Abstract

BAYESIAN M ECHANISMS IN SPATIAL COGNITION: TOWARDS REAL-WORLD CAPABLE COMPUTATIONAL COGNITIVE MODELS OF SPATIAL MEMORY

Tamas Madl

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy, 2015

Computational cognitive models of spatial memory often neglect difficulties posed by the real world, such as sensory noise, uncertainty, and high spatial complexity. However, since cognition and its neural bases have been shaped by the structure and challenges of the physical world, cognitive models should take these into account as well.

This work takes an interdisciplinary approach towards developing a cognitively plausible spatial memory model able to function in real-world environments, despite the sensory noise and high spatial complexity. We investigated how spatially rele- vant brain areas might maintain an accurate location estimate of mammals, despite accumulating sensory noise, hypothesizing that hippocampal place cells might per- form Bayesian cue integration, and that hippocampal reverse replay might play a role in cognitive map correction. We proposed biologically plausible mechanisms facili- tating these statistically near-optimal mechanisms, and reported modelling results of single-neuron recordings from rats and behaviour data from humans acquired outside this PhD to support the former, and sketch map accuracy data collected in experiments performed online supporting the latter hypothesis.

In addition to dealing with sensory noise and uncertainty, in realistic environments, large-scale representations also have to be stored and used efficiently. Hierarchical spatial representations help dealing with large amounts of spatial information by in- creasing the speed and efficiency of retrieval search and of route planning, as well as facilitating economical storage. It has been suggested that cognitive maps in humans are hierarchical, but the computational principles underlying these hierarchies have

7

received little attention. We investigated features influencing cognitive map structure using collected spatial memory data concerning real-world and virtual reality environ- ments, and proposed a computational mechanism (clustering in psychological space) which might give rise to sub-map structures. We validated our proposed mechanism using spatial memories of human subjects in over a hundred cities world-wide, and im- plemented a computational model able to predict, in advance, their sub-map structures based on our hypothesis.

Based on these insights, we developed a spatial memory module for a general cog- nitive architecture (the LIDA model of cognition), integrating it with the other cogni- tive mechanisms built into LIDA. We demonstrated the ability of the resulting model to deal with the challenges of the real world by running it in simulated environments, modelled after our participants actual urban environments, using high-fidelity robotic simulation software (including a physics engine) which provides the same interfaces as a real robot. Our LIDA-based spatial memory model could reproduce the spatial representation errors of human participants in different recreated environments, sub- stantiating the plausibility of the computational implementation of our hypotheses.

Chapter 1

Introduction

Brains have evolved to move bodies through space in order to increase the chances of survival and reproduction, through numerous complex behaviours such as fleeing from threats or searching for nutrients or potential mates. The ability to remember spatial information, e.g. previously encountered food sources or shelters, has provided sufficient evolutionary advantage that all known organisms with brains (and even some without, such as the slime mold - Reid et al. (2012)) have at least a rudimentary ability to utilize representations of space for more efficient navigation. Higher mammals have evolved a network of brain areas implementing spatial memory, a system for storing and recalling spatial information about the environment and about their location in it.

Representing spatial information accurately in the real world is hard, for several reasons. Sensors and actuators are limited, erroneous and noisy (in the sense of noise interfering with the signal). There are additional sources of uncertainty or unknown in- formation, such as external events, actions of other organisms, unperceived or currently unperceivable objects or events. Furthermore, physical environments can be highly complex, and yet cognitive resources (amount of memory, processing power, time and energy available) are necessarily limited by biological and physical constraints.

In artificial intelligence (AI) and robotics research, probabilistic models have pro- vided key tools for dealing with such challenges, facilitating the quantitative charac- terization of beliefs and uncertainty in the form of probability distributions, and the machinery of Bayesian inference for updating them with new data. They have also inspired the ‘Bayesian brain’ (Knill & Pouget, 2004) and ‘Bayesian cognition’ (Chater et al., 2010) paradigms in the cognitive sciences. These paradigms have been suc- cessful in explaining human behaviour in tasks as diverse as the integration of sensory cues (Ernst, 2006) including spatial information (Cheng et al., 2007; Nardini et al.,

2008), sensorimotor learning (Ko¨ rding & Wolpert, 2004), visual perception (Yuille & Kersten, 2006) or reasoning (Oaksford & Chater, 2007). Their success suggests an answer to what biological cognition might be doing to cope with the above-mentioned challenges: approximate Bayesian inference.

Slime molds are able to avoid previously explored areas using externalized spatial memories, and to solve mazes using nutrient gradients

1.1. MOTIVATION 13

1.1 Motivation

Despite of this success and of the suitability of probabilistic models to deal with un- certain and noisy spatial information, there have been few attempts to use them for modelling spatial memory within cognitive modelling, the branch of cognitive science concerned with computationally simulating mental processes. There is a gap in the litera- ture between probabilistic spatial models in robotics (called Simultaneous Localization and Mapping or SLAM) (Thrun & Leonard, 2008), which are capable of dealing with real-world noise, uncertainty, and complexity to some extent, but are cognitively im- plausible, and between computational cognitive models of spatial memory, which are designed to model biological spatial cognition, but cannot deal with all of these chal- lenges, and are thus confined to simplistic simulations (see Chapter 2 for a review).

In addition, although spatial representations in humans have been argued early to be hierarchical (Hirtle & Jonides, 1985; McNamara et al., 1989; Greenauer & Waller,

2010), similarly to some robotic implementations having to deal with large, complex environments (Kuipers, 2000; Wurm et al., 2010), it is not known how (by which pro- cess) these hierarchical spatial maps might be structured. Although many computa- tional models of spatial memory running in simplified environments exist, there is a lack of biologically and psychologically plausible ‘algorithms’ serving as models of human cognitive computations related to spatial information processing which can function in realistic, uncertain, complex environments.

The deprioritization of the problems of uncertainty and noise in favour of tractably modelling other human cognitive mechanisms is also pronounced in cognitive archi- tectures, which try to account for a large number of mental processes in a unified, comprehensive, systems-level model (as opposed to computational cognitive models, which usually focus on a single phenomenon). In their overview of the field, Langley et al. (2009) argue that “ we should attempt to unify many findings into a single theoret- ical framework, then proceed to test and refine that theory”, supporting the arguments of Newell (1973) that “you can’t play 20 questions with nature and win”, highlighting the importance of systems-level research in the cognitive sciences. Although a few such cognitive architectures do model spatial mechanisms in navigation space (Har- rison et al., 2003; Schultheis & Barkowsky, 2011; Sun & Zhang, 2004), they all run in simple, noise-free environments. According to a comparative table of cognitive ar- chitectures (Samsonovich, 2011) available in updated form online, there is currently no cognitive architecture implementing both Bayesian update and an empirically vali- dated, psychologically plausible ‘cognitive map’ at the same time.

In our usage of the terms, a computational model is ‘psychologically plausible’ (or ‘cognitively plausible’) to the extent that it is consistent with psychological findings and can accurately reproduce psychology data, i.e. behaviours. Analogously, it is ‘biologically plausible’ (or ‘neurally plausible’) to the extent that it is consistent with neuroscience and can reproduce neural data, e.g. single-cell recordings or brain imaging results.

<http://bicasociety.org/cogarch/architectures.htm>

14 CHAPTER 1. INTRODUCTION

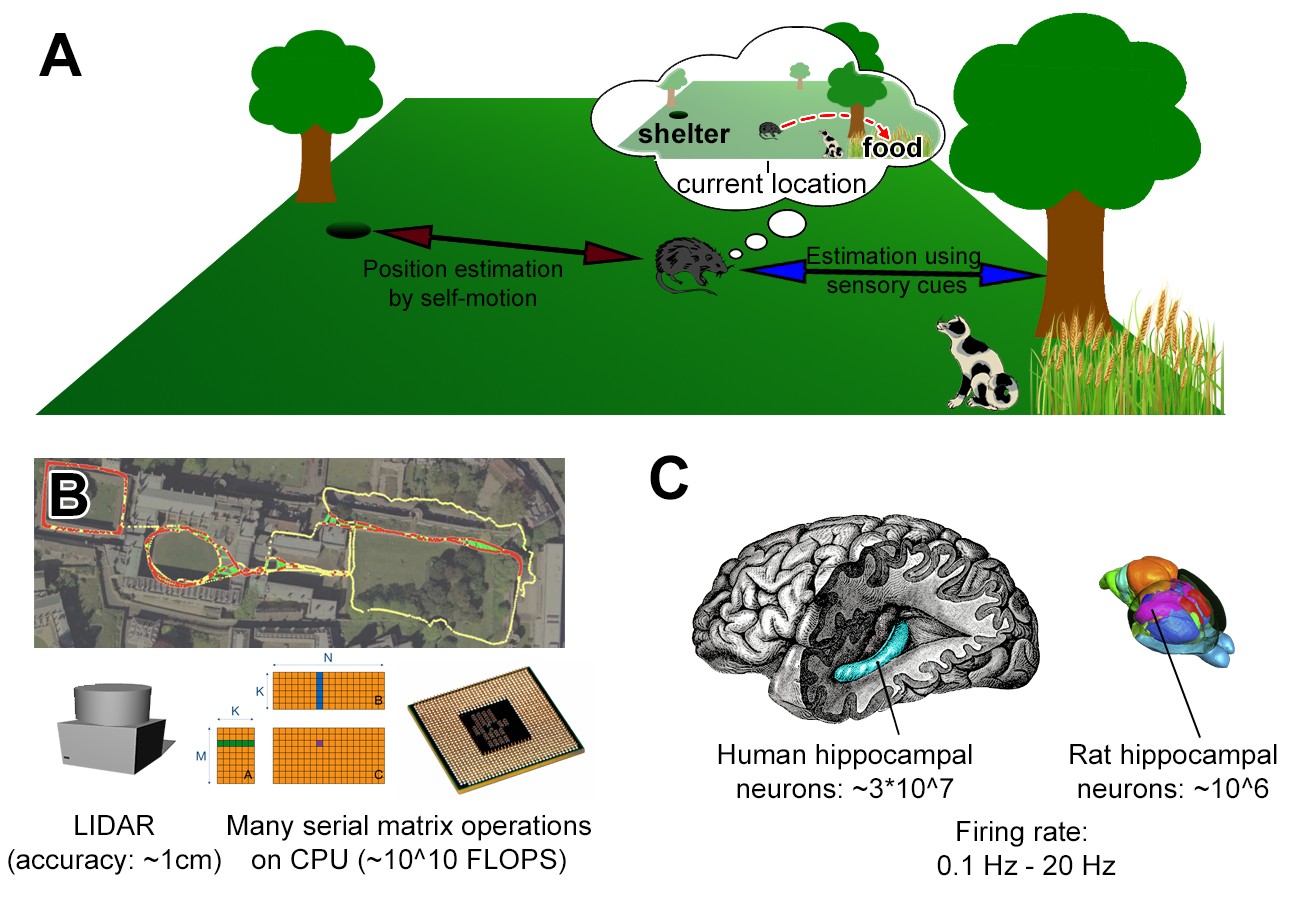


Figure 1.1: Motivation for proposing new computational cognitive models of spa- tial memory. A: Learning representations of the space around animals confers signif- icant advantages, such as the ability to plan a detour out of sight (dashed red arrow) to reach a food source while avoiding danger in this example. In real environments, this task is made more difficult by the unreliability, errors and noise inherent in both the estimation of position by integrating self-motion and in estimated object distances (e.g. based on vision). Most existing cognitive models of spatial memory neglect these challenges. B: State of the art SLAM models in robotics are able to estimate loca- tions and learn maps accurately, but rely on sensors and computations which are very different from biology - e.g. higher measurement accuracy using laser-based distance sensors (LIDAR), centralized control and coordination, and high number of serial op- erations per second - up to 1010 floating-point operations per second (FLOPS) needed for state of the art SLAM systems (Machado Santos et al., 2013). C: In contrast, the hippocampus - the major brain area involved in world-centered spatial representations

- contains only a few million neurons, of which only a subset is active at a time, each firing only a few times per second (Rapp & Gallagher, 1996; Sˇ imic´ & Bogdanovic´,

1997); and relies on noisy, inaccurate sensory measurements. Although many models of spatial memory in brains exist, there is a lack of computational mechanisms which are both neurally and psychologically plausible, and can work in realistic environments and with noisy sensors. (Example SLAM data in Panel B from (Newman et al., 2011), and 3D rat brain in Panel C from (Calabrese et al., 2013), with permission.)

1.1. MOTIVATION 15

The present work was motivated by these gaps in the literature, and aims to take com- putational cognitive models of navigation-scale spatial memory one step closer to mod- elling behaviour in realistic environments, such as high-fidelity robotic simulations or physical environments. It aims to do so by means of proposing probabilistic mecha- nisms of spatial cognition which are implementable in brains and can reproduce be- haviour data, and by computationally implementing these mechanisms, in the form of cognitive models and within an existing cognitive architecture. Situated within the computational sub-fields of cognitive science (cognitive modelling and cognitive ar- chitectures), the goal of this work is to contribute to the understanding of information processing in human cognition. As such, although it is computational in nature, the extent of its success is determined by its ability to predict and explain the kinds of behaviour data it is intended to model, as well as its consistency with established find- ings in psychology and neuroscience. It is not aiming for performance, or accuracy of learned spatial representations (these are the domains of robotics), or for maximizing neurobiological fidelity at the cellular level or below. Although building on neuro- scientific evidence, our concern is modelling spatial information processing on Marr’s algorithmic level of analysis (Marr & Poggio, 1976; Poggio & Marr, 1977), as opposed to e.g. biological neural networks - see Table 1.1 -, with a single exception.

|  |  |  |
| --- | --- | --- |
| ↓ Level of analysis | Description | In this work |
| 1. Computational | What problem(s) does the system solve, and why? | Localization,  Map error correction, Map structuring |
| 2. Algorithmic/ Representational | How might it solve them? (Using what representations and processes?) | Cognitive models of spatial memory |
| 3. Implementation | How is it implemented physically? | Place, grid, head- direction, border cells,  ... (Hartley et al., 2014) |

Table 1.1: Investigating spatial mechanisms on Marr’s (1976) levels of analysis. The present work is mostly concerned with the second level.

Unlike the rest of our work, we have investigated the plausibility of Bayesian spatial cue integration both on Marr’s algorithmic (Chapter 5) and implementation level (Chapter 3), in order to maintain the desirable criteria of both psychological and neural plausibility for our other models. Although this mechanism has been empir- ically substantiated on a behavioural level before (Cheng et al., 2007; Nardini et al.,

2008), its neural implementation has remained in doubt, with current mechanistic mod- els of Bayesian inference in brains making assumptions not fully consistent with the

Human cognition needs to keep track of the space of navigation as well as the spaces immediately around the body (e.g. reachable objects) and of the body (e.g. body-part configurations). Although uncertainty and noise play are important in the latter two spaces as well, we will confine ourselves to navigation-scale spatial mechanisms in this work.

16 CHAPTER 1. INTRODUCTION

anatomy or activity of the hippocampal complex (the major brain areas representing world-centered spatial information) - see next Section. This doubt of biological im- plementability has motivated our investigation of single-cell electrophysiological data (acquired outside this PhD) to provide the first evidence for Bayesian updating in the hippocampus on a neuronal level, and our proposal of a plausible mechanism for im- plementing it. This evidence, presented in Chapter 3, affords a degree of biological plausibility to the models utilizing Bayesian mechanisms in the rest of our work (which is concerned with processes on the algorithmic/representational level).

1.2 Probabilistic models of space in brains and minds

Although the focus of most of this work is on modelling behaviour data, we would like the employed mechanisms to be plausibly implementable in the parts of the brain they functionally correspond to. Apart from the lack of neuronal-level evidence that the hippocampal complex may perform Bayesian inference or even represent uncertainty, the possibility of the implementation of such a mechanism given the anatomical and electrophysiological constraints of this network of brain cells is also unclear.

Below, we briefly review probabilistic neural spatial models which have been pro- posed in the literature (Chapter 2 provides more general review of computational cognitive models of spatial memory). We start with normative models of dealing with spatial uncertainty, which derive optimal solutions to the problem a system might be solving (Marr’s computational level), and then continue describing mechanistic (implemen- tation level) models which might facilitate these, and their consistency with what is known about the hippocampal complex. Amore extensive review of Bayesian models in brains can be found in (Pouget et al., 2013; Vilares & Kording, 2011). There is currently little experimental support for any of the proposed neural uncertainty repre- sentations.

Models of probabilistic estimation of spatial information have been pioneered by (Bousquet et al., 1997), who suggested the use of a Kalman filter to model localization in the hippocampus. A Kalman filter is a dynamic Bayesian inference algorithm for estimating the values of unknown, not directly observable variables (such as location) from noisy observations, yielding statistically optimal estimates if the noise is normally distributed (Kalman, 1960). MacNeilage et al. (2008) also put forth arguments for dy- namic Bayesian inference as a model of spatial orientation, mentioning both Kalman filters and particle filtering (a related Bayesian filtering algorithm using samples in- stead of parameters to represent probability distributions) and leaving the question of their neural implementation open. Particle filter-based models of localization on the algorithmic level have been suggested by (Fox & Prescott, 2010; Cheung et al., 2012). Osborn (2010) went beyond localization, suggesting a Kalman filtering approach to also account for localizing objects in the environment. Recently, Penny et al. (2013) argued that if one presupposes the existence of ‘observation’ and ‘dynamic’ models,

Observation models and dynamic models are mathematical functions mapping from true states to observed states, and from pre-motion to post-motion states, respectively.

1.2. PROBABILISTIC MODELS OF SPACE IN BRAINS AND MINDS 17

required by Kalman filters, one might as well extend the inference to also use them for model selection (‘which environment am I in?’), motor planning (‘how do I get to place X?’), and to construct sensory imagery (‘what does place X look like?’) in addition to localization. They have combined these functions in a single probabilistic model, and argued that it is consistent with findings of pattern replay in the brain. An even more general probabilistic formulation based on dynamic Bayesian inference is the Free-Energy Principle (Friston et al., 2006), which aspires to provide a unified theory of brain function, and has been argued to be consistent with aspects of hippocampal processing (Friston et al., 2011).

Despite their considerable theoretical elegance, the above-mentioned models do not provide a final and complete answer to the motivating question of this thesis (Sec- tion 1.1), which can be summarized as: ‘how does biological cognition learn represen- tations of navigation space from noisy sensors in an uncertain world?’, for two reasons. First, none of them try to reproduce or show quantitative consistence with either be- havioural or neural data concerning spatial cognition (although qualitative consistence with anatomical and neural findings is pointed out by the authors). Although these models provide explanations, their predictions regarding spatial processing have not been quantitatively evaluated.

Second, in addition to the lack of quantitative validation, their neural implementa- tion is not known, and far from straightforward. For example, implementing the kinds of large matrix inversions and multiplications required by Kalman filters (Kalman,

1960) is easy on a computer, with centrally coordinated, serial, ‘fast’ computations, but difficult with the kind of distributed, parallel, ‘slow’ (on the level of single neu- rons, which only spike up to a few dozen times per second) computation performed by the brain. In the domain of world-centered, navigation-scale spatial mechanisms, any suggested neural implementation has to conform with not only the limitations imposed by biological neural networks, but also with the specific connectivity and activity ob- served in the hippocampal complex, in order to be considered biologically plausible.

In addition to such normative models, a number of mechanistic (implementation- level) models of how uncertainty and inference could be implemented in brains have also been proposed. They can be roughly grouped into three categories - see (Pouget et al., 2013; Vilares & Kording, 2011) for reviews. We briefly summarize these groups below, together with their consistency with what is known about the hippocampus.

• Probabilistic population codes (PPC) (Ma et al., 2006) encode probability dis- tributions in the logarithmic domain by means of a set of coefficients of cor- responding exponential basis functions, each coefficient encoded by the activ- ity (spike count) of a neuron. They assume neural variability is independent and Poisson-distributed. However, hippocampal neurons exhibit more variabil- ity than a Poisson process (Fenton & Muller, 1998; Barbieri et al., 2001). Also, if Bayesian inference were implemented in the hippocampus via a PPC, the en- coded probability distributions would strongly depend on the firing rate of hip- pocampal neurons: increased firing rates should mean decreased levels of un- certainty. But empirically, this is not the case - for example, firing rates increase

18 CHAPTER 1. INTRODUCTION

with movement speed (Maurer et al., 2005), which would mean the lowest uncer- tainties when running fastest (however, faster movements are harder to control and should thus lead to higher uncertainty).

• Instead of an encoding in the logarithmic domain, codes in which firing rates are proportional to probabilities have also been proposed, e.g. by Koechlin et al. (1999); Barber et al. (2003). The problem with their implementation in hip- pocampal neurons is that the firing rates of these neurons are also influenced by factors unrelated to probability, such as where the animal is headed (Ferbinteanu

& Shapiro, 2003) or trial dependent features (Allen et al., 2012), and can change substantially if either the shape or colour of an environment is altered (Leutgeb et al., 2005). These influences would strongly interfere with the outcome of the Bayesian inference, if it were implemented in a code that directly utilizes firing rates.

• Sampling-based codes represent probability distributions with a set of samples drawn from them (Fiser et al., 2010). They are asymptotically correct with in- finitely many samples, and approximations otherwise. Apart from being able to represent complex, multi-modal distributions, not having to rely on any fixed- form parametrization such as Gaussians, this also allows reducing their accuracy and computational demands by restricting the number of samples used. This property has been used e.g. by (Shi et al., 2010) to explain the deviations from the statistical optimum in an exemplar model of a reproduction task. It is difficult to make a general statement as to the implementability of this class of models in the hippocampal complex, as there is a wide variety of suggested concrete neu- ral implementations in non-spatial domains (Sanborn (2015) provides a review), and some applied to navigation space, e.g. (Fox & Prescott, 2010; Cheung et al.,

2012). None of them have been quantitatively validated by neural (electrophys- iological) measurements, although most of them are supported by behavioural observations.

How the brain might encode and utilize uncertainty is still an open question (Pouget et al., 2013), but based on the observations regarding the hippocampus outlined above, we argue that a sampling-based code is most suitable in this brain area; in terms of violating as few empirical observations as possible. We will provide electrophysio- logical evidence of Bayesian inference from single neurons, and propose a possible sampling-based mechanism, in Chapter 3 (and in more detail in Appendix TODO).

1.3 Hypotheses

To achieve goals in a spatially extended, realistic environment, at a minimum, an agent (e.g. a biological agent such as an animal, or an artificial agent such as a robot) must be able to 1) move, and keep track of its movements, 2) sense, and interpret its sen- sations, 3) represent spatial locations in its environment, e.g. of itself and its goal, 4)

1.3. HYPOTHESES 19

update these representations when changes occur in the environment, and 5) utilize these representations to achieve its goals (e.g. navigate to its goal location, avoiding dangers). Extensive work on all levels of analysis has been carried out for 1)-3), with the most recent Nobel prize in physiology or medicine awarded on the topic of 3) to John O’Keefe, May-Britt Moser and Edvard I Moser for the discovery of ‘cells that constitute a positioning system in the brain’ (Burgess, 2014) - see Chapter 2 below.

We have argued above that despite of the variety of existing models regarding

4)+5), new models are needed to move towards biological and psychological plausi- bility as well as real-world capability at the same time (since biological cognition has been shaped by the constraints and challenges of the real world, these should not be ne- glected in models of cognition). In particular, in accordance with the ‘Bayesian brain’ (Knill & Pouget, 2004) and ‘Bayesian cognition’ (Chater et al., 2010) paradigms, we have suggested approximate Bayesian inference to be a well-suited mechanism for tackling these challenges. Models on Marr’s algorithmic (and implementation) level which utilize such a mechanism require a number of underlying assumptions, some of which can be stated and evaluated as hypotheses.

We summarize major hypotheses in one place in Table 1.2 below, and expand on them in the respective results chapters below. The first two concern the representa- tion and manipulation of uncertainty in the hippocampus (required for maintaining approximately accurate location estimates despite noisy sensors and accumulating er- rors). Hypothesis 3 is needed since unless all remembered landmark locations are corrected at every moment (which would likely be intractable), a discrepancy between remembered and actual locations might arise when revisiting a location encountered previously (when traversing a ‘loop’ in the environment). This discrepancy necessi- tates a backward correction of multiple recent self and landmark locations to main- tain consistent representations. The last two are needed to formulate a computational mechanism of spatial representation structure. Structured, hierarchical representations provide clear computational advantages, such as increased speed and efficiency of re- trieval search, and economical storage. However, although strong neural (Derdikman

& Moser, 2010) and behavioural (Hirtle & Jonides, 1985; McNamara et al., 1989; Greenauer & Waller, 2010) evidence exists for such structure, underlying computa- tional principles have remained largely unknown.

|  |  |  |
| --- | --- | --- |
| Hypothesis | Prediction | Empirical support |
| 1 Hippocampal place cells  can perform approximate  Bayesian inference | Place field size depends on  uncertainty (e.g. proxim- ity of landmarks) in a Bayesian fashion | Place field sizes (recorded  from hippocampal neurons of behaving rats) are cor- related with uncertainties predicted by a Bayesian model (Chapter 3) |
| 2 Spatial uncertainty is  represented as the size of place cell firing fields |

20 CHAPTER 1. INTRODUCTION

|  |  |  |
| --- | --- | --- |
| 3 When revisiting a place,  estimates of recently traversed locations and encountered landmarks are updated in an approx. Bayes-optimal fashion | After revisiting parts of an  environment, place fields should shift, and recently active place cells should re-activate. Errors should conform to Bayesian pre- dictions | Neural: none in this work,  but place fields seem to shift after revisits (Mehta et al., 2000), and recently active place cells do reacti- vate (‘replay’) (Carr et al.,  2011). Behavioural: er- rors correlate with predic- tions (Chapter 5) |
| 4 The structure of spatial  representations arises from clustering | Landmarks which are  co-represented (belong together) in participants’ spatial memory should be closer in these features than those not belonging together | Neural: none in this work.  Behavioural: the probabil- ity of two landmarks being co-represented is strongly correlated with distances along these specific fea- tures. These distances al- low prediction of partici- pant representation struc- ture (Chapter 4) |
| 5 This clustering mecha-  nism operates on features including Euclidean distance, path distance, boundaries, visual and functional similarity |

Table 1.2: Hypotheses of the models presented in this work, and empirical sup- port. Place cell electrophysiological recording data was acquired outside this PhD. All other data has been collected by the author, unless otherwise specified.

1.4 Outline and Contributions

This thesis is presented in the Alternative Format allowed by the University of Manch- ester presentation of theses policy , which allows incorporating sections in a format suitable for publication in peer-reviewed journals. We chose the alternative format to more easily accommodate already published work, to reduce risks of self-plagiarism, and because of the largely self-contained nature of our individual results chapters. Thus, in what follows, the literature review (Chapter 2) and the three chapters (3-5) reporting the results, are copies of papers either accepted by or submitted to peer- reviewed journals. The following list summarizes these papers and the contributions therein:

• Chapter 2: Madl T., Chen K., Montaldi D. & Trappl R., 2015. Computational

[http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=7420](http://documents.manchester.ac.uk/DocuInfo.aspx)

In all publications, Madl wrote the draft of the paper, developed the software and/or designed the experiments, recruited and tested the participants, and analysed the data. Corrections suggested by Chen, Montaldi, and Franklin were incorporated into the final drafts by Madl after discussions with these co- authors. All publications were supervised by Chen and Montaldi, with Chen mainly commenting on mathematical and computational issues, and Montaldi on psychological and neuroscientific issues.

1.4. OUTLINE AND CONTRIBUTIONS 21

cognitive models of spatial memory in navigation space: A review. Neural Net- works, 65, 18-43.

Contributions: 1) a systematic review of representative cognitive models con- cerned with navigation-scale spatial memory, falling into symbolic, neural net- work, or cognitive architecture models, including a comparative table of the characteristics of these models.

• Chapter 3: Madl T., Franklin S., Chen K., Montaldi D. & Trappl R., 2014.

Bayesian Integration of Information in Hippocampal Place Cells. PLoS ONE

9(3), e89762

Contributions: 2) first quantitative electrophysiological validation of the repre- sentation of spatial uncertainty in the brain, and of Bayesian integration of spa- tial information in the brain, in three different environments (using data acquired outside this PhD). 3) Formulation and empirical support for an inference mech- anism based on coincidence detection (falling into the camp of sampling-based models of neural inference)

• Chapter 4: Madl T., Franklin S., Chen K., Trappl R. & Montaldi D., submitted.

Exploring the structure of spatial representations. Cognitive Processing Contributions: 4) behavioural evidence for clustering as the normative princi- ple underlying spatial representation structure, and 5) the first computational model of navigation-scale spatial representation structure on the individual level (able to predict this structure in participants’ long-term spatial memory from the geospatial properties of an environment)

• Chapter 5: Madl T, Franklin S, Chen K, Montaldi D & Trappl R, submitted.

Towards real-world capable spatial memory in the LIDA cognitive architecture.

Biologically Inspired Cognitive Architectures

Contributions: 6) integration of three spatial mechanisms capable of dealing with uncertainty and noise into a comprehensive cognitive architecture (localization, map structuring, map correction), and 7) embodying this architecture on a robot, allowing demonstration of the model functionality in a realistic robotic simula- tor. 8) proposal of a biologically plausible mechanism for correcting errors in learned maps when revisiting an already known place (the ‘loop closure’ prob- lem, well known in robotics, but neglected in cognitive science), and evaluation against behaviour data regarding cognitive map accuracy in human subjects.

The model best accounting for spatial memory structure presented in Chapter 4 also constitutes a novel kind of metric learning in machine learning, based on the idea of learning a similarity function in the space of absolute pairwise differences (as opposed to e.g. a Mahalanobis distance function). Although proposed before in a similar form for person re-identification in the computer vision community (Zheng et al., 2011), the insight that this space contains neglected information which can be utilized to improve

LIDA stands for Learning Intelligent Distribution Agent, and is reviewed in a paper co-authored during this PhD but not included in this thesis: (Franklin et al., 2014)

22 CHAPTER 1. INTRODUCTION

performance in general (not just on image data), and the general formulation allowing arbitrary constituent models for learning a metric in this space, are a novel contribution (9). Since it is too far from the topic of this thesis, metric learning in absolute pairwise difference space is only described briefly (to the extent required to model cognitive map structure) in Chapter 4. Applications and results on other kinds of data, with other constituent models, and in a semi-supervised setting, and are briefly summarized in Appendix TODO.

After presenting the mentioned papers constituting the literature review and re- sults chapters, we present an analysis of the methods employed during this research in Chapter 6. We continue to discuss the implications of our results and the neural imple- mentability of the mechanisms for which a concrete implementation has not already been presented in the results in Chapter 7. We conclude in Chapter 8 with a conclusion and an outline of potential future work opened up by this research.

We note that the line of criticism mentioned regarding the neural implementability of the high-level probabilistic models of localization in the previous section also apply to our proposed mechanism of cognitive map structuring (Chapter 4). Although it is intended to be a cognitive and not a neural model, we have argued that consistency with the underlying neuroscience can and should play a role in constraining the space of possible models, and evaluating models, even on the algorithmic level. But the map structuring mechanism in Chapter 4 is, to our knowledge, the first formal model of the observed structure in cognitive maps, both on Marr’s computational and algorithmic levels. We did not have the time and resources to extend it down to include a plausible neural implementation within this PhD.

Bibliography

Allen, K., Rawlins, J. N. P., Bannerman, D. M., & Csicsvari, J. (2012). Hippocampal place cells can encode multiple trial-dependent features through rate remapping. The Journal of Neuroscience, 32, 14752–14766.

Barber, M. J., Clark, J., & Anderson, C. H. (2003). Neural representation of proba- bilistic information. Neural Computation, 15, 1843–1864.

Barbieri, R., Quirk, M. C., Frank, L. M., Wilson, M. A., & Brown, E. N. (2001). Con- struction and analysis of non-poisson stimulus-response models of neural spiking activity. Journal of neuroscience methods, 105, 25–37.

Bousquet, O., Balakrishnan, K., & Honavar, V. (1997). Is the hippocampus a kalman filter? In Proceedings of the Pacific Symposium on Biocomputing (pp. 655–666).

Burgess, N. (2014). The 2014 nobel prize in physiology or medicine: A spatial model for cognitive neuroscience. Neuron, 84, 1120–1125.

Calabrese, E., Johnson, G. A., & Watson, C. (2013). An ontology-based segmenta- tion scheme for tracking postnatal changes in the developing rodent brain with mri. NeuroImage, 67, 375–384.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. Nature neuro- science, 14, 147–153.

Chater, N., Oaksford, M., Hahn, U., & Heit, E. (2010). Bayesian models of cognition.

Wiley Interdisciplinary Reviews: Cognitive Science, 1, 811–823.

Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian inte- gration of spatial information. Psychological bulletin, 133, 625.

Cheung, A., Ball, D., Milford, M., Wyeth, G., & Wiles, J. (2012). Maintaining a cog- nitive map in darkness: the need to fuse boundary knowledge with path integration. PLoS Comput. Biol, 8, e1002651.

Derdikman, D., & Moser, E. I. (2010). A manifold of spatial maps in the brain. Trends in cognitive sciences, 14, 561–569.

30

BIBLIOGRAPHY 31

Ernst, M. O. (2006). A bayesian view on multimodal cue integration. Human body perception from the inside out, (pp. 105–131).

Fenton, A. A., & Muller, R. U. (1998). Place cell discharge is extremely variable during individual passes of the rat through the firing field. Proceedings of the National Academy of Sciences, 95, 3182–3187.

Ferbinteanu, J., & Shapiro, M. L. (2003). Prospective and retrospective memory cod- ing in the hippocampus. Neuron, 40, 1227–1239.

Fiser, J., Berkes, P., Orba´n, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. Trends in cognitive sciences,

14, 119–130.

Fox, C., & Prescott, T. (2010). Hippocampus as unitary coherent particle filter. In Neural Networks (IJCNN), The 2010 International Joint Conference on (pp. 1–8). IEEE.

Franklin, S., Madl, T., D’Mello, S., & Snaider, J. (2014). Lida: A systems-level architecture for cognition, emotion, and learning. Autonomous Mental Development, IEEE Transactions on, 6, 19–41. doi:10.1109/TAMD.2013.2277589.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain.

Journal of Physiology-Paris, 100, 70–87.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference.

Biological cybernetics, 104, 137–160.

Greenauer, N., & Waller, D. (2010). Micro-and macroreference frames: Specifying the relations between spatial categories in memory. Journal of Experimental Psy- chology: Learning, Memory, and Cognition, 36, 938.

Harrison, A. M., Schunn, C. D. et al. (2003). ACT-R/S: Look ma, no” cognitive-map.

In International conference on cognitive modeling (pp. 129–134).

Hartley, T., Lever, C., Burgess, N., & O’Keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 369, 20120510.

Hirtle, S., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. Memory & Cognition, 13, 208–217.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.

Journal of Fluids Engineering, 82, 35–45.

Knill, D. C., & Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. TRENDS in Neurosciences, 27, 712–719.

32 BIBLIOGRAPHY

Koechlin, E., Anton, J. L., & Burnod, Y. (1999). Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area mt. Biological cybernetics, 80, 25–44.

Ko¨ rding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learn- ing. Nature, 427, 244–247.

Kuipers, B. (2000). The spatial semantic hierarchy. Artificial intelligence, 119, 191–

233.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. Cognitive Systems Research, 10, 141–160.

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., & Moser, M.-B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. Science, 309, 619–623.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. Nature neuroscience, 9, 1432–1438.

Machado Santos, J., Portugal, D., & Rocha, R. P. (2013). An evaluation of 2d slam techniques available in robot operating system. In Safety, Security, and Rescue Robotics (SSRR), 2013 IEEE International Symposium on (pp. 1–6). IEEE.

MacNeilage, P. R., Ganesan, N., & Angelaki, D. E. (2008). Computational approaches to spatial orientation: from transfer functions to dynamic bayesian inference. Jour- nal of neurophysiology, 100, 2981–2996.

Marr, D., & Poggio, T. (1976). From Understanding Computation to Understanding

Neural Circuitry.. Technical Report DTIC Document.

Maurer, A. P., VanRhoads, S. R., Sutherland, G. R., Lipa, P., & McNaughton, B. L. (2005). Self-motion and the origin of differential spatial scaling along the septo- temporal axis of the hippocampus. Hippocampus, 15, 841–852.

McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. Journal of Experimental Psychology: Learning, Memory, and Cognition,

15, 211.

Mehta, M. R., Quirk, M. C., & Wilson, M. A. (2000). Experience-dependent asym- metric shape of hippocampal receptive fields. Neuron, 25, 707–715.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. Current biology, 18, 689–693.

Newell, A. (1973). You can’t play 20 questions with nature and win: Projective com- ments on the papers of this symposium, .

BIBLIOGRAPHY 33

Newman, P., Chandran-Ramesh, M., Cole, D., Cummins, M., Harrison, A., Posner, I., & Schroeter, D. (2011). Describing, navigating and recognising urban spaces- building an end-to-end slam system. In Robotics Research (pp. 237–253). Springer.

Oaksford, M., & Chater, N. (2007). Bayesian rationality: The probabilistic approach to human reasoning. Oxford University Press.

Osborn, G. W. (2010). A kalman filtering approach to the representation of kinematic quantities by the hippocampal-entorhinal complex. Cognitive neurodynamics, 4,

315–335.

Penny, W., Zeidman, P., & Burgess, N. (2013). Forward and backward inference in spatial cognition. PLoS Comput Biol, 9, e1003383.

Poggio, T., & Marr, D. (1977). From understanding computation to understanding neural circuitry. Neurosciences Research Program Bulletin, 15, 470–488.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains:

knowns and unknowns. Nature Neuroscience, 16, 1170–1178.

Rapp, P. R., & Gallagher, M. (1996). Preserved neuron number in the hippocampus of aged rats with spatial learning deficits. Proceedings of the National Academy of Sciences, 93, 9926–9930.

Reid, C. R., Latty, T., Dussutour, A., & Beekman, M. (2012). Slime mold uses an externalized spatial memory to navigate in complex environments. Proceedings of the National Academy of Sciences, 109, 17490–17494.

Samsonovich, A. V. (2011). Comparative analysis of implemented cognitive architec- tures. In BICA (pp. 469–479).

Sanborn, A. N. (2015). Types of approximation for probabilistic cogni- tion: Sampling and variational. Brain and Cognition, . URL: http:

[//www.sciencedirect.com/science/article/pii/S0278262615300038.](http://www.sciencedirect.com/science/article/pii/S0278262615300038) [doi:http://dx.doi.org/10.1016/j.bandc.2015.06.008](http://dx.doi.org/10.1016/).

Schultheis, H., & Barkowsky, T. (2011). Casimir: an architecture for mental spatial knowledge processing. Topics in Cognitive Science, 3, 778–795.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing bayesian inference. Psychonomic Bulletin & Review,

17, 443–464.

Sˇ imic´, G., & Bogdanovic´, N. (1997). Volume and number of neurons of the human hippocampal formation in normal aging and alzheimer’s disease. Journal of Com- parative Neurology, 379, 482–494.

34 BIBLIOGRAPHY

Sun, R., & Zhang, X. (2004). Top-down versus bottom-up learning in cognitive skill acquisition. Cognitive Systems Research, 5, 63–89.

Thrun, S., & Leonard, J. J. (2008). Simultaneous localization and mapping. In Springer handbook of robotics (pp. 871–889). Springer.

Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncer- tainty, behavior, and the brain. Annals of the New York Academy of Sciences, 1224,

22–39.

Wurm, K. M., Hornung, A., Bennewitz, M., Stachniss, C., & Burgard, W. (2010).

Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems. In Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation. volume 2.

Yuille, A., & Kersten, D. (2006). Vision as bayesian inference: analysis by synthesis?

Trends in cognitive sciences, 10, 301–308.

Zheng, W.-S., Gong, S., & Xiang, T. (2011). Person re-identification by probabilistic relative distance comparison. In Computer Vision and Pattern Recognition (CVPR),

2011 IEEE Conference on (pp. 649–656). IEEE.