscience*, 15(11):7079–7094, Nov. 1995. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.15-11-07079.1995.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M. user-defined body parts with deep learning. *Nature Neuroscience*, 21(September), 2018. ISSN 1546-1726. doi: 10.1038/s41593-018-0209-y. Publisher: Springer US.

Mattar, M. G. and Daw, N. D. Prioritized memory access explains planning and hippocampal replay. Technical report, 2018.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: Insights from

the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457, July 1995. ISSN 1939-1471. doi: 10.1037/0033-295X.102.3.419.

McDonald, R. J. and White, N. M. A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, 107(1): 3–22, 1993. ISSN 1939-0084, 0735-7044. doi: 10.1037/0735-7044.107.1.3.

McHugh, T. J., Jones, M. W., Quinn, J. J., Balthasar, N., Coppari, R., Elmquist, J. K., Lowell, B. B., Fanselow, M. S., Wilson, M. A., and Tonegawa, S. Dentate Gyrus NMDA Receptors Mediate Rapid Pattern Separation in the Hippocampal Network. *Science*, 317(5834):94–99, July 2007. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1140263.

McNaughton, B., Barnes, C., and O'Keefe, J. The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. *Experimental Brain Research*, 52(1), Sept. 1983. ISSN 0014-4819, 1432-1106. doi: 10.1007/BF00237147.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M. B. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, 2006. ISSN 1471003X. doi: 10.1038/nrn1932.

Mehta, M. R., Barnes, C. A., and McNaughton, B. L. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94(16):8918–8921, Aug. 1997. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.94.16.8918.

Miller, K. J., Botvinick, M. M., and Brody, C. D. Dorsal hippocampus contributes to

model-based planning. 20(9), 2017. doi: 10.1038/nn.4613.

Minderer, M., Harvey, C. D., Donato, F., and Moser, E. I. Virtual reality explored. *Nature*, 533(7603):324–325, May 2016. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature17899.

Mitchell, M. *An Introduction to Genetic Algorithms*. The MIT Press, 1998. ISBN 978-0-262-28001-3. doi: 10.7551/mitpress/3927.001.0001.

Moita, M. A. P., Rosis, S., Zhou, Y., LeDoux, J. E., and Blair, H. T. Putting Fear in Its Place: Remapping of Hippocampal Place Cells during Fear Conditioning. *The Journal of Neuroscience*, 24(31):7015–7023, Aug. 2004. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.5492-03.2004.

Moore, A. W. and Atkeson, C. G. Prioritized sweeping: Reinforcement learning with less data and less time. *Machine Learning*, 13(1):103–130, Oct. 1993. ISSN 0885-6125, 1573-0565. doi: 10.1007/BF00993104.

Morgan, C. L. In - *ON THE STUDY OF ANIMAL INTELLIGENCE*, volume II. Mind os-XI, 1886.

Morgan, 1852-1936, C. L. *An introduction to comparative psychology*. Scribner, New York, scott edition, 1894.

Morris, R. G. M., Garrud, P., Rawlins, J. N. P., and O'Keefe, J. Place navigation impaired in rats with hippocampal lesions. *Nature*, 297(5868):681–683, June 1982. ISSN 0028-0836, 1476-4687. doi: 10.1038/297681a0.

Moser, E. I., Kropff, E., and Moser, M.-B. Place Cells, Grid Cells, and the Brain's Spatial Representation System. *Annual Review of Neuroscience*, 31(1):69–89, July 2008. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev.neuro.31.

061307.090723.

Muller, R., Bostock, E., Taube, J., and Kubie, J. On the directional firing properties of hippocampal place cells. *The Journal of Neuroscience*, 14(12):7235–7251, Dec. 1994. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.14-12-07235.1994.

Muller, R. U. and Kubie, J. L. The Effects of Changes in the Environment on the Spatial Firing of Hippocampal Complex-Spike Cells. Technical report, 1987. Publication Title: The Journal of Neuroscience Volume: 7 Issue: 7.

Nieh, E. H., Schottdorf, M., Freeman, N. W., Low, R. J., Lewallen, S., Koay, S. A., Pinto, L., Gauthier, J. L., Brody, C. D., and Tank, D. W. Geometry of abstract learned knowledge in the hippocampus. *Nature*, 595(7865):80–84, July 2021. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-021-03652-7.

Nitz, D. A. Path shape impacts the extent of CA1 pattern recurrence both within and across environments. *Journal of Neurophysiology*, 105(4):1815–1824, Apr. 2011. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00573.2010.

Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., and Buzsáki, G. Replay and Time Compression of Recurring Spike Sequences in the Hippocampus. *The Journal of Neuroscience*, 19(21):9497–9507, Nov. 1999. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.19-21-09497.1999.

O'Keefe, J. Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, 51(1):78–109, Jan. 1976. ISSN 00144886. doi: 10.1016/0014-4886(76)90055-8. ISBN: 0014-4886.

O'Keefe, J. A review of the hippocampal place cells. *Progress in Neurobiology*, 13

(4):419–439, Jan. 1979. ISSN 03010082. doi: 10.1016/0301-0082(79)90005-4.

O'Keefe, J. and Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature*, 381(6581):425–428, May 1996. ISSN 0028-0836, 1476-4687. doi: 10.1038/381425a0.

O'Keefe, J. and Burgess, N. Dual phase and rate coding in hippocampal place cells: Theoretical significance and relationship to entorhinal grid cells. *Hippocampus*, 15(7):853–866, Jan. 2005. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20115.

O'Keefe, J. and Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–175, Nov. 1971. ISSN 00068993. doi: 10.1016/0006-8993(71)90358-1.

O'Keefe, J. and Nadel, L. *The Hippocampus as a Cognitive Map*. Oxford University Press, Oxford, UK, 1978.

Olton, D. S. and Papas, B. C. Spatial memory and hippocampal function. *Neuropsychologia*, 17(6):669–682, Jan. 1979. ISSN 00283932. doi: 10.1016/0028-3932(79)90042-3.

Omer, D. B., Maimon, S. R., Las, L., and Ulanovsky, N. Social place-cells in the bat hippocampus. *Science*, 359(6372):218–224, Jan. 2018. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aao3474.

Ormond, J., Serka, S. A., and Johansen, J. P. Enhanced Reactivation of Remapping Place Cells during Aversive Learning. *The Journal of Neuroscience*, 43 (12):2153–2167, Mar. 2023. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1450-22.2022.

O'Neill, J., Boccara, C., Stella, F., Schoenenberger, P., and Csicsvari, J. Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. *Science*, 355(6321):184–188, Jan. 2017. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aag2787.

Pachitariu, M., Sridhar, S., Pennington, J., and Stringer, C. Spike sorting with Kilosort4. *Nature Methods*, Apr. 2024. ISSN 1548-7091, 1548-7105. doi: 10.1038/s41592-024-02232-7.

Packard, M. G. and McGaugh, J. L. Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. *Neurobiology of Learning and Memory*, 65(1):65–72, Jan. 1996. ISSN 10747427. doi: 10.1006/nlme.1996.0007.

Papale, A. E., Zielinski, M. C., Frank, L. M., Jadhav, S. P., and Redish, A. D. Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. *Neuron*, 92(5):975–982, Dec. 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.10.028.

Pareto, V. *Cours d'économie politique*. Lausanne, f. rouge edition, 1896.

Park, E. H., Dvorak, D., and Fenton, A. A. Ensemble place codes in hippocampus: CA1, CA3, and dentate gyrus place cells have multiple place fields in large environments. *PLoS ONE*, 6(7), 2011. ISSN 19326203. doi: 10.1371/journal.pone.0022349.

Passecker, J., Mikus, N., Malagon-Vina, H., Anner, P., Dimidschstein, J., Fishell, G., Dorffner, G., and Klausberger, T. Activity of Prefrontal Neurons Predict Future Choices during Gambling. *Neuron*, 101(1):152–164.e7, Jan. 2019. ISSN

08966273. doi: 10.1016/j.neuron.2018.10.050.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science*, 321(5894):1322–1327, Sept. 2008. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1159775.

Pavlides, C. and Winson, J. Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *The Journal of Neuroscience*, 9(8):2907–2918, Aug. 1989. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.09-08-02907.1989.

Penfield, W. and Rasmussen, T. *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*. Macmillan, Oxford, 1950.

Pennartz, C. M. A., Lee, E., Verheul, J., Lipa, P., Barnes, C. A., and McNaughton, B. L. The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by Hippocampal Ripples. *The Journal of Neuroscience*, 24(29):6446–6456, July 2004. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0575-04.2004.

Penny, W. D., Zeidman, P., and Burgess, N. Forward and Backward Inference in Spatial Cognition. *PLoS Computational Biology*, 9(12):e1003383, Dec. 2013. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1003383.

Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., and Battaglia, F. P. Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature Neuroscience*, 12(7):919–926, July 2009. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.2337.

Pezzulo, G., Rigoli, F., and Chersi, F. The Mixed Instrumental Controller: Using

Value of Information to Combine Habitual Choice and Mental Simulation. *Frontiers in Psychology*, 4, 2013. ISSN 1664-1078. doi: 10.3389/fpsyg.2013.00092.

Pfeiffer, B. E. and Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, May 2013. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature12112.

Phelps, E. A. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, 14(2):198–202, Apr. 2004. ISSN 09594388. doi: 10.1016/j.conb.2004.03.015.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999, Aug. 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature07140.

Pitkänen, A., Pikkarainen, M., Nurminen, N., and Ylinen, A. Reciprocal Connections between the Amygdala and the Hippocampal Formation, Perirhinal Cortex, and Postrhinal Cortex in Rat: A Review. *Annals of the New York Academy of Sciences*, 911(1):369–391, June 2000. ISSN 0077-8923, 1749-6632. doi: 10.1111/j.1749-6632.2000.tb06738.x.

Preston, A. R. and Eichenbaum, H. Interplay of Hippocampus and Prefrontal Cortex in Memory. *CURBIO*, 23(17):R764–R773, 2013. ISSN 0960-9822. doi: 10.1016/j.cub.2013.05.041. Publisher: Elsevier.

Preston, A. R., Shrager, Y., Dudukovic, N. M., and Gabrieli, J. D. Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, 14(2):148–152, Jan. 2004. ISSN 1050-9631, 1098-1063. doi:

10.1002/hipo.20009.

Qasim, S. E., Miller, J., Inman, C. S., Gross, R. E., Willie, J. T., Lega, B., Lin, J.-J., Sharan, A., Wu, C., Sperling, M. R., Sheth, S. A., McKhann, G. M., Smith, E. H., Schevon, C., Stein, J. M., and Jacobs, J. Memory retrieval modulates spatial tuning of single neurons in the human entorhinal cortex. *Nature Neuroscience*, 22(12):2078–2086, Dec. 2019. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-019-0523-z.

Quirk, G., Muller, R., and Kubie, J. The firing of hippocampal place cells in the dark depends on the rat's recent experience. *The Journal of Neuroscience*, 10(6):2008–2017, June 1990. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.10-06-02008.1990.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045):1102–1107, June 2005. ISSN 0028-0836. doi: 10.1038/nature03687.

Ranck, J. Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats. Technical report, Soc. Neurosci. Abstr. 10:599, 1984.

Redish, A. D. Vicarious trial and error. *Nature Reviews Neuroscience*, 17(3):147–159, Mar. 2016. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn.2015.30.

Richardson, R., Ledgerwood, L., and Cranney, J. Facilitation of Fear Extinction by D-Cycloserine: Theoretical and Clinical Implications. *Learning & Memory*, 11(5):510–516, Sept. 2004. ISSN 1072-0502, 1549-5485. doi: 10.1101/lm.78204.

Ritchey, M., Dolcos, F., and Cabeza, R. Role of Amygdala Connectivity in the Persistence of Emotional Memories Over Time: An Event-Related fMRI Investi-

gation. *Cerebral Cortex*, 18(11):2494–2504, Nov. 2008. ISSN 1047-3211, 1460-2199. doi: 10.1093/cercor/bhm262.

Rollenhagen, A. and Lübke, J. H. R. The morphology of excitatory central synapses: from structure to function. *Cell and Tissue Research*, 326 (2):221–237, Nov. 2006. ISSN 0302-766X, 1432-0878. doi: 10.1007/s00441-006-0288-z.

Rolls, E. T. Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, 9(4):467–480, 1999. ISSN 1050-9631, 1098-1063. doi: 10.1002/(SICI)1098-1063(1999)9:4⟨467::AID-HIPO13⟩3.0.CO;2-F.

Rolls, E. T. The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience*, 7, 2013. ISSN 1662-5137. doi: 10.3389/fnsys.2013.00074.

Romanes, G. J. *Animal Intelligence*. Trench & Company, London, k. paul edition, 1886.

Rozeboom, W. W. Secondary extinction of lever-pressing behavior in the albino rat. *Journal of Experimental Psychology*, 54(4):280–287, 1957. ISSN 0022-1015. doi: 10.1037/h0046073.

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., and Daw, N. D. *Predictive representations can link model- based reinforcement learning to model-free mechanisms*. 2017. ISBN 1-111-11111-1.

Sanders, H., Wilson, M. A., and Gershman, S. J. Hippocampal remapping as hidden state inference. *eLife*, 9:1–31, 2020. ISSN 2050084X. doi: 10.7554/eLife.51140.

Sarel, A. and Finkelstein, A. Vectorial representation of spatial goals in the hippocampus of bats. 180(January):176–180, 2017.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., and Moser, E. I. Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. *Science*, 312(5774):758–762, May 2006. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1125572.

Schmidt, B., Papale, A., Redish, A. D., and Markus, E. J. Conflict between place and response navigation strategies: Effects on vicarious trial and error (VTE) behaviors. *Learning & Memory*, 20(3):130–138, Mar. 2013. ISSN 1549-5485. doi: 10.1101/lm.028753.112.

Schmitt, W. B., Deacon, R. M. J., Seeburg, P. H., Rawlins, J. N. P., and Bannerman, D. M. A Within-Subjects, Within-Task Demonstration of Intact Spatial Reference Memory and Impaired Spatial Working Memory in Glutamate Receptor-A-Deficient Mice. *The Journal of Neuroscience*, 23(9):3953–3959, May 2003. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.23-09-03953.2003.

Schuck, N. W., Cai, M. B., Wilson, R. C., Niv, Y., Schuck, N. W., Cai, M. B., Wilson, R. C., and Niv, Y. Human Orbitofrontal Cortex Represents a Cognitive Map of State Space Article Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, 91(6):1402–1412, 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.08.019. Publisher: Elsevier Inc.

Schultz, W., Dayan, P., and Montague, P. R. A neural substrate of prediction and reward. *Science (New York, N.Y.)*, 275(5306):1593–9, Mar. 1997. ISSN 0036-8075.

Schwartenbeck, P., Baram, A., Liu, Y., Mark, S., Muller, T., Dolan, R., Botvinick, M., Kurth-Nelson, Z., and Behrens, T. Generative replay underlies compositional inference in the hippocampal-prefrontal circuit. *Cell*, 186(22):4885–4897.e14, Oct. 2023. ISSN 00928674. doi: 10.1016/j.cell.2023.09.004.

Scoville, W. B. and Milner, B. LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1):11–21, Feb. 1957. ISSN 0022-3050. doi: 10.1136/jnnp.20.1.11.

Shalev-Shwartz, S. and Wexler, Y. Minimizing the maximal loss: How and why. In Balcan, M. F. and Weinberger, K. Q., editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 793–801, New York, New York, USA, 20–22 Jun 2016. PMLR.

Siegle, J. H., López, A. C., Patel, Y. A., Abramov, K., Ohayon, S., and Voigts, J. Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *Journal of Neural Engineering*, 14(4):045003, Aug. 2017. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2552/aa5eea.

Simon, D. A. and Daw, N. D. Neural Correlates of Forward Planning in a Spatial Decision Task in Humans. *The Journal of Neuroscience*, 31(14):5526–5539, Apr. 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.4647-10.2011.

Singer, A., Carr, M., Karlsson, M., and Frank, L. Hippocampal SWR Activity Predicts Correct Decisions during the Initial Learning of an Alternation Task. *Neuron*, 77(6):1163–1173, Mar. 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.01.027.

Singer, A. C. and Frank, L. M. Rewarded Outcomes Enhance Reactivation of Experience in the Hippocampus. *Neuron*, 64(6):910–921, Dec. 2009. ISSN 08966273. doi: 10.1016/j.neuron.2009.11.016.

Sinnamon, H. M., Freniere, S., and Kootz, J. Rat hippocampus and memory for places of changing significance. *Journal of Comparative and Physiological Psychology*, 92(1):142–155, 1978. ISSN 0021-9940. doi: 10.1037/h0077446.

Sivanandam, S. and Deepa, S. Genetic Algorithms. In *Introduction to Genetic Algorithms*, pages 15–37. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-73189-4. doi: 10.1007/978-3-540-73190-0_2.

Skaggs, W. E. and McNaughton, B. L. Replay of Neuronal Firing Sequences in Rat Hippocampus During Sleep Following Spatial Experience. *Science*, 271(5257): 1870–1873, Mar. 1996. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.271.5257.1870.

Skaggs, W. E. and McNaughton, B. L. Spatial Firing Properties of Hippocampal CA1 Populations in an Environment Containing Two Visually Identical Regions. *The Journal of Neuroscience*, 18(20):8455–8466, Oct. 1998. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.18-20-08455.1998.

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., and Barnes, C. A. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus*, 6(2):149–172, 1996. ISSN 1050-9631, 1098-1063. doi: 10.1002/(SICI)1098-1063(1996)6:2⟨149::AID-HIPO6⟩3.0.CO;2-K.

Skinner, B. *Science and human behavior*. The Macmillan Company, New York, 1953.

Smittenaar, P., FitzGerald, T., Romei, V., Wright, N., and Dolan, R. Disruption of Dorsolateral Prefrontal Cortex Decreases Model-Based in Favor of Model-free Control in Humans. *Neuron*, 80(4):914–919, Nov. 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.08.009.

Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., and Moser, E. I. Representation of Geometric Borders in the Entorhinal Cortex. *Science*, 322(5909):1865–1868, Dec. 2008. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1166466.

Spiers, H. J. and Gilbert, S. J. Solving the detour problem in navigation: a model of prefrontal and hippocampal interactions. *Frontiers in Human Neuroscience*, 9, Mar. 2015. ISSN 1662-5161. doi: 10.3389/fnhum.2015.00125.

Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, Nov. 2017. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4650.

Stangl, M., Topalovic, U., Inman, C. S., Hiller, S., Villaroman, D., Aghajan, Z. M., Christov-Moore, L., Hasulak, N. R., Rao, V. R., Halpern, C. H., Eliashiv, D., Fried, I., and Suthana, N. Boundary-anchored neural mechanisms of location-encoding for self and others. *Nature*, 589(7842):420–425, Jan. 2021. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-020-03073-y.

Stemmler, M., Mathis, A., and Herz, A. V. M. Connecting multiple spatial scales to decode the population activity of grid cells. *Science Advances*, 1(11):e1500816, Dec. 2015. ISSN 2375-2548. doi: 10.1126/science.1500816.

Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B., and Moser, E. I. The entorhinal grid map is discretized. *Nature*, 492(7427):72–78, Dec. 2012.

ISSN 0028-0836, 1476-4687. doi: 10.1038/nature11649.

Sun, C., Yang, W., Martin, J., and Tonegawa, S. Hippocampal neurons represent events as transferable units of experience. *Nature Neuroscience*, 23(5):651–663, May 2020. ISSN 1097-6256. doi: 10.1038/s41593-020-0614-x. Publisher: Springer US.

Sutton, R. and Barto, A. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 9(5):1054–1054, Sept. 1998. ISSN 1045-9227, 1941-0093. doi: 10.1109/TNN.1998.712192.

Sutton, R. S. Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin*, 2(4):160–163, July 1991. ISSN 0163-5719. doi: 10.1145/122344.122377.

Suzuki, W. L. and Amaral, D. G. Perirhinal and parahippocampal cortices of the macaque monkey: Cortical afferents. *Journal of Comparative Neurology*, 350(4):497–533, Dec. 1994. ISSN 0021-9967, 1096-9861. doi: 10.1002/cne.903500402.

Taube, J. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *The Journal of Neuroscience*, 15(1):70–86, Jan. 1995. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.15-01-00070.1995.

Taube, J. S., Muller, R. U., and Ranck, J. B. Head-Direction Cells Recorded from the Postsubiculum in Freely Moving Rats. I. Description and Quantitative Analysis. Technical report, 1990. Publication Title: The Journal of Neuroscience Issue: 2.

Tavares, R., Mendelsohn, A., Grossman, Y., Williams, C., Shapiro, M., Trope, Y.,

and Schiller, D. A Map for Social Navigation in the Human Brain. *Neuron*, 87(1): 231–243, July 2015. ISSN 08966273. doi: 10.1016/j.neuron.2015.06.011.

Theves, S., Fernandez, G., and Doeller, C. F. The Hippocampus Encodes Distances in Multidimensional Feature Space. *Current Biology*, 29(7):1226–1231.e3, Apr. 2019. ISSN 09609822. doi: 10.1016/j.cub.2019.02.035.

Thorndike, E. L. Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2 (4):i–109, 1898. ISSN 0096-9753. doi: 10.1037/h0092987.

Thorndike, E. L. *Animal Intelligence*. New York, 1911.

Tolman, E. C. *Purposive behavior in animals and men*. Univ of California Press, 1932.

Tolman, E. C. Sign-Gestalt or conditioned reflex. *Psychological Review*, 40(3): 246–255, May 1933. ISSN 1939-1471, 0033-295X. doi: 10.1037/h0071472.

Tolman, E. C. Cognitive maps in rats and men. *Image and Environment: Cognitive Mapping and Spatial Behavior*, 55(4):189–208, 1948a. ISSN 0033-295X. doi: 10.4324/9780203789155-11. ISBN: 9781351513647.

Tolman, E. C. THE PSYCHOLOGICAL REVIEW COGNITIVE MAPS IN RATS AND MEN *. Technical report, 1948b. Volume: 55 Issue: 4.

Tolman, E. C. and Gleitman, H. Studies in learning and motivation: I. Equal reinforcements in both end-boxes, followed by shock in one end-box. *Journal of Experimental Psychology*, 39(6):810–819, 1949. ISSN 0022-1015. doi: 10.1037/h0062845.

Tolman, E. C., Ritchie, B. F., and Kalish, D. Studies in spatial learning. I. Orientation

and the short-cut. *Journal of Experimental Psychology*, 36(1):13–24, Feb. 1946. ISSN 0022-1015. doi: 10.1037/h0053944.

Tsien, J. Z., Huerta, P. T., and Tonegawa, S. The Essential Role of Hippocampal CA1 NMDA Receptor–Dependent Synaptic Plasticity in Spatial Memory. *Cell*, 87(7):1327–1338, Dec. 1996. ISSN 00928674. doi: 10.1016/S0092-8674(00) 81827-9.

Ulanovsky, N. and Moss, C. F. Hippocampal cellular and network activity in freely moving echolocating bats. *Nature Neuroscience*, 10(2):224–233, Feb. 2007. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn1829.

Ulanovsky, N. and Moss, C. F. Dynamics of hippocampal spatial representation in echolocating bats. *Hippocampus*, 21(2):150–161, Feb. 2011. ISSN 1050-9631, 1098-1063. doi: 10.1002/hipo.20731.

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., and Lega, B. Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences*, 117(45): 28463–28474, Nov. 2020. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas. 2013250117.

Van Beest, E. H., Bimbard, C., Fabre, J. M. J., Dodgson, S. W., Takács, F., Coen, P., Lebedeva, A., Harris, K. D., and Carandini, M. Tracking neurons across days with high-density probes. *Nature Methods*, Sept. 2024. ISSN 1548-7091, 1548-7105. doi: 10.1038/s41592-024-02440-1.

Van Der Meer, M. A. A., Kemere, C., and Diba, K. Progress and issues in second-order analysis of hippocampal replay. *Philosophical Transactions of the Royal*

Society B: Biological Sciences, 375(1799):20190238, May 2020. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2019.0238.

Vertechi, P., Lottem, E., Sarra, D., Godinho, B., Treves, I., Quendera, T., Oude Lohuis, M. N., and Mainen, Z. F. Inference-Based Decisions in a Hidden State Foraging Task: Differential Contributions of Prefrontal Cortical Areas. *Neuron*, 106 (1):166–176.e6, Apr. 2020. ISSN 10974199. doi: 10.1016/j.neuron.2020.01.017. Publisher: Cell Press.

Vertes, R. and Kocsis, B. Brainstem-diencephalo-septohippocampal systems controlling the theta rhythm of the hippocampus. *Neuroscience*, 81(4):893–926, Oct. 1997. ISSN 03064522. doi: 10.1016/S0306-4522(97)00239-X.

Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess, N., and Daw, N. D. Hippocampal Contributions to Model-Based Planning and Spatial Memory. *Neuron*, 102(3):683–693.e4, May 2019. ISSN 10974199. doi: 10.1016/j.neuron.2019.02.014. Publisher: Cell Press.

Warren, W. H., Rothman, D. B., Schnapp, B. H., and Ericson, J. D. Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition*, 166:152–163, Sept. 2017. ISSN 00100277. doi: 10.1016/j.cognition.2017.05.020.

Wehner, R. and Srinivasan, M. V. Searching behaviour of desert ants, genus-Cataglyphis (Formicidae, Hymenoptera). *Journal of Comparative Physiology ? A*, 142(3):315–338, 1981. ISSN 0340-7594, 1432-1351. doi: 10.1007/BF00605445.

Widloski, J. and Foster, D. J. Flexible rerouting of hippocampal replay sequences around changing barriers in the absence of global place field remapping. *Neu-*

ron, 110(9):1547–1558.e8, May 2022. ISSN 08966273. doi: 10.1016/j.neuron.2022.02.002.

Wiener, S., Paul, C., and Eichenbaum, H. Spatial and behavioral correlates of hippocampal neuronal activity. *The Journal of Neuroscience*, 9(8):2737–2763, Aug. 1989. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.09-08-02737.1989.

Wilent, W. B. and Nitz, D. A. Discrete Place Fields of Hippocampal Formation Interneurons. *Journal of Neurophysiology*, 97(6):4152–4161, June 2007. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.01200.2006.

Wills, T. J., Cacucci, F., Burgess, N., and O'Keefe, J. Development of the hippocampal cognitive map in preweanling rats. *Science*, 328(5985):1573–1576, June 2010. ISSN 00368075. doi: 10.1126/science.1188224.

Wilson, M. and McNaughton, B. Reactivation of hippocampal ensemble memories during sleep. *Science*, 265(5172):676–679, July 1994. ISSN 0036-8075. doi: 10.1126/science.8036517.

Wilson, M. A. and McNaughton, B. L. Dynamics of the Hippocampal Ensemble Code for Space. *Science*, 261(5124):1055–1058, Aug. 1993. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.8351520.

Wilson, R., Watkins, J. J., and Wilson, R. Combinatorics : Ancient and Modern. (October 2019):1–21, 2013. doi: 10.1093/acprof. ISBN: 9780199656592.

Witter, M., Groenewegen, H., Lopes Da Silva, F., and Lohman, A. Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region. *Progress in Neurobiology*, 33(3):161–253, Jan. 1989. ISSN 03010082. doi:

10.1016/0301-0082(89)90009-9.

Witter, M. P., Wouterlood, F. G., Naber, P. A., and Van Haeften, T. Anatomical Organization of the Parahippocampal-Hippocampal Network. Technical report, 2000.

Wood, R. A., Bauza, M., Krupic, J., Burton, S., Delekate, A., Chan, D., and Keefe, J. O. hippocampal-dependent spatial navigation. *Nature Publishing Group*, 554 (7690):102–105, 2018. ISSN 0028-0836. doi: 10.1038/nature25433. Publisher: Nature Publishing Group.

Yin, H., Knowlton, B., and Balleine, B. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behavioural Brain Research*, 166(2):189–196, Jan. 2006. ISSN 01664328. doi: 10.1016/j.bbr.2005.07.012.

Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, 22(2):513–523, July 2005. ISSN 0953-816X, 1460-9568. doi: 10.1111/j.1460-9568.2005.04218.x.

Yu, A. J. and Dayan, P. Uncertainty, Neuromodulation, and Attention. *Neuron*, 46 (4):681–692, May 2005. ISSN 08966273. doi: 10.1016/j.neuron.2005.04.026.

Zeithamova, D., Dominick, A. L., and Preston, A. R. Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, 75(1):168–179, July 2012. ISSN 1097-4199. doi: 10.1016/j.neuron.2012.05.010.

Zhu, S. L., Lakshminarasimhan, K. J., Arfaei, N., and Angelaki, D. E. Eye move-

ments reveal spatiotemporal dynamics of visually-informed planning in navigation. 2021. doi: 10.1101/2021.04.26.441482.

Zielinski, M. C. The role of replay and theta sequences in mediating hippocampal-prefrontal interactions for memory and cognition. (October 2017):60–72, 2020. doi: 10.1002/hipo.22821.

Ziv, Y., Burns, L. D., Cocker, E. D., Hamel, E. O., Ghosh, K. K., Kitch, L. J., Gamal, A. E., and Schnitzer, M. J. Long-term dynamics of CA1 hippocampal place codes. *Nature Neuroscience*, 16(3):264–266, Mar. 2013. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.3329.

Zola-Morgan, S., Squire, L., and Amaral, D. Lesions of the hippocampal formation but not lesions of the fornix or the mammillary nuclei produce long-lasting memory impairment in monkeys. *The Journal of Neuroscience*, 9(3):898–913, Mar. 1989. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.09-03-00898.1989.

Zugaro, M. B., Berthoz, A., and Wiener, S. I. Background, But Not Foreground, Spatial Cues Are Taken as References for Head Direction Responses by Rat Anterodorsal Thalamus Neurons. *The Journal of Neuroscience*, 21(14):RC154–RC154, July 2001. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.21-14-j0001.2001.

Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., and Spiers, H. J. Hippocampal place cells construct reward related sequences through unexplored space. *eLife*, 4:e06063, June 2015. ISSN 2050-084X. doi: 10.7554/eLife.06063.

Ólafsdóttir, H. F., Carpenter, F., and Barry, C. Coordinated grid and place cell replay

during rest. *Nature Neuroscience*, 19(6):792–794, Apr. 2016. ISSN 15461726. doi: 10.1038/nn.4291. Publisher: Nature Publishing Group.

Ólafsdóttir, H. F., Bush, D., and Barry, C. The Role of Hippocampal Replay in Memory and Planning. *Current biology: CB*, 28(1):R37–R50, Jan. 2018. ISSN 1879-0445. doi: 10.1016/j.cub.2017.10.073.

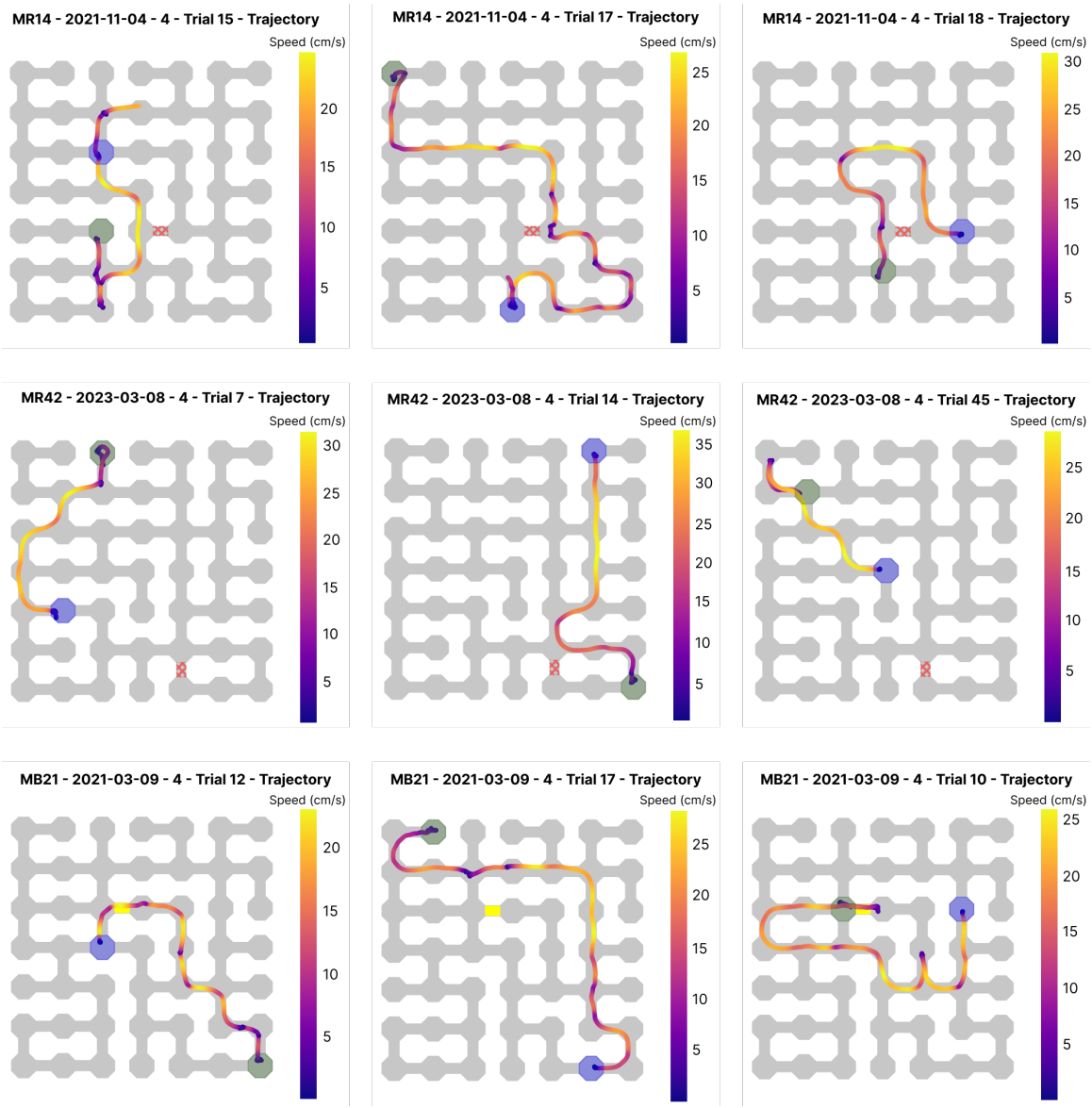
A

Supplements

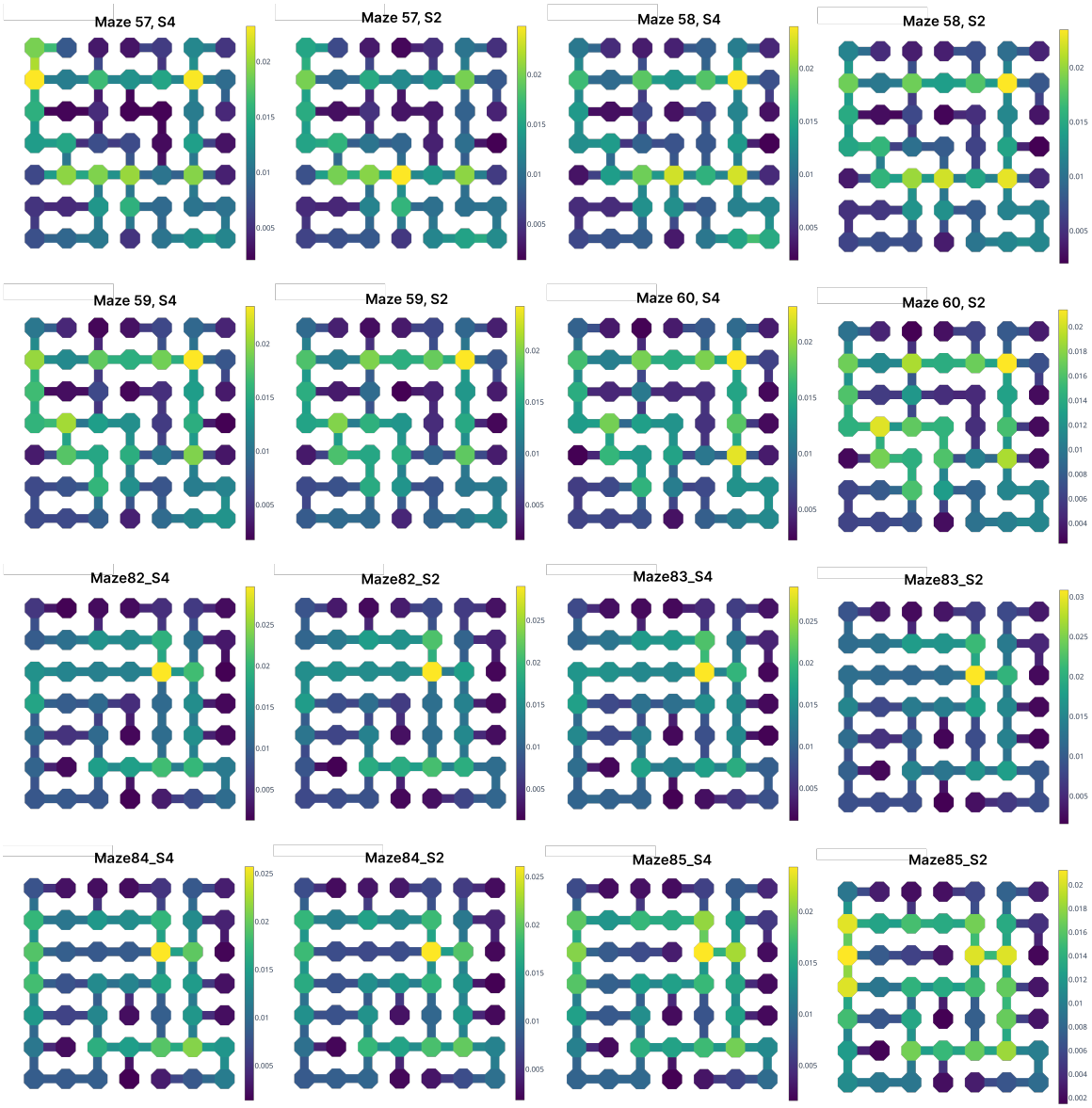
This section offers supplementary material that complements the main text and figures, providing more insights about this project. It contains the following sections:

- Chapter 4 - Mice quickly integrate new spatial information into their navigation strategy under a goal-directed paradigm - and other behavioural signatures
- Chapter 5 - Hippocampal representations in a complex and dynamic maze environment
- Chapter 6 - Online and offline replay of an intricate hippocampal cell coding system

Supplements to Chapter 4

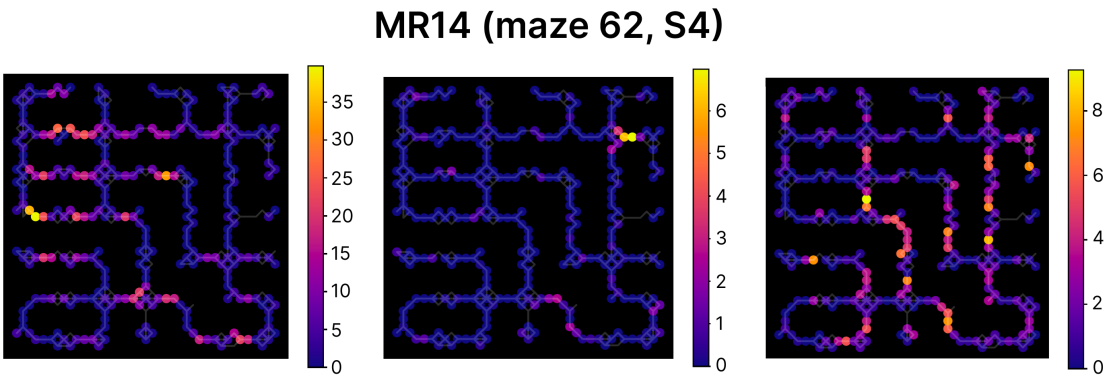


Supplementary Figure A.1: Other examples of navigation trials in the 7x7 grid maze. Three different mice for three different maze structures where a bridge has been removed (top and middle rows) or added (bottom row) in that session.

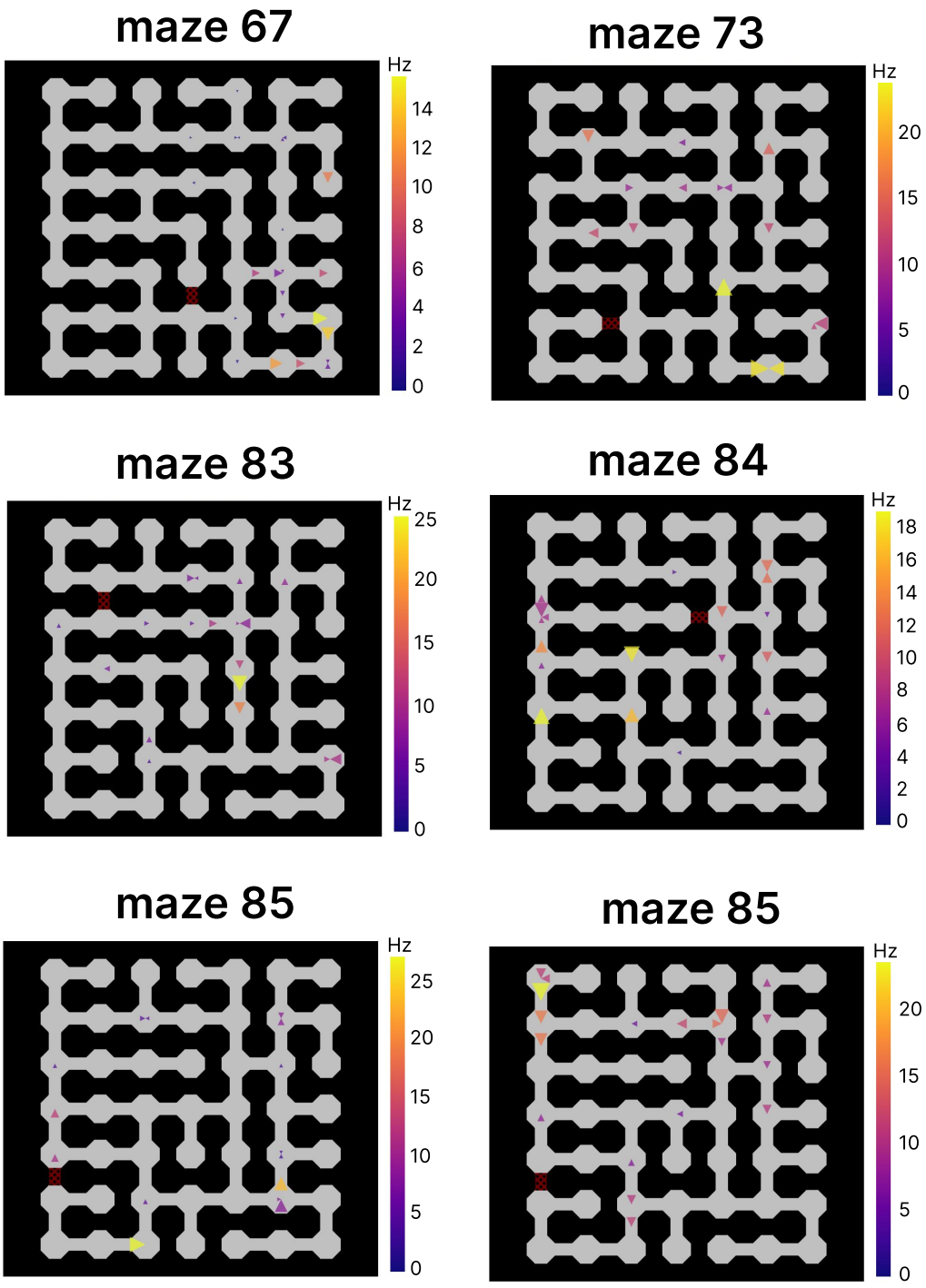


Supplementary Figure A.2: Other examples of visit occupancy across mice in distinct maze sessions. The top six maze sessions correspond to structures experienced by the mice in early days of experiment, while the six bottom maze sessions correspond to some of the last structures experienced across the duration of the experiment.

Supplements to Chapter 5

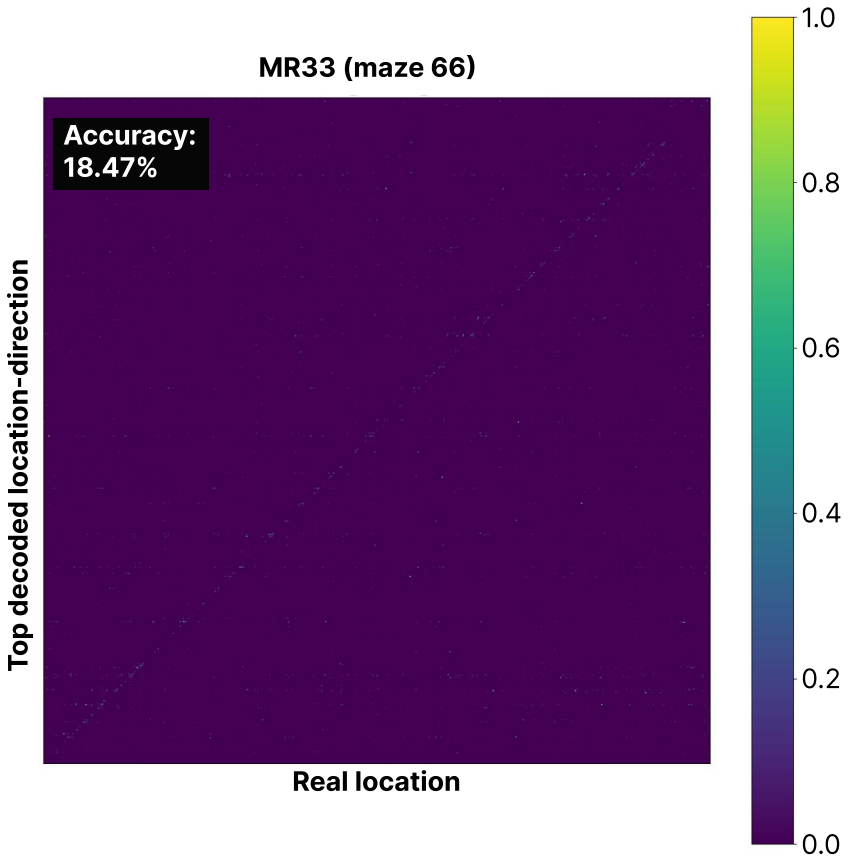


Supplementary Figure A.3: Other examples of hippocampal neuronal firing maps in the linearised version of the maze.

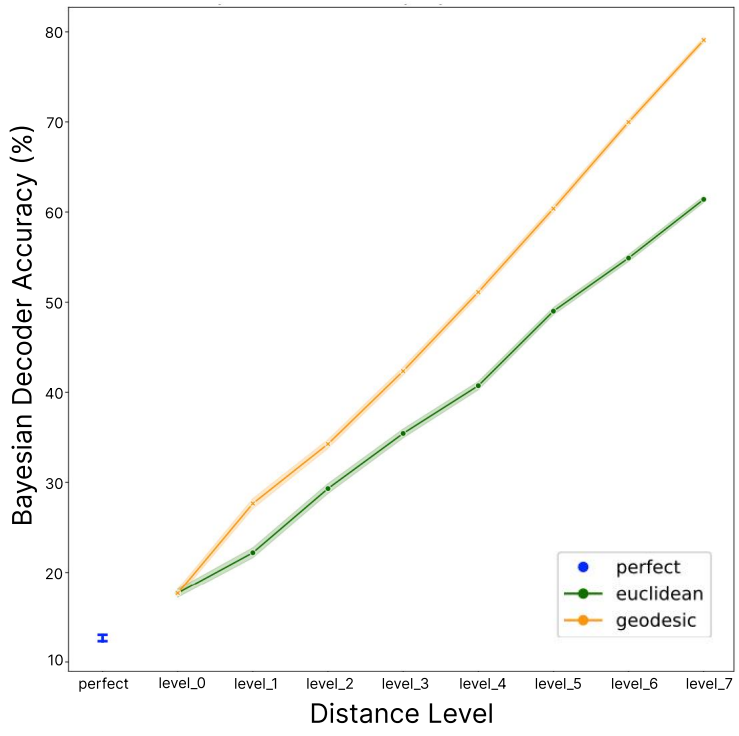


Supplementary Figure A.4: dCA1 neuronal directional firing maps in the 7x7 grid maze. The activity of the neurons presented elicits trajectories of subpaths along the maze structure for specific running orientations. The last neuron, however, seems to represent an approximation to a ”move inwards” in the maze policy, which we have seen across many neurons, at different scales. In these maps, colour represents the average firing rate at the location independent of direction (increasing activity from dark to light colours), and arrow size indicates directionally selective firing rate proportion in the direction the arrow is pointing to.

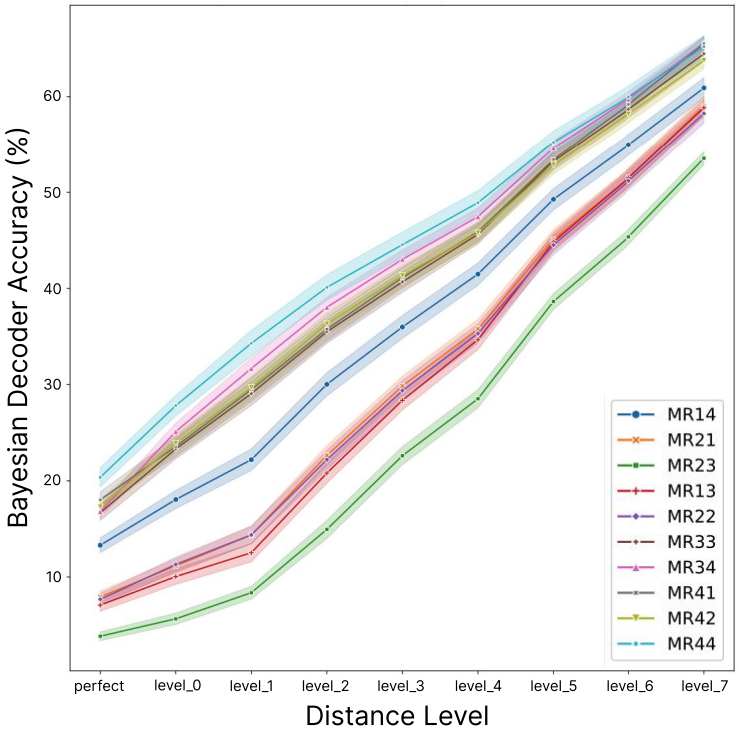
Supplements to Chapter 6



Supplementary Figure A.5: Recorded dCA1 neural activity allows decoding of the mouse location-direction (state action) pairs in the maze. Confusion matrices of animal decoded location-direction (node or bridge x travelling direction (N, S, E, W)) (y-axis) from neural activity of example mouse (MR33) for mazes 66; respective accuracy values are depicted for the decoded positions that match the real position of the mouse, with lighter colours indicating higher accuracy values. Chance level between 0.245-0.238%, for a minimum (408) and maximum (420) number of available states experienced.



a



b

Supplementary Figure A.6: Variation of Bayesian decoder accuracy for location-direction pairs. Variation of the mean (a) and mouse (b) Bayesian decoding accuracy considering perfect decoding (blue) and decoding of locations within various levels of Euclidean and geodesic distance from the real location. $Level_0$ corresponds to decoding the right location irrespective of the direction, and thus already exhibits increased activity vs perfect location-direction decoding.

Table A.1: Correlation coefficient statistics for mouse decoded location between D1S5-D1S3 ripple activity and D2S2-D1S4 behavioural occupancy, using D1S2 as basis for decoding. The absolute set of calculations is done considering the positive ripple differences and the absolute values of the relative occupation differences. For each mean correlation coefficient, we calculated the statistical t-test and correspondent p-value.

		real		absolute	
		t-test	p-value	t-test	p-value
all sessions	-	1.5622	1.202e-01	-4.4050	1.929e-05 ***
	current added edge	2.2609	2.644e-02 *	-2.6804	8.905e-03 **
	current removed edge	0.1255	9.005e-01	-3.5798	5.955e-04 ***
	previous added edge	0.1045	9.171e-01	-2.3080	2.372e-02 *
	previous removed edge	2.0087	4.795e-02 *	-3.8670	2.232e-04 ***
by maze	current added edge	-0.8504	1.550e-02 *	-3.1184	8.882e-03 **
	current removed edge	2.8186	4.094e-01	-2.4982	2.572e-02 *
	previous added edge	-0.6074	5.549e-01	-1.7530	1.051e-01
	previous removed edge	1.0228	3.251e-01	-2.9432	1.114e-02 *
by mouse	current added edge	2.4587	3.624e-02 *	-2.8125	2.030e-02 *
	current removed edge	0.5546	5.927e-01	-7.9806	2.257e-05 ***
	previous added edge	0.3668	7.223e-01	-2.8453	1.924e-02 *
	previous removed edge	1.8216	1.019e-01	-4.0920	2.710e-03 **

Table A.2: Correlation coefficient statistics for mouse decoded location-direction pairs between D1S5-D1S3 ripple activity and D2S2-D1S4 behavioural occupancy, using D1S2 as basis for decoding. The absolute set of calculations is done considering the positive ripple differences and the absolute values of the relative occupation differences. For each mean correlation coefficient, we calculated the statistical t-test and correspondent p-value.

		real		absolute	
		t-test	p-value	t-test	p-value
all sessions	-	1.2680	2.066e-01	-1.9555	5.227e-02
	current added edge	1.8878	6.263e-02	-0.3697	7.126e-01
	current removed edge	0.0224	9.822e-01	-2.6433	9.919e-03 **
	previous added edge	-0.4387	6.621e-01	-1.5530	1.246e-01
	previous removed edge	1.7432	8.513e-02	-0.9789	3.306e-01
by maze	current added edge	1.8447	8.989e-02	-0.6591	5.223e-01
	current removed edge	-0.5887	5.655e-01	-2.5792	2.184e-02 *
	previous added edge	-1.0329	3.220e-01	-1.9278	7.788e-02
	previous removed edge	1.3931	1.869e-01	-1.2840	2.215e-01
by mouse	current added edge	2.6483	2.655e-02 *	-0.6908	5.071e-01
	current removed edge	0.4590	6.571e-01	-2.8762	1.829e-02 *
	previous added edge	-0.0706	9.452e-01	-2.5379	3.182e-02 *
	previous removed edge	2.1380	6.122e-02	-1.4622	1.777e-01

Table A.3: Correlation coefficient statistics for mouse decoded location-direction pairs between D1S5 ripple activity and D2S2-D1S4 behavioural occupancy, using D1S4 as basis for decoding. The absolute set of calculations is done considering the positive ripple differences and the absolute values of the relative occupation differences. For each mean correlation coefficient, we calculated the statistical t-test and correspondent p-value.

		real		absolute	
		t-test	p-value	t-test	p-value
all sessions	-	-0.2835	7.772e-01	1.4367	1.530e-01
	current added edge	-0.5318	5.965e-01	0.2351	8.148e-01
	current removed edge	0.1250	9.009e-01	1.8843	6.362e-02
	previous added edge	0.1942	8.466e-01	-0.8908	3.761e-01
	previous removed edge	-0.6749	5.019e-01	2.7189	8.203e-03 **
by maze	current added edge	-0.5664	5.815e-01	0.4328	6.729e-01
	current removed edge	0.5352	6.009e-01	1.1453	2.713e-01
	previous added edge	0.4722	6.452e-01	-0.7575	4.634e-01
	previous removed edge	-0.8116	4.316e-01	0.8886	3.904e-01
by mouse	current added edge	-0.7649	4.639e-01	0.7164	4.919e-01
	current removed edge	0.1087	9.158e-01	2.1816	5.703e-02
	previous added edge	-0.1094	9.153e-01	-0.7197	4.900e-01
	previous removed edge	-0.5817	5.815e-01	4.2245	2.225e-03 **

Table A.4: Correlation coefficient statistics for mouse decoded location-direction pairs between D1S3 ripple activity and D2S2-D1S4 behavioural occupancy, using D1S2 as basis for decoding. The absolute set of calculations is done considering the positive ripple differences and the absolute values of the relative occupation differences. For each mean correlation coefficient, we calculated the statistical t-test and correspondent p-value.

		real		absolute	
		t-test	p-value	t-test	p-value
all sessions	-	-0.2835	7.772e-01	1.4367	1.530e-01
	current added edge	-0.5318	5.965e-01	0.2351	8.148e-01
	current removed edge	0.1250	9.009e-01	1.8843	6.362e-02
	previous added edge	0.1942	8.466e-01	-0.8908	3.761e-01
	previous removed edge	-0.6749	5.019e-01	2.7189	8.203e-03 **
by maze	current added edge	-0.5664	5.815e-01	0.4328	6.729e-01
	current removed edge	0.5352	6.009e-01	1.1453	2.713e-01
	previous added edge	0.4722	6.452e-01	-0.7575	4.634e-01
	previous removed edge	-0.8116	4.316e-01	0.8886	3.904e-01
by mouse	current added edge	-0.7649	4.639e-01	0.7164	4.919e-01
	current removed edge	0.1087	9.158e-01	2.1816	5.703e-02
	previous added edge	-0.1094	9.153e-01	-0.7197	4.900e-01
	previous removed edge	-0.5817	5.815e-01	4.2245	2.225e-03 **