School of Computing
FACULTY OF ENGINEERING

UNIVERSITY OF LEEDS

A Topological Interpretation of Place Cell Activity in the Rat Hippocampus

Hugh Osborne

Submitted in accordance with the requirements for the degree of
MSc Advanced Computer Science (Intelligent Systems)

2016/17

The candidate confirms that the following have been submitted.

Items	Format	Recipient(s) and Date
Deliverables 1 & 2 : Data Analysis Report	Report	Supervisor (01/07/17)
Deliverables 4 & 6 : Progress Report and Results Presentation	Report	Supervisor, Assessor (04/08/17)
Deliverables 3 & 5	Code Submission	Supervisor, Assessor (06/09/17)
Deliverable 7	Final Report	SSO (06/09/17)

Type of project: Exploratory Software (ESw)

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

(Signature of Student) ______________________

Summary

Previous work has shown that, in principle, it is possible to analyse the electrical activity of so called ‘place cells’ in the brain of a rat, as it explores a maze, to infer various geometric and spatial properties about that maze. This is achieved by analysing the spike times of multiple place cells which fire when the rat passes through associated regions of its environment. A logical description of the spatial relations between these regions, called place fields, can be used to derive the topology of the maze and identify features such as loops.

A solution has been produced previously which works successfully with simulated data. This project shows that the principles behind that solution also work well with real experimental data by employing a number of techniques to reduce sources of error. These techniques include smoothing the data, eliminating problematic cells and introducing an alternative definition of the spatial relation between place fields. Two separate data sets are used to evaluate the solution for different maze shapes. As an extension, the solution is applied to rest data to show that topology of earlier mazes can be inferred from ‘replay’ of place cells while the rat is asleep.

Acknowledgements

I wish to thank my supervisor, Anthony Cohn, for his guidance and support throughout the project and Yuri Dabaghian for providing links to the experimental data sources and offering perspective on the expectations of real world cell recordings. I would also like to thank David Harrison for an interesting conversation early on regarding alternative methods for inference.

Contents

1. Introduction

1.1 Context

David Eagleman, Associate Professor of Neuroscience at Baylor College of Medicine, says in his book, The Brain [1], “The most complicated thing we’ve discovered in the universe is the human brain. How does, what happens in these three pounds of jello-like material, become us?”. Understanding how the brain works is imperative to knowing how we learn about and understand the world and for our understanding of mental health. Seventeen percent of adults in England meet the criteria for a common mental health disorder [2]. 850,000 people suffer from dementia in the UK caused by multiple diseases in the brain [3].

The hippocampus has long been understood as a crucial component in the way we learn and remember [4, 5]. It is one of the few places in the brain where we have observed adult neurogenesis, the growth of new neurons [6]. Without it, human short term memory is severely impaired and long term memory becomes “read only” [7]. Navigation is also impaired as is the ability to remember sequences [8].

Beyond issues associated with physical damage, other pathologies such as atrophy and plaques in the temporal cortex and hippocampus are an early indication for the onset of Alzheimer’s disease and Dementia [9]. Investigating the structure and behaviour of cells in the hippocampus, then, is of immediate importance. From a computational standpoint, as systems such as IBM Watson [10] and Google Deep Mind’s AlphaGo [11] push the boundaries of machine cognition, we must look for further inspiration from the biological state of the art. What is more, the question of whether the brain is a (not so simple) computer [12] has implications for how we evaluate AI systems and how we define what intelligence is. For many years, artificial intelligence research was primarily concerned with knowledge representation using formal logical calculi [13]. This was intuitive based on human propensity for structure and pattern recognition as well as the mathematical beauty of many logical systems. Recently, AI research has shifted emphasis to that of biologically inspired neural networks and regression of large volumes of data [14, 15]. Where possible, it is preferable to interpret bio-inspired systems or indeed biological processes themselves using logic. It links our high level intuition to a biologically plausible model and helps us to understand how our everyday perception emerges from the structure of the brain. Furthermore, in order to support neurological research, computer science projects can help to provide novel interpretations of experimental results as well as produce applications to open up new methods and avenues of experimentation.

1.2 Spatial Reconstruction Experiment

This project is born from a single thought experiment outlined by Dabaghian, et al. [16]. An experiment is setup in which a rat is running round a maze, isolated from an experimenter who may only observe signals from electrodes in the rat's hippocampus. What can be deduced from these signals about the geometrical and spatial properties of the rat's environment? Previous work on this problem [17], described later, successfully inferred the topology of a maze with high accuracy from simulated cell data. There was, however, little success when directly applying the same solution to experimental data. The aim of this project is to see if the solution can be extended to improve performance with real world data. Firstly, a program will be implemented to visualise and to perform analysis on the data. Once the main obstacles to correctly deducing topology have been identified, a set of processes or techniques can be designed and implemented to improve the performance of the solution. If successful, there are a number of possible extensions to this work such as investigating the topology of large open spaces or applying the solution to signals received while the rat is resting.

The literature review chapter will explore what led to the formulation of the Spatial Reconstruction Experiment, examine results from neurological investigations which may affect what can be deduced from hippocampal cells, and cover computational research topics required to implement a solution. Finally, the work done in the previous project will be examined in detail as this will form the basis to the implementation in this project.

2. Literature Review

2.1 Towards the Spatial Reconstruction Experiment

Place Cells and Place Fields

Memory and intelligence testing of patients who had undergone resection of the medial temporal lobe (so called lobotomies) in the Fifties, identified the hippocampus as a centre for short term memory in the brain [7]. Later, O'Keefe and Dostrovsky [18] investigated spike firing rates of cells in the rat hippocampus by recording from electrodes inserted into the brain. They found that the activity of cells in a region of the hippocampus known as CA1 responded only when the rat was in a certain location within a testing platform (and for certain other cells, facing in a particular direction). During further experimentation, O'Keefe [19] found that many of these "place cells" respond solely to the rat being in a single associated region of its environment. He hypothesised that each "place field", defined as the spatial region in which there is heightened activity of the place cell, is based on some internal navigation system. The location of place fields are dependent on the locations of certain spatial cues relative to each other [20]. In particular, visual cues such as lights or picture cards, sound and feel cues have been shown to inform the location of a place field. Due to the stability of place fields in an unchanging environment, Brown et al. [21] showed that it is possible to decode the spike trains from place cells to estimate the location of a rat within a testing platform with reasonable accuracy. This relies on the rat being familiar with the environment and for a reasonable coverage of place fields to be found.

Experimental Methods

Many of the experiments mentioned here involve the implantation of electrodes into the rat's hippocampus under anaesthetic. Initially, these were dual microelectrodes [19] across which a potential difference could be measured, but in later experiments were replaced by tetrodes [22] which allowed for the triangulation, using potential difference, of individual cells.

As described by operant conditioning techniques [23], rats may be trained to perform a ground task such as performing an action or moving to a certain location using food rewards. Rats are usually starved to ˜85% body weight to increase the effective reward of the food. Generally, training platforms consist of connected walled corridors as well as open areas depending on the specific experiment and are often contained within a roughly metre squared area. The experiments by O'Keefe and Conway [20] provide a good example of what is typical in this area. An alternative method for motivating spatial learning is to use the so called Morris water maze [24]. This involves a pool full of opaque water and a platform positioned just below the surface so it is invisible from above. A rat is placed into the pool and naturally attempts to find the platform. The benefit of this technique is that the rat does not need to be starved before training and, indeed, training has been found to be quicker. The downside, however, for experiments involving electrodes is obvious.

Place Cells in Humans

There is some experimental evidence that humans also have place responsive cells in the hippocampus. Ekstrom et al. [25] and Miller et al. [26] took advantage of patients with drug resistant epilepsy undergoing invasive monitoring with electrodes in the left and right temporal lobes. In both experiments, patients were asked to play a computer game involving moving around a town. In the Ekstrom experiment, as the patients moved between shops in the town to pick up taxi passengers, 11% of cells were found to respond to place fields in a similar manner to rat place cells. Other cells responded to goals and local objects in the game. In the Miller experiment, 25.6% of cells across the hippocampus, entorhinal cortex and amygdala were identified as place responsive with a majority sensitive to direction. In addition, it was found that these place responsive cells were reactivated when items in the game were recalled later without the visual cues of the game world. While it is apparent that there is more going on in the hippocampus of humans in terms of cell responses, the existence of place responsive cells in humans further justifies the research into rat place cells.

Internal Spatial Representation

O'Keefe and Nadel [27] postulated that the hippocampus of humans, and other animals including rats, encode a "cognitive map". This map is an internal representation of the environment, not relative to the body and is perceived to remain stationary as the body moves. To investigate whether place cells fire simply in response to stimuli at the current location of the rat or if there is indeed a memorised internal representation, O'Keefe and

Speakman [28] performed a spatial memory task in which a rat would be trained to reach a goal in the presence of spatial cues. The cues would then be removed and the rat would need to find the goal again. The ability of the rats to do this task and the fact that 90% of place fields were retained after the spatial cues were removed provides good evidence for an internal representation which is not simply a response to current stimuli.

Place Field Size

During O'Keefe's [19] initial investigation into place cells, he found that the size of place fields varies greatly. In his experiments, sizes ranged between $10cm^2$ and nearly half of the $2m^2$ testing area. In experiments where the testing area was increased over time, it was shown that place fields can stretch to fill the space, although this stretching is limited and beyond a certain size, additional place fields emerge [29–31]. It has also been shown [32] that the size of a place field can depend on the location of the place cell itself in the hippocampus.

Remapping

Muller and Kubie [33] investigated the effect of changes in a rat's environment on its place fields. They found that when all visual cues were rotated in the testing platform, the place cells also rotated by the same degree. When the testing platform was scaled up, there was some degradation of the fields, that is, some place fields would change location or the place cell would not fire at all, however, some place fields did scale with the environment. Changing the shape of the testing platform (for example from circular area to a square one) or introducing barriers to the area, caused all place fields to change position or disappear. This demonstrates an important aspect of place cells, their propensity to "remap" to partially or entirely different configurations of place fields when certain changes to the environment are made.

Another way to cause remapping of place fields, is for the behavioural task of the rat to change in the same environment [34]. In an experiment where a rat switched from a foraging task (randomly located food pellets) to a goal oriented task (move to a specific pile of food), additional directed place fields were shown to emerge.

In 2004, Moita et al. [35] showed that inducing a fear response in a rat, which had learned a spatial environment and whose place fields were stable, would cause some remapping to occur.

Remapping due to spatial changes in the environment can be avoided if those changes are made in suitably small increments. O'Keefe and Burgess [29] while investigating the geometric determinants of place fields, saw that changing the scale in either or both dimensions of an explorable area caused similar scaling of the place fields and no remapping was observed.

Gothard et al. [31] observed similar scaling along a track with food rewards at either end. One end was movable, thereby extending the length of the track. Small changes to track length, led to a scaling of the place fields whereas larger changes caused some place cells to cease firing. Similarly, Fenton et al. [36] reported distortions in place fields in relation to the movement of black and white cards in a cylindrical testing platform.

Muller et al. [37] categorised this behaviour in three ways. A **Null remapping** describes the case where place fields remain associated with their respective place cells and may change shape or position as long as the relative position of all place fields do not change. A **Full remapping** occurs when all place cells either cease firing in any location, begin firing having not done so before or whose place fields have shifted to entirely different positions relative to others. Finally, **partial remappings** describe the case where only some place fields change.

Topological Ordering

The null remapping of place cells, during which place fields maintain adjacency relations with each other and relative ordering, led Dabaghian et al. [16] to hypothesise that place cells code for a topological representation of the environment. An answer to their Spatial Reconstruction Experiment would then be that, from the spike firing activity of place cells, the topology of the maze could be deduced but not the absolute shape or scale. For example, a W shaped maze would look identical to a Y shaped maze as each has a single junction with three arms. The size of each maze would also have no effect. This interpretation is supported by the results of a number of experiments on the hippocampus with respect to its “sequence coding” behaviour. Chiba et al. [38] showed that hippocampal lesions severely impair a rat’s performance in spatial ordering tasks. The hippocampus is further implicated in non spatial ordering tasks such as remembering sequences of odours [39, 40]. This behaviour is relevant to the idea of topological ordering in which sequencing is more of a concern than absolute position.

2.2 Reasoning about Place Fields

The Region Connection Calculus

Dabaghian et al. describe inferring geometric information about the environment by reasoning about the spatial relations between place fields. To do this, they suggest the use of the Region Connection Calculus formulated by Randell et al. [41]. Allen relations [42] define a representation of time intervals in terms of the different ways in which two intervals can overlap. The RCC is a similar “interval logic” but can be applied to both temporal and spatial regions. The eight base relations of RCC are presented in figure 2.1 and derive from the 'connected' relation which says that two regions, x and y, are connected if there exists a point in x which is also in y. The overlapping circles in the figure illustrate the various ways in which two regions can be connected (or not connected). There are simplified versions of these RCC8 relations such as RCC5, shown in figure 2.2.

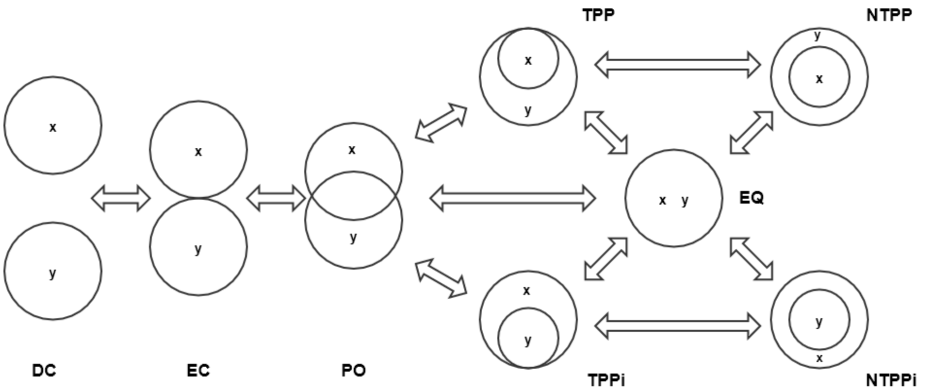


Figure 2.1: ***DC*** *: Disconnected,* ***EC*** *: Externally Connected,* ***PO*** *: Partial Overlap,* ***TPP (TPPi)*** *: Tangential Proper Part (inverse),* ***NTTP (NTTPi)*** *: Non-tangential Proper Part (inverse),* ***EQ*** *: Equal. Arrows show possible transitions between relations. Transition diagram reproduced from [41].*

RCC5 is appropriate for models in which the edges or borders of regions are poorly defined such that there is no need to identify regions which are just touching. As explained in [17], because the edges of place fields are defined between locations with no cell activity, and locations with some activity which is subject to variance and noise, RCC5 provides the appropriate set of relations to describe them.

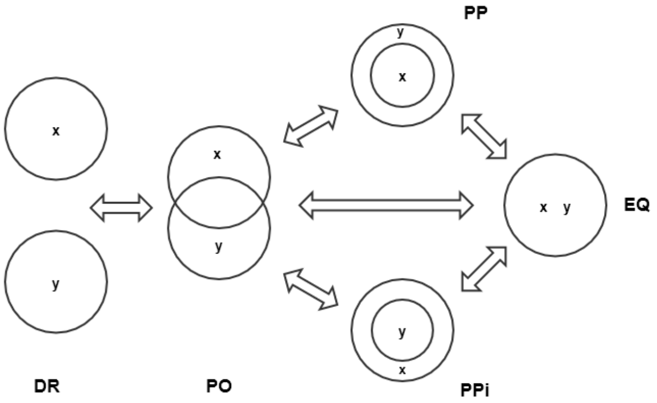


Figure 2.2: ***DR*** *: Discrete,* ***PO*** *: Partial Overlap,* ***PP (PPi)*** *: Proper Part (inverse),* ***EQ*** *: Equal. Arrows show possible transitions between relations.*

Simplicial Complexes

Having created a description of the connectedness of place fields using RCC5, the original work from Dabaghian et al. described the use of a simplicial complex to derive topological characteristics of the environment. A simplicial complex is a structure used in an area of mathematics called algebraic topology to describe shapes or surfaces (see figure 2.3). A simplex is a generalised triangle in k-dimensions. A triangle is a 2-simplex, that is, a simplex in two dimensions. A 3-simplex is a tetrahedron or triangular based pyramid. A simplicial complex as described in [43] is a collection of simplices connected via faces, edges or vertices. A simplicial complex can be built from a graphical representation of place fields such that fields are represented by vertices with edges between fields which satisfy the RCC5 PO relation. From this structure, certain functions, which are beyond the scope of this document, can be applied to give a clearly defined representation of the space. This includes identifying the Betti number which is the number of 'holes' in the space. As an example, an O shaped maze would have a (1 dimensional) Betti number of 1 and a figure-8 maze would have a Betti number of 2. This project focuses on a different method for identifying certain topological attributes, described later.

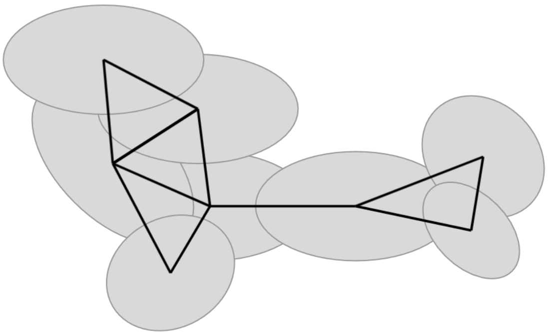


Figure 2.3: *The grey oblongs represent a collection of place fields. Vertices in the black simplicial complex are located at place field centres and edges denote PO. The individual edge is valid in a simplicial complex as it is defined as a 1-simplex.*

2.3 Further Relevant Characteristics of Place Cells

Ensemble Coding (Multiple Place Field Cells)

A key question to answer if attempting to infer the relative locations of place fields is, does a place cell encode a single unique place field in the environment or more than one? Fenton et al. [44] answered this question with a study of place fields within a much larger testing area than before ($2.1m^2$). They showed that place cells can code for multiple distinct place fields in a large enough area and that this supports a so called "ensemble coding" in which a vector of place cell signals uniquely identifies a location in the environment.

If it is the case that place cells code for multiple fields, then one should expect to see locally repeated patterns of place fields in areas which are identical with respect to the spatial cues which govern place cell firing. This is dependent on the absence of so called "path integration", some mechanism upstream from place cells in the hippocampus to track where the rat has come from. There is some disagreement to what extent path integration affects place fields. Skaggs and McNaughton [45], monitored the place fields in two identical boxes connected by a corridor. Despite replacing one box with the other or changing the starting location, there was almost no similarity of place field locations between each box. This supported the hypothesis of a path integration process.

Evidence contradicting Skaggs and McNaughton was compiled by Spiers et al. [46] who conducted an experiment with rats in a maze of four identical compartments connected by a corridor. They found strong evidence for similar place field locations in each compartment represented in the activity of a single place cell.

Observing increased activity in a "Multiple Place Field" (MPF) cell may indicate that the rat is in one of a number of locations and the presence of these cells, therefore, is significant for this project.

Context Sensitivity of Place Cells

Frank et al. [47] showed that in some cases, place cells fire based on the rat's current activity while travelling through its place field. In a W shaped maze, for example, a CA1 place cell may fire in the central arm if the rat is heading out towards its food goal but not when it is returning. Similarly, Wood et al. [48] modified a T shaped maze to a figure-8 configuration so that the rat would continuously travel up the central arm and be trained to turn left and then right on alternating loops. Despite using much less accurate stereotrodes, place fields were identified in the central arm which responded only in the context of an upcoming left turn or an upcoming right turn.

Other experiments [34, 49] have further extended the idea of context sensitive place fields. Markus et al. confirmed that place cells are more selective of directionality in mazes with corridors compared to more open areas. Moreover, directionality becomes more of a factor in place fields when the rat is working towards a goal indicating that planning can be associated with certain place fields.

Even early work by O'Keefe [19] pointed to the context specificity of place cells when some so called "misplace cells" increased firing activity for place fields where a reward was expected and there was none or alternatively when a reward was present where previously there had been none. Indeed, there is further evidence [31, 50] of place fields attached to the food goal of tasks for which a rat was trained and also fields that are associated with walls and other obstacles within an arena. This, again, indicates that place cell activity represents an internal cognitive map. It also means that observing increased place cell activity may not always indicate the rat is in its place field and furthermore, observing no activity does not always guarantee the rat is outside of its place field.

Learning and Place Field Formation

On encountering a novel environment, some time is required for place fields to form and stabilise. The internal representation of a cognitive map through place cells, coupled with repeated place cell activity during sleep, points to a mechanism involving spatial coding for learning the environment. Foster and Wilson [51] showed that for a brief time immediately after exploration of a new environment, while the rat is awake, place cells show activation in reverse order to that which led to the current goal location. This is postulated to be evidence for reinforcement learning based on back-propagation of reward.

Monaco et al. [52] showed that place field potentiation occurs as a rat explores a new environment and as the internal representation is updated. In particular, it was shown that this happens while the rat performs head-scanning movements to survey the immediate spatial cues.

It seems, however, that not all place fields emerge over time during exploration. Hill [53] found that for many place cells, associated place fields are apparent on the first traversal by a rat of a novel environment. Frank et al. [54] confirmed this but expanded the investigation to find that some place cells start off silent, showing no activity, but later develop to produce a place field. They suggested that place field stability is only reached after 5 to 6 minutes of exploration.

2.4 Avenues of Investigation beyond Topology Inference

Wake vs Sleep

An exciting area of research derived from O'Keefe and Dostrovsky's discovery is the activity of place cells during sleep. Pavlides and Winson [55] who confirmed and extended work by Ranck [56] found that place cells activated as a rat explored a maze when awake, showed heightened activity during slow wave sleep and to a lesser extent, REM sleep. Further exploration [57, 58] showed that the order in which place cells fire, as the rat travels through a number of place fields while awake, leads to increased activity in the same order during sleep states. This indicates that there may be some form of replay of activity during sleep. It may be the case that the topology of the maze being replayed may also be found using the same technique as when the rat is awake.

Other Cells in CA1

Place cells are not the only functional unit in CA1 or the wider hippocampal formation. Hartley et al. [59] give an excellent summary of the major types of spatial cell in addition to place cells. Head direction cells, as the name implies, show increased activity when the animal is facing a certain allocentric direction. That is, the direction is not relative to the body. Head direction cells have been found in both the presubiculum and entorhinal cortex [60, 61] and have been shown to influence the allocentric orientation of place fields [62].

In 2005, research led by May-Britt and Edvard Moser [63] uncovered the existence of grid cells in the entorhinal cortex. An individual grid cell shows increased activity in certain locations in the environment similar to place cells but there are many more place fields associated with one cell, organised into a grid like structure with scale, periodicity and orientation. These cells are understood to be upstream of place cells in the hippocampus and have less context sensitivity [64].

Finally, cells which respond to locations near walls or other obstructions in the local environment, known as boundary cells have been discovered in many areas of the hippocampal formation [65–67] with the exception of the hippocampus itself [59].

Continuous Attractor Networks

Further understanding of how path integration, for planning and navigation, is performed within the hippocampal region is an ongoing topic of research. One model for understanding how a network of neurons can code for navigation is called the continuous attractor neural network model [68]. The model describes a network of interconnected neurons such that more excitatory connections between two neurons implies they are functionally proximal (for example, a head direction cell has more excitatory connections to another cell which codes for a similar head direction). The model defines a collection of active place cells as a chart which, when combined with theoretical “attractor map” cells and cross referenced with other spatial inputs such as head direction cell activity, provides a contextual cognitive map directly associated with cues in the world. Although Barry and Burgess [69] cite little evidence for this model among the place cells in CA1, they discuss its validity among entorhinal grid cells, the head direction cell network and place cells in CA3. Recent work by Meshulam et al. [70] reinforces the application of this model to cells in both CA3 and CA1 when applied to both place cells and non-place cells.

2.5 Previous Project

Although there were other areas of investigation and various topological inference algorithms were tested in the previous project, this section concentrates on the use of RCC to describe place field connectivity and an algorithm used to find junctions in the topology of the maze as these will be relevant in later chapters.

Connectivity Graph

According to the Spatial Reconstruction Experiment, an application acting as the observer has no knowledge of the shape of the maze and the only information available is the time at which each place cell spikes. With the assumption that each cell fires when the animal is within the place field represented by that cell, the previous project [17] used RCC5 to describe deductions that can be made about the relationships between place fields. During a session, as a rat runs round a maze, if two place cells show increased spiking at the same time, it can be deduced that the place fields of these two cells must at least partially overlap (PO). In contrast, if there is never a time when two cells cofire, their place fields must be disconnected. By making these deductions for all pairs of place fields in a session, a graph of connectivity can be built where each node represents a place field and an arc represents partial overlap or connectedness.

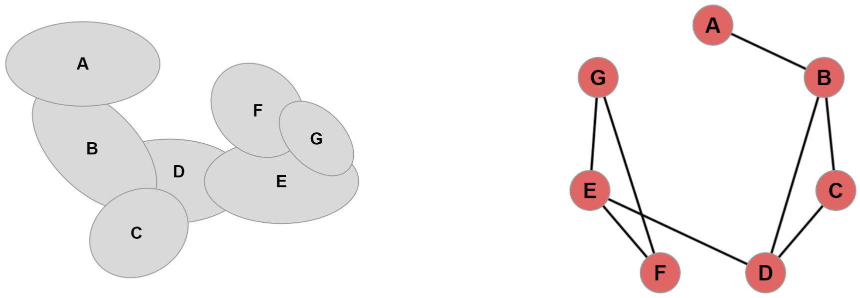
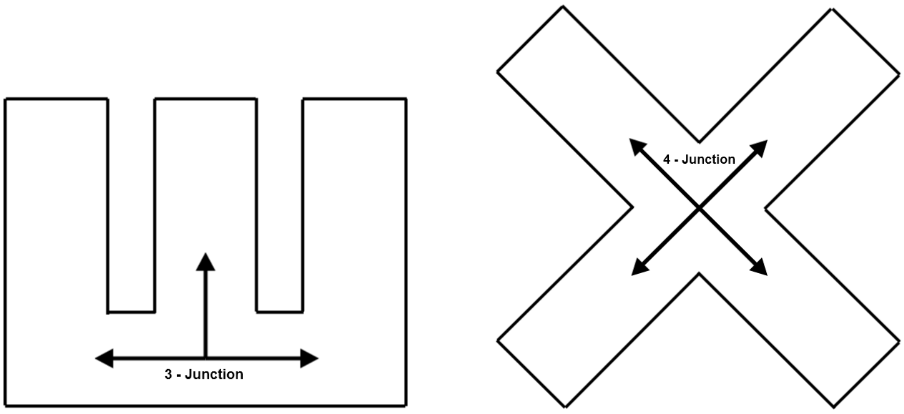


Figure 2.4: *(left) Partial overlap between place fields. (right) The derived connectivity graph.*

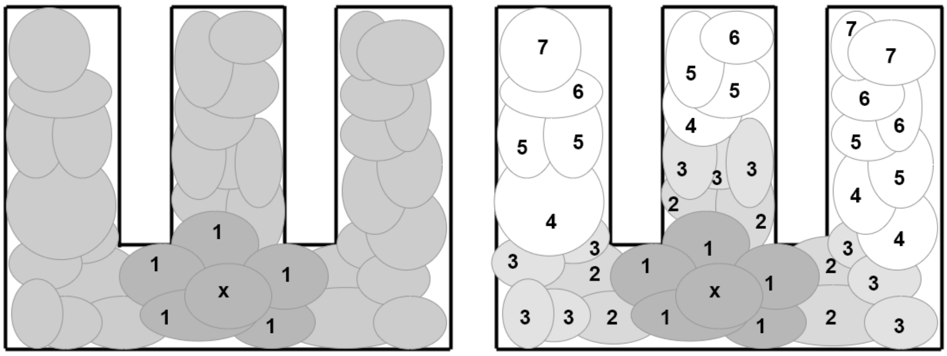
Junction Finding

Once the connectivity graph has been generated, an algorithm is implemented to find the number of junctions in the maze and the number of corridors leading away from each (the degree of the junction). A junction is defined as an area of a maze with three or more corridors leading away from it. For example, in the W shaped maze, there is a single junction of degree three which connects the three arms. A plus shaped maze would have a single junction with degree four and an H shaped maze would have two junctions of degree three. The junction finding algorithm starts from any place field in the maze which is then labelled x. The next step is to find all fields which overlap with x. This is equivalent to finding all neighbours of x in the connectivity graph. These fields are labelled 1 to denote their distance from x as shown in figure 2.6.

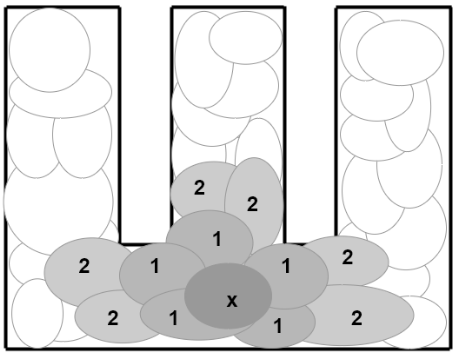


(a) *W-maze with a 3-junction.* (b) *Plus shaped maze with a 4-junction.*

Figure 2.5



(a) *The first set of connected fields.* (b) *The fully labelled maze.*



(c) *The 2-labelled fields form three disconnected subsets.*

Figure 2.6

All fields which are neighbours of any 1 fields and not already labelled are given the label 2 as they are all 2 jumps away from x. The algorithm continues counting distances from x until all fields have been labelled.

For each set of fields who all have the same label, the number of disconnected subsets is counted. A set has disconnected subsets if there exists a pair of fields in the set such that there is no path between them. Figure 2.6c shows an example of a set of 2 fields with three disconnected subsets surrounding field x and its 1 fields.

If a set has three or more disconnected subsets, this indicates that there must be a junction in the maze of that degree.

Initial Attempts with Experimental Data

The connectivity graph and junction finding algorithms were shown to be highly successful with the simulated data and further investigations were carried out into identification of other topological features including loops and larger open spaces called arenas. With only a short time left, testing was performed on experimental data and although generally unsuccessful, the challenges posed to the algorithms with real world data were well documented and will be covered later in section 3.3.

3. Data Analysis and Visualisation

3.1 Data Sources

In order to assess the feasibility of deriving a maze topology from experimental data, it was necessary to develop a simple application to visualise the location of the rat as it ran around a maze and the locations at which spiking occurred. The expectation was, for each place cell, a distinct cluster of spiking activity would be visible as the rat passed through the corresponding place field. An assessment could then be made of the general size, shape, reliability and coverage of place fields for each experiment. Additionally, any noise or unexpected activity could be noted and perhaps explained using results from the background research. The two available experimental data sets are described below.

Loren Frank

The experimental data set used by the previous project is from the lab of Loren Frank at UCSF. The data are recordings from hippocampus regions CA1 and CA3 in seven Long-Evans rats and from the MEC and CA1 regions in two further rats. The methodology is summarised in [71]. The rats were deprived of food to 85% of their initial body weight. Initially, they were trained to run from one end of a linear track to the other to receive a reward at each end. The animals then underwent surgery to implant microdrives containing 30 tetrodes.

After recovery, each day during two or three 15 minute sessions in a W shaped track, the animals were allowed to behave freely and learned through trial and error to perform an alternation task (figure 3.1) starting from the central arm then moving along the left arm before returning back to the centre, then the right arm and back and so on. This behaviour would result in a reward given at the end of each arm provided the correct order was followed. The animals were given rest sessions, during which cell recordings were also taken, in a black box before and after each run. Two geometrically identical W tracks in different parts of the room (separated by a large black divider) were used for the sessions. For the first two days, track A was used for two distinct 15 minute sessions alternating with three rest sessions. From day three onward, both tracks were used, two sessions of track B and one session of track A, with five alternating rest sessions.

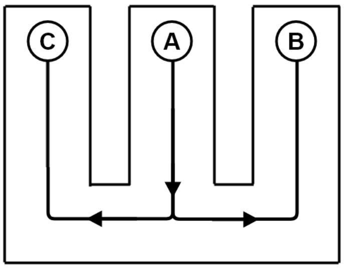


Figure 3.1: *The rat's path repeatedly follows A to C to A to B to A.*

Although other data is available, the pertinent data used for this project includes the position of the rat during each session, derived from tracking of an LED attached to the rat's head. The position data is assumed to be unavailable to the observer in the spatial reconstruction experiment but it is useful to be able to visualise the rat's behaviour and for analysis of the spike data. For the spike data itself, distinct cells in the hippocampus were identified from the signals at each tetrode and the times at which each cell produced a spike was recorded. The LED video processing to derive position and cell and spike time identification were performed by the Frank team and made available as Matlab data files which was the starting point for this project.

Gyorgy Buzsáki

In order to test the solution with different shaped mazes, a further data set was obtained from the Gyorgy Buzsáki lab at NYU. These recordings were also taken from the CA1 and CA3 regions as rats explored various different mazes.

- A single linear track
- A plus or cross shaped maze
- A Figure 8 maze with a wheel in which the rat is trained to alternate turning right or left before returning to the wheel for a period of time.
- Two open square areas of differing size ($180cm^2$ and $120cm^2$)

As with the Frank data, both the position and occurrence of spikes was recorded during each session and made available in an already processed form.

3.2 The Animation Visualiser

Both datasets were converted to a consistent format for use throughout the project. A MATLAB script was implemented to extract the position and spiking data of the Frank set and a python script was implemented to convert the .csv formats of the Buzsáki set. For each experimental session, then, there is a run.txt file which lists the rat's position at each timestep (the Frank data was at 30Hz and the Buzsáki data was at 39.06Hz). There is also a file, for each identified cell, listing the times at which that cell fired. An application was implemented to display a mouse icon which moves according to the position data in run.txt. It deposits a single sprite, a gaussian radial gradient, in the the rat's location in the event of a spike. The application was implemented in C++ utilising OpenGL GLUT due to its ease of setup, cross platform nature, extensive online support, and flexibility when compared to languages such as Java or Python. The position data was normalised so that the extent of the maze corresponded to the size of the visualiser window. Although the shape of the maze is not displayed, it can usually be inferred from the movement of the rat or location and shape of the place fields.

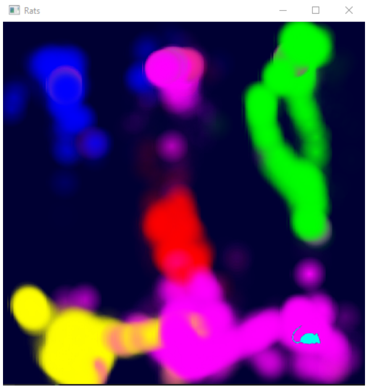


Figure 3.2: *The visualiser application showing the mouse icon and five place fields derived from the spiking data of five different cells during a session in the w-maze.*

In order to obtain more quantitative results beyond an animation, a single snapshot of all spikes at their corresponding locations was saved to a file and the screen was split into a grid. Each grid square was given a value corresponding to the number of spikes which occurred while the rat was in that square. This 'spike rate' value was normalised to between 0 and 255 and displayed as a red colour in the image as well as recorded to a file. An example is shown in figure 3.3

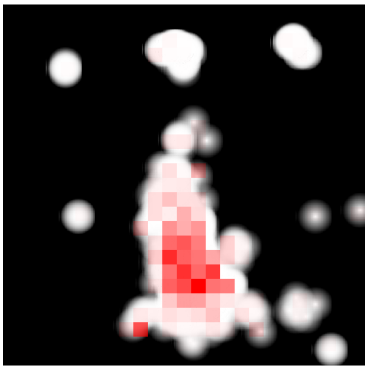


Figure 3.3: *A snapshot image of all spiking activity from a single cell's spike file with the red 'spike rate' grid overlay. The grid gives a good indication of where a place field is located and its extent by taking into account the number of spikes in a given area. This is in contrast to the white gaussian deposits which display dormant or spurious firing with the same intensity as place field firing.*

3.3 Analysis

Python was used to further analyse this data due to its more functional nature and the simplicity of the matplotlib library [72] for plotting graphs.

Figure 3.4 summarises the number of cells showing spiking activity during a single session for six rats from the Frank experiments. In total there are eight rats, six of whose data shall be used for testing and evaluation. The remaining two rats have recordings from cells in different regions of the hippocampus. As mentioned earlier, the spike rate grids give a good indication of where a place field is located and its extent by taking into account the number of spikes in a given area. These grids were therefore used primarily for understanding the nature of the place fields in each experiment.

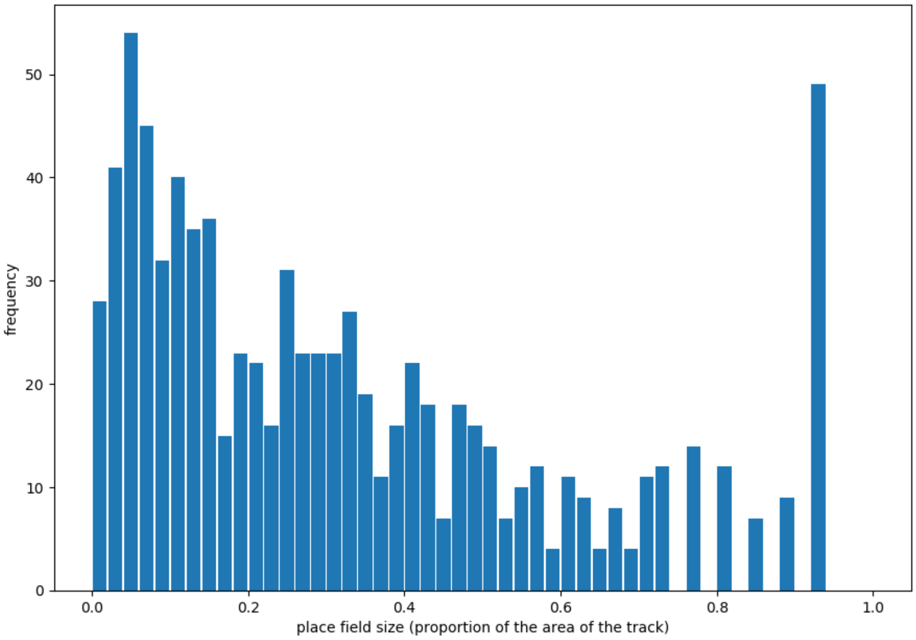
Figure 3.4: Number of place cells available in a single session

Name	Mean spiking units	max	min
Bond	16.6	30	9
Conley	8.5	18	3
Miles	12.5	19	4
Frank	21.8	41	4
Dudley	5.9	12	2
Eight	20.2	27	12

Place Field Size

Figure 3.5 shows the histogram of sizes of all place fields during run sessions on track A and track B undertaken by three rats across all days and sessions of the w-maze experiments.

Figure 3.5: Histogram of Place Field Area (W-maze)



As a proportion of the size of each track, the place fields are generally quite large but with high variance. For the w-maze, there is an outlying number of place fields covering almost the entire track. Without expert knowledge of the functioning of the electrodes and behaviour of cells in the hippocampus, it is difficult to say if these fields come from valid place cells or not but it is probable that they are, in fact, interneurons. Interneurons do not act like place cells and have a very high spike rate. This high spike rate can be interpreted as constant cell activity resulting in a 'place field' which covers the entire track. There are also a high proportion of place fields which take up almost no space at all on the track. Certain cells spike sporadically with no cohesive place field and these are recorded as spike rate grids with no bright red squares. It can be said that these are either not place cells or they code for a location which does not lie within the maze. These outlier cells will likely be a source of noise later on.

Multiple Place Fields and Spurious Firing

As described in [17] and in the literature review, there are a number of Multiple Place Field (MPF) cells in the datasets. These cells produce field images such as in figure 3.6.

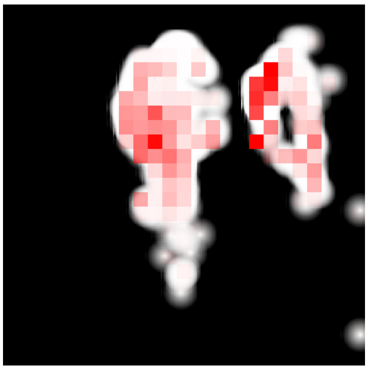


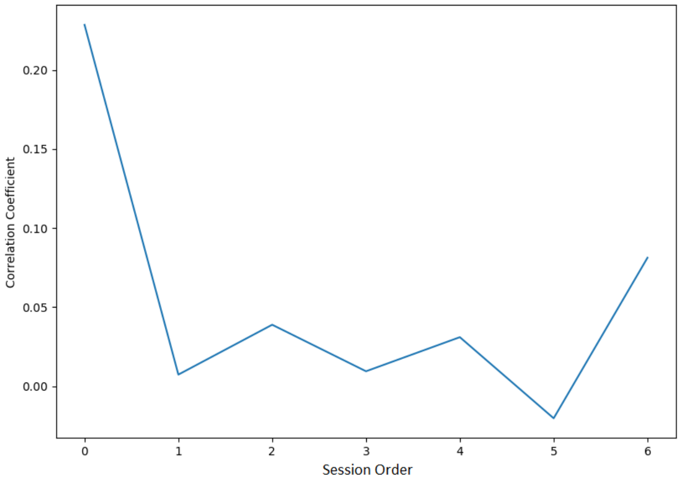
Figure 3.6: *An MPF cell produces more than one distinct place field in the track.*

It will be necessary to find a solution to deal with these cells as connectivity with other fields will be ambiguous in this case. Similarly, many cells show spurious firing at the ends of each arm of the w-maze (shown previously in figure 3.3). This is most likely due to increased overall activity in the hippocampus while eating or potentially as a result of replay on achieving a goal (in this case, reaching the food in the correct sequence). In general, there is a large amount of firing outside of the place field. It is probable that this can be reduced using a data smoothing technique.

Stability of Place Fields

For the w-maze experiments, each rat was required to perform the alternation activity in two different mazes of the same shape over a number of days and sessions. If no place field remapping occurred between each of these activities, it would be useful to be able to take spike data from consecutive sessions to increase the confidence of the solution. To measure the degree to which place fields remain stable across sessions, the 25x25 spike rate grids from the same cell in two consecutive sessions were flattened to vectors of length 625. The two vectors were then zipped together with each tuple treated as an observation from a pair of discrete random variables. If it were the case that the two grids were identical, if the field was exactly the same shape from one session to the next, calculating the correlation coefficient between the two random variables would yield a value of 1. If they were completely different, perhaps the place field switched to a different location entirely in the second session, the correlation coefficient would be close to -1. Figure 3.7 shows the average correlation coefficient between the first session and subsequent sessions for each available cell during track A alternation activities for the three rats. In many cases, consecutive cell spike data is unavailable due to a negligible number of recorded spikes. This alone is an indication that remapping may be occurring between sessions, although movement or failure in the tetrodes could also account for this.

Figure 3.7: Average Correlation Coefficient between the First Session and Subsequent Sessions



The correlation between the first and second sessions' place field is generally higher, due to place fields remaining stable for the same track on the same day (for three of the six rats, the same activity was performed twice each day). After this, however, the average correlation drops to nearly zero. When inspecting the field images of individual cells across multiple sessions, it appears that though there are some fields which remain stable over many days, many others do not. This explains the near zero correlation average and indicates that while there may be some value to using data from multiple sessions, additional work would be required to identify only those fields which haven't changed. For now, only spiking data from individual sessions will be used.

Place Field Coverage

One factor which would surely affect the ability of the application to estimate topology correctly is whether the place fields recorded in a single session cover the track fully. If it is the case that one arm of the w-maze has no place fields, identifying this arm (and thus an arc from a junction node) would be impossible. To confirm acceptable coverage, the spike rate grids from each cell were summed and displayed as a heat map. For each session, two maps were generated, one with all place fields and one with only place fields smaller than 40% of the full size of the track. The stipulation of smaller place fields allow confirmation that full coverage is not only as a result of large place fields spanning the entire track due to the presence of interneurons. Figure 3.8 shows examples from different rats, mazes and sessions for all field sizes. Figure 3.9 shows coverage of cells with fields smaller than 40% of the area of the track.

Figure 3.8: Place Field Coverage : All Place Fields

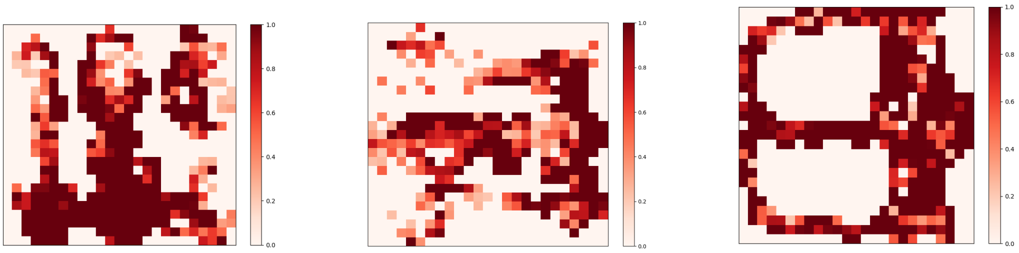
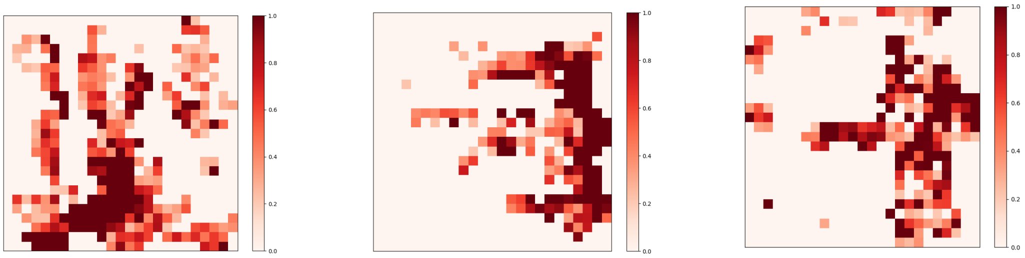


Figure 3.9: Place Field Coverage : Place Fields $< 40\%$



From left to right the maze shapes are a W, a backwards E, and a figure 8. Red colour indicates activity from any cell in that location of the maze. White indicates no activity. The 40% or smaller fields show that, while there are gaps, place fields are spread across all areas of each maze.

For the smaller sized fields, there are occasions where there is a break in between the available place fields as shown in figure 3.10 . In this case, no partial overlap will be apparent across the gap and so the connectivity graph may show a set of disconnected place fields.

The figure also highlights the question of coverage of the rat's path around the maze. If the rat does not visit an area of the maze, it is guaranteed that no cell will fire in that location. This could lead to a part of the maze topology being missed or topology being incorrectly inferred. In the figure, the rat frequently travels up the middle of the central arm of the maze then on separate occasions turns right and follows the right hand wall back down or turns left to follow the left hand wall. As the left and right sides are covered by different place fields and they do not overlap in the middle, it may be deduced that the left and right sides are two separate corridors which is incorrect.

Finally, although not visible in the still images of firing over an entire session, it is the case that there are cells which do not consistently fire when the rat is in the associated place field. This could be due to error in the recording equipment or formation of a place field during the session due to spatial learning.

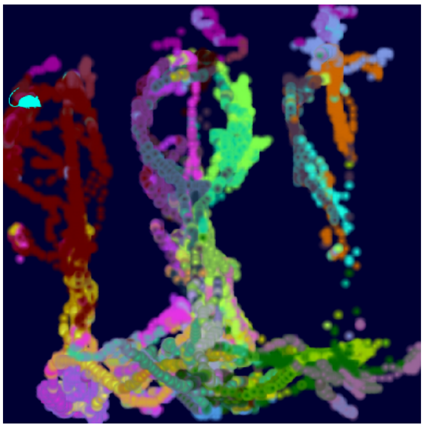


Figure 3.10: *There is a gap between place fields in the right arm and those in the rest of the track. In addition, the central arm shows three adjacent fields which are not connected due to the rat's path.*

In summary, the following facts were confirmed from the analysis

- In general, place fields present themselves in all areas of the maze although there are sometimes gaps between them which may hinder the use of 'Partial Overlap' as a means to derive connectedness.
- The range of place field size is larger than the simulated version and there are far fewer place cell recordings available.
- There are instances of place cells which fire in two or more distinct places - so called multiple place field (MPF) cells.
- There is a high incidence of spurious firing and firing at important locations such as food sources in addition to the place field.
- There are instances of cells which fire constantly throughout the experiment so no useful location information can be deduced from these.
- For many place cells, there is little or no similarity of location of their place fields in an experiment from one day to the next.
- A few place cells show inconsistent firing when the rat is in the corresponding place field.
- The path of the rat does not cover the whole maze leaving gaps between place fields which may be misinterpreted.

4. Design and Implementation

4.1 Extending Junction Finding

The junction finding algorithm from the previous project returns a list of labelled sets of fields and their disconnected subsets. From this list, a network graph can be generated such that each node represents a subset and each arc represents connectedness of each subset. Not only does this aid in visualising the results of the algorithm but the network identifies subsets of fields which lie at the ends of corridors, nodes with degree 1, denoting dead ends. It also identifies subsets of fields lying along a corridor with degree 2 and junction nodes with degree 3 or higher.

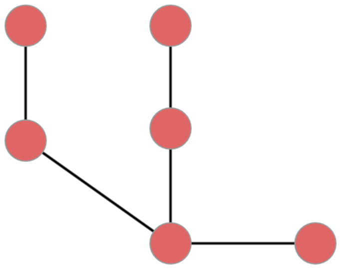


Figure 4.1: *A graph representing topology of a W-maze.*

In fact, the algorithm can be altered to simply build this 'topology graph' without generating the list of subsets first.

Starting with field x in the **connectivity** graph, find each disconnected subset among the neighbours of x and create a node in a new **topology** graph for each. The fields in each of these nodes are connected via x so create arcs between each of the new nodes.

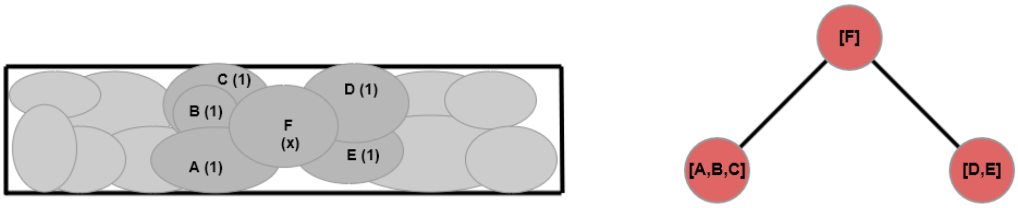


Figure 4.2: *(left) Place fields overlapping with x. (right) The partial graph of sets of connected fields.*

Iteratively, for each disconnected subset x', find each disconnected subset among the neighbours of x' (not including those already visited) and create a node in the topology graph for each.

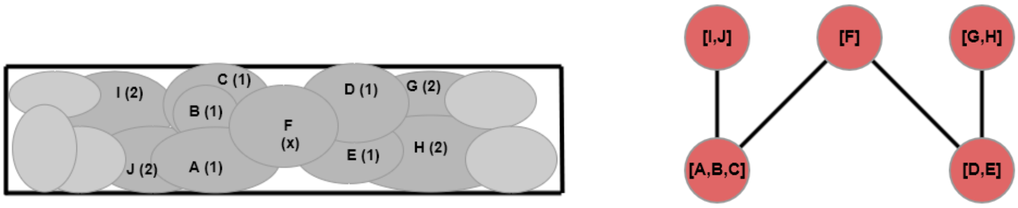


Figure 4.3: *(left) Place fields overlapping with the neighbours of x. D and E form a disconnected subset and A,B and C another. (right) The next step in the partial graph of sets of connected fields.*

The stopping criterion of the iterative process is if there are no unvisited adjacent fields.

Identifying Loops

In this form, the junction finding algorithm is unable to identify loops in the maze. If it were the case that the iterative step found a path back to previously visited fields, the stopping criterion would be satisfied and the final arc linking the start and end of the path would not be added. An observation can be made that the only place where an arc should be added to create a cycle in the graph is between the node at the end of a branch, that is a node of degree 1, and any other node in the graph whose fields are connected. Figure 4.4 shows the two appropriate locations for ‘cycle arcs’ in a figure-8 topology graph.

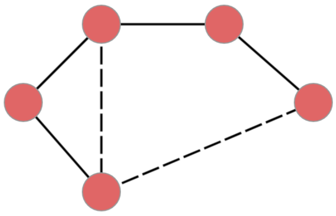


Figure 4.4: *The dotted arcs are cycle bridging arcs identified by the loop identifying step.*

An additional step was added to the junction finding algorithm after the end of the iterative process, once the topology graph is generated. For each node in the graph of degree 1, search through all other nodes in the graph. If there are place fields in the degree 1 node which are connected to fields in the other node then create an arc between them thus creating a cycle.

Trimming the Topology Graph

The lack of coverage of the rat's path around the maze and generally a lack of total coverage of place fields raises a potential problem. A small branch of partially disconnected fields which in reality should overlap with a number of neighbours can be construed as a unique arm or corridor of its own. This would appear in the topology graph as an additional path leading away from a junction node. It is expected that these paths would be short and made up of only a few fields whereas genuine corridors will be denoted by much longer paths with many more fields. With this in mind, another step was added to the end of the junction finding algorithm. All nodes, in the completed topology graph, of degree 1 are removed if the number of fields in that subset is less than some parameterised value. This 'branch cut' value will be explored in the evaluation.

4.2 Connectivity Graph

Avoiding MPF Cells and Large Place Fields with Intersections

From the results of the data analysis, it was necessary to make some changes to the RCC algorithm used to derive connectvity. Figure 4.5 shows a typical problem which can occur due to MPF cells in the dataset.

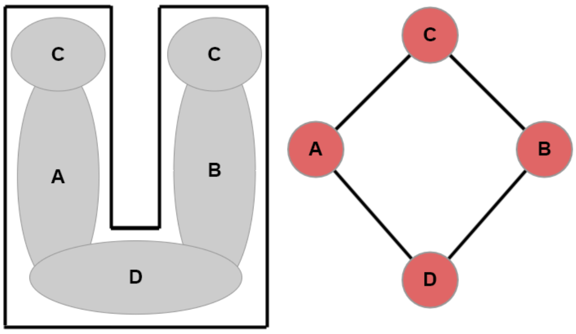


Figure 4.5: *A schematic showing the effect of MPF cells on connectivity.*

Field D overlaps with Fields A and B. Field C also overlaps with Fields A and B but is from an MPF cell and is split between both arms of the maze. The connectivity graph would show that A is connected to C and B is also connected to C. The topology graph would therefore deduce that there is a path from A to B via C. This is incorrect.

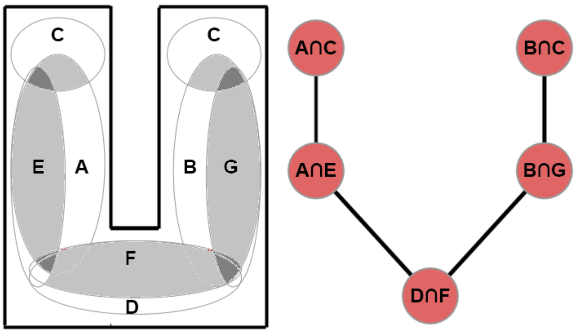


Figure 4.6: *(left) Additional overlaps can be used to avoid MPF confusion. (Right) The corrected connectivity graph of 2-field intersections.*

Figure 4.6 shows the same case except now fields E,F and G have been included. If instead of deriving connectedness from PO of individual fields, it was derived from the PO of intersections of strictly two, the case where C fires on its own, would not be included in the connectivity graph. The topology graph algorithm would deduce, correctly, the path in the figure.

This solution is unlikely to scale well with larger numbers of cells. As soon as two MPF cells cofire in two or more place fields, the stipulation of using intersections of two fields fails for the same reason as it did with individual fields and there would be a requirement of strictly three field intersections. As the number of intersections required increases, the size of the available fields will decrease to the point where there are no longer any overlapping fields. The number of field intersections to be used to derive the connectivity graph will be a parameter explored in the evaluation section.

A similar problem to that of MPF cells occurs when a cell has a very large place field which extends across and beyond a junction. Figure 4.13 shows an example where fields A, B and C would appear to be arms from the junction field D. However, field E overlaps them all and so the connectivity graph shows them all to be connected. The topology graph algorithm is then unable to derive a junction, only a single field connected to all cells.

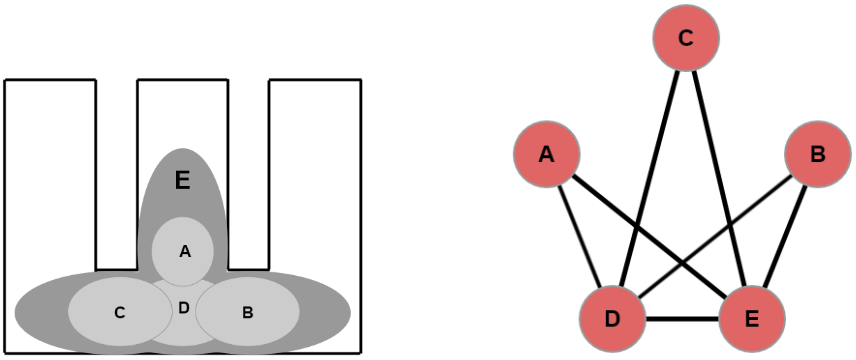


Figure 4.7: *(left) Place fields entirely subsumed by a single large field. (right) The resulting connectivity graph.*

As before, this problem can be partially solved with the application to field intersections instead of individual fields.

A Threshold to Eliminate Spurious and Inconsistent Firing

One additional factor was added to the connectivity graph algorithm to tackle the issue of inconsistent firing of a place cell as the rat moves through its place field, as discussed in section 3.3, and to reduce the effect of spurious firing. A weight was added to each arc of the connectivity graph to represent the number of occurrences of cofiring between two cells during the session. In each session, the rat traverses the maze multiple times and it would be expected that each traversal would cause cofiring of the cells at least twice, once on the way out and once on the way back. By including a 'weight cutoff' parameter such that arcs with a weight below this threshold are removed, spurious occurrences of cofiring outside of place fields and inconsistent place cells can be eliminated from the graph.

Connectivity using Timing

A different way to define connectedness of place fields is to observe that if cell A fires at time, t, then cell B fires at t+1, the rat must have passed through the field corresponding to cell A and moved into the field corresponding to cell B and they must therefore be adjacent. This is a looser definition of connection than partial overlap and solves two main problems.

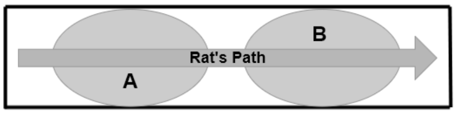


Figure 4.8: *Over time, as the rat moves along its path, each cell spikes in order.*

Firstly, in the case that there is a gap between fields in the maze which would show no overlaps due to cofiring, consecutive firing over time would still be able to capture this connectedness and include the 'orphan' fields in the graph. This assumes that times during which no firing occurs are ignored. Secondly, in the scenario described earlier where very small field intersections are being used to avoid problems with MPF cells, the reduction in overlapping occurrences would not hinder the connectivity algorithm as timing adjacency could be used instead. In fact, the entire maze area could be partitioned into intersections of unique combinations of place fields. Each set of intersections would be unique to that location. There would be no overlap at all as each set would never fire anywhere else, however, connectedness could be derived using the timing method.

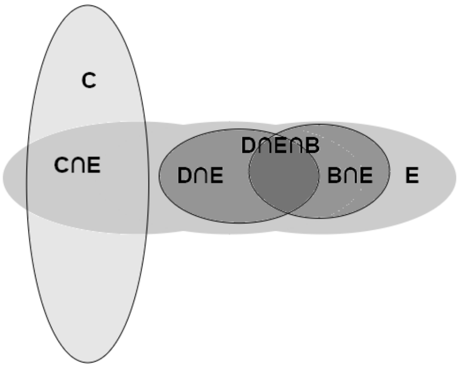


Figure 4.9: *Each unique combination of intersections represents a small region of the track. Note that all regions defined in this way are disjoint.*

There is also the possibility to use both connectivity algorithms to derive a single connectivity graph. In an ideal case, both the RCC and Timing algorithms should produce the same connectivity graph. However, both algorithms have strengths which the other does not. In a case where two fields overlap and the rat only runs through the intersection (not the first, then the intersection, then the second) the RCC algorithm would infer connectedness due to cofiring but the Timing algorithm would not. Similarly, separated fields with no overlap would only be connected in the timing algorithm not the RCC. By first generating the connectivity graph using the RCC algorithm then augmenting that graph by adding missing arcs from the Timing algorithm, a better result may be possible. Choosing to use the RCC algorithm alone, the Timing algorithm alone, or this combination represents a parameter to the solution which will be tested to see which is most effective.

4.3 Further Data Preprocessing

In order to give the connectivity and junction finding algorithms the best chance at correctly identifying topology, additional pre-processing was performed on the spike timing data.

Hidden Markov Model

It was expected that, in order to aid the estimation of topology, some smoothing would be required on the spiking data to minimise the effect of spurious firing. To achieve this, the time series of spikes from each cell was split into discrete consecutive time bins and a hidden markov model [73] was estimated with the following parameters:

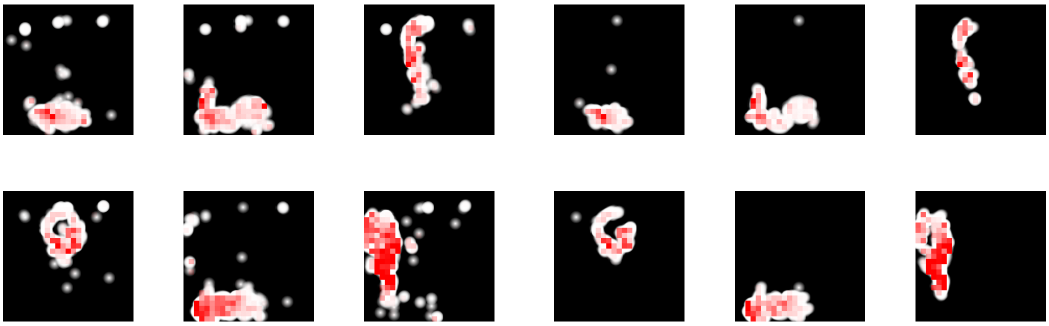
- The size of each time bin.
- The probability of moving into a place field given the rat is currently outside.
- The probability of moving out of a place field given the rat is currently inside.
- The probability of observing a spike while inside a place field.
- The probability of observing a spike while outside of a place field.

For each consecutive time bin, the probability that the rat is in each place field is estimated based on the observed number of spikes in that bin, the above parameters, and the probability estimated from the previous time bin.

This process is followed for each cell spiking file. For each time bin, if the probability that the rat is outside of the place field is greater than inside the place field, the spikes in that bin are discarded. A new smoothed cell file is then generated and stored.

As well as the Markov model smoothing, a threshold was introduced for the minimum number of spikes required for a time bin to be included in the dataset. This aim of this threshold value is to eliminate both spurious firing and firing near food sources as in these cases, the spike rate is significantly lower than in the place field itself. The parameters to the Markov model, spike rate threshold and time bin size were chosen to tightly limit the extent of place fields while retaining their general shape. With more time, a full sensitivity analysis should be performed on these parameters.

Figure 4.10



(a) *Place fields before HMM smoothing.*

(b) *Place fields after HMM smoothing.*

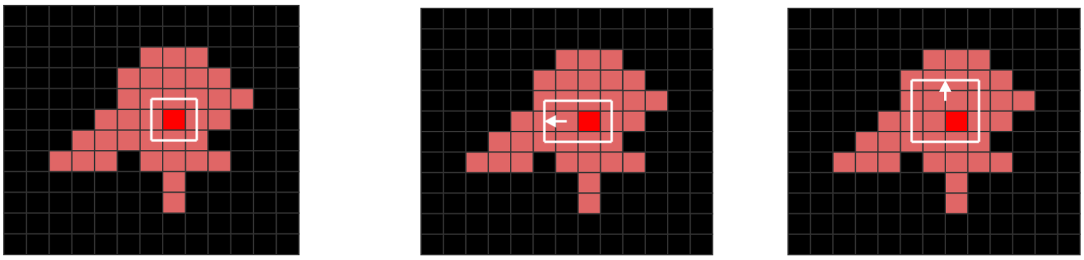
Using Position Data

The grid of red squares overlaid on the spike firing images gives a clear definition of the location of place fields. Spurious firing is generally not enough to elicit a high spike rate in a certain square in contrast to the place field itself. Additionally, the 25 x 25 grid gives a less granular measure of spike location allowing for further quick analysis to be performed on the data. In particular, distinct place fields belonging to MPF cells can be identified.

Using the position data to perform pre-processing of the input for the application (instead of purely analytical work) violates the conditions of the Spatial Reconstruction Experiment. It may be instructive, however, to see what improvement can be made when spurious firing and MPF cells are eliminated from the dataset. Eliminating spurious firing was performed by simply removing all spikes from the dataset which occurred in squares with a negligible spiking rate.

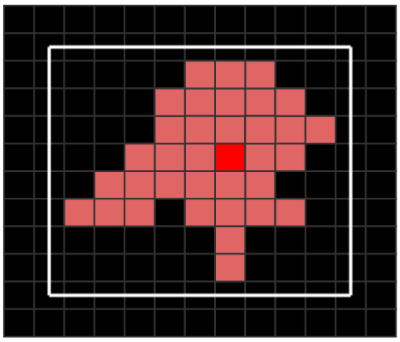
To identify MPF cells, the following algorithm was devised (figure 4.11). Find a red square in the grid and define two rows and two columns either side of this square. Each row or column will be three squares long defining a box surrounding the initial square. Scan along each row and column to see if there are any red squares. If there are, move the row one square away from the centre. Continue scanning rows and columns and extending the box until all rows and columns contain no red squares. This is the bounds of the place field.

Figure 4.11



(a) *A small bounding box is defined around the initial square.*

(b) *Rows are shifted outwards while there are red squares along the edge.*



(c) *The final bounding box.*

If there exists another red square in the grid which sits outside this bounding box, there must be more than one place field for this cell.

Ideally, finding the presence of more than one distinct place field should be enough to eliminate that cell from the data set entirely. This would be equivalent to defining an MPF cell as different to a place cell and so could plausibly be removed from the experiment. However, because there is still noise and spurious firing has not been eliminated entirely, most cells show more than one location which may be interpreted as a place field. Removing these cells entirely would leave the data set almost completely empty with no chance of making a good topology inference. It is possible that more work to consider the relative size of fields might help to separate genuine MPF cells from noisy place cells but for now, the place field with the biggest area is chosen and spikes occurring in this field are used in the dataset.

Using this bounding box technique, it is also possible to identify very large place fields which might be caused by an interneuron instead of a place cell. As discussed, interneurons fire throughout the entire session and therefore produce a place field the same size and shape of the maze.

Figure 4.12: *The place field of an interneuron covers the entire maze. The bounding box algorithm finds a box the size of the entire grid.*

This cell provides no location information and only serves to make all fields appear connected. An interneuron can be identified by the bounding box of its ‘place field’ whose area will be almost the same as that of the maze itself. As an interneuron is not a place cell, it is acceptable to remove these from the data set without violating the SRE.

In the evaluation, comparisons will be made of the performance of the application on different combinations of pre-processing techniques to gain an understanding about the limitations of the data with respect to the algorithms used.

Different Maze Topologies

Having developed the application using mainly the Frank w-maze data, it is important to be able to test it on different maze topologies as well. The second data set from the Buzsáki Lab at NYU contains spiking and position data for a number of different experiments with rats. Although the data is generally much noisier than the Frank data, success of the application with different topologies will be a key metric for evaluation.

4.4 Rest Period Data

As mentioned in the literature review (section 2.4), it has been shown that place cells firing in the hippocampus when the rat is resting or asleep, fire in the same order as in a track the rat has seen. This leads to the idea that some kind of replay is occurring among place cells after experiencing a maze. The Frank w-maze data has rest sessions before and after each track session. Testing to see if topology of a w-maze can be determined from these rest sessions is irresistible. During each rest session, the rat is placed in a high walled box for 20 minutes. The rat may be awake or asleep during this time. There are two main obstacles to deriving meaningful results from applying the algorithm to rest data. First, it is not known at what time or for how long, replay occurs in any given session. There may be periods where there are no firing place cells, or periods where they do fire but not due to replay. Only from periods of enough place cell firing, in the correct order, might we be able to derive the topology of that replay. This leads to the second obstacle which is there is no ground truth with which to confirm the result. During these sessions, the position data gives no useful information. It is also difficult to derive a statistical model for a null hypothesis in which no replay occurs without going back to simulating spike data. This is, sadly, beyond the scope of this project. An experiment by Karlsson and Frank [74] uses the same experimental data and has shown that significant replay events occur during these rest sessions while the rat is both awake and asleep. Without knowing the precise times of the replay events, it was necessary to define a time window capturing only a small section of spike times within the entire rest session. From this time window, connectivity and topology graphs could be generated in the same way as with the track runs although, without any position data, interneurons and MPFs are not identifiable from the rest data. The time window was then scanned across the full spike time series and a number of different window sizes were examined. Time windows generating topology graphs matching the topology of the previous run maze could be identified and checked to confirm that cell order was correct. In order to check correctness, paths of length greater than 2 in the rest window topology graphs were compared to see if equivalent paths were present in the topology graphs of the previous run.

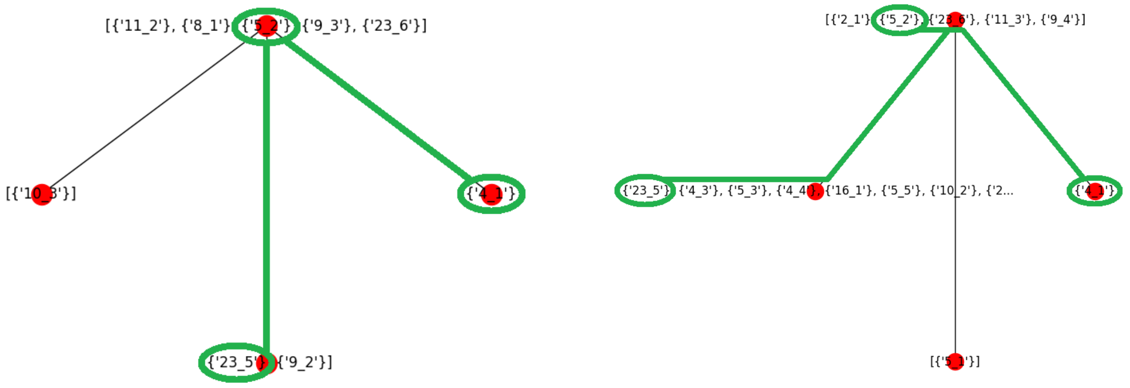


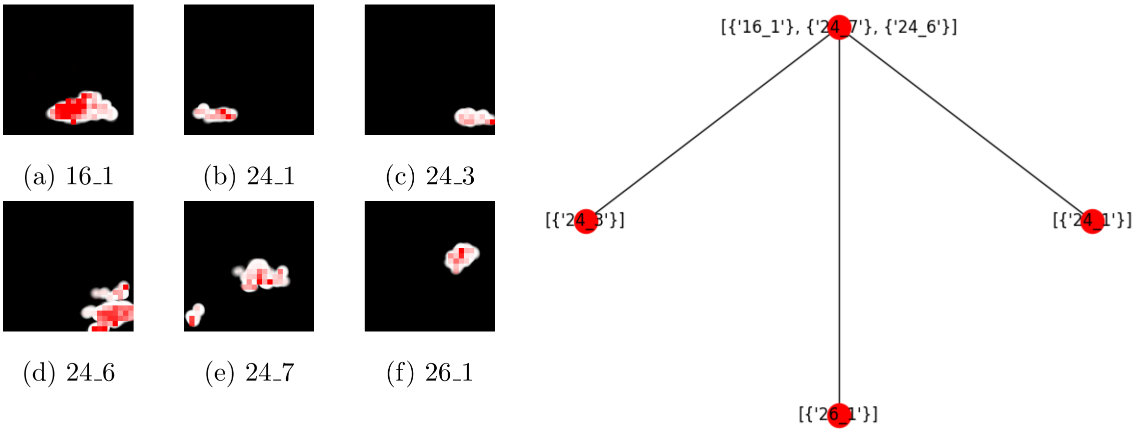
Figure 4.13: *(left) A topology graph from a run session. (right) A topology graph from the next rest session. Green lines and circles identify a matching path between the two.*

The existence of any matching paths might indicate that this window contained a replay event and that the application had at least partially deduced the topology.

5. Results and Evaluation

5.1 Methodology

Initially, evaluation of the application was planned to be carried out in a similar fashion to that of a classifier. Running the application on all maze shapes together with a definition of what the topology graph should look like for each, would enable the creation of a confusion matrix. Then various metrics such as recall and precision could be used to evaluate performance. However, there are a number of ways in which error can be introduced (as summarised in the Data Analysis, section 3.3) and the similarities between each correct shape topology are such that any small error would immediately change the 'classification'. The significance of a topology graph with a 4-junction, generated from a w-maze session, is not that it has been mistaken for a Plus-shaped maze. Furthermore, the success of the application is likely to be reliant on the quality of the input data which is extremely variable across all experiments and sessions. For these reasons, a more qualitative approach to evaluation is required, so for each set of results, common causes of error will be examined. To provide a quantitative measure of success, a definition of 'correct' topology will be given for the topology of each maze shape and an accuracy value will be calculated based on how many correct topology graphs are generated from the data set. Although for convenience this will be referred to as 'accuracy', correctness will only be dependent on the number and degree of nodes in the topology graph. Because of this, there is an opportunity for false positives to occur where the correctness definition is satisfied but the actual ordering of fields is incorrect. Figure 5.1 shows an example where it is possible for the application to derive a topology graph with a single junction of degree 3 for the w-maze, which would be declared correct, even though the session in question has no place fields in one of the arms and the application mistakenly counted a third branch elsewhere.


(a) 16_1 (b) 24_1 (c) 24_3

(d) 24_6 (e) 24_7 (f) 26_1

Figure 5.1: *Fields labelled (a)-(f) cover only the bottom arm and central junction of a backwards E maze yet in certain circumstances, a topology graph with a 3-junction can be incorrectly derived.*

This value will be particularly useful for comparing the various techniques and parameters used for each result set. The evaluation will proceed by analysing results produced from the same data set but with differing pre-processing techniques. These are

- W-maze data with no pre-processing, spike time files taken directly from the experiments.
- W-maze data with HMM smoothing.
- W-maze data with HMM smoothing and exclusion of interneurons.
- W-maze data with HMM smoothing, exclusion of interneurons and the reduction of MPFs to single fields using position data.

Finally, performance on different shaped mazes with the most successful preprocessing techniques will be evaluated as well as results from their application to the rest data.

Expected Maze Topologies

For the quantitative measure of performance, the definitions of 'correct' topology are given here.

W-maze

A correct topology graph for the w-maze will have a single 3-junction, three 1-junctions (dead ends) and $(n-4)$ 2-junctions where n is the number of nodes in the graph.

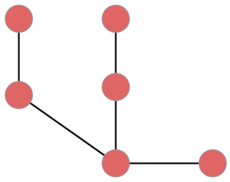


Figure 5.2: *A correct W shaped maze topology graph.*

Linear

A linear maze, for example, one which is a straight corridor and nothing else or perhaps a U shaped maze will have two 1-junctions and $(n-2)$ 2-junctions. An O shaped maze, which has no dead ends and therefore all nodes are 2-junctions, is not considered here as there are no examples of this in the data.

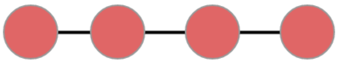


Figure 5.3: *A correct linear maze topology graph.*

Plus

A plus or cross shaped maze will have a single 4-junction, four 1-junctions and $(n-5)$ 2-junctions.

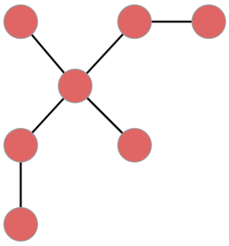


Figure 5.4: *A correct plus shaped maze topology graph.*

Figure-8

The figure-8 maze has two potential configurations depending on whether the central crossover area is considered a single 4-junction such as with the infinity symbol, ∞, or two separate 3-junctions such as with a B shape or an 8 on a digital clock. Due to the two loops, there should be no 1-junctions (no dead ends) but $(n-1)$ or $(n-2)$ 2-junctions. As the cycle finding step is a separate process and may be evaluated separately from the topology graph algorithm, any number of 1-junctions will be permitted with only the 3 or 4-junctions required for a correct graph.

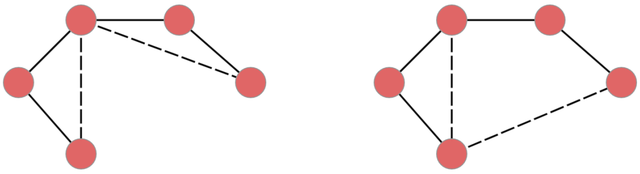


Figure 5.5: *(left) Figure-8 with a single 4-junction. (right) Figure-8 with two 3-junctions.*

Arenas

It is not obvious what the correct definition for topology should be for a large open area and there may be no strict number of junctions. A purely qualitative analysis will be carried out for arenas.

Algorithm Parameters

As discussed throughout chapter 4 the following parameters will be tested on each data set and the average accuracy across all experimental sessions for all combinations of parameters will be recorded.

Place Field Intersections

As discussed in 4.2, it is possible to stipulate that the connectivity algorithms use intersections of fields, rather than the fields themselves, for generating a connectivity graph. To explore the impact of this stipulation, each result set will be derived from connectivity between 1-field (i.e no intersections), 2-field or 3-field intersections. For these values, every node in the connectivity graph will list 1, 2 or 3 cells respectively.

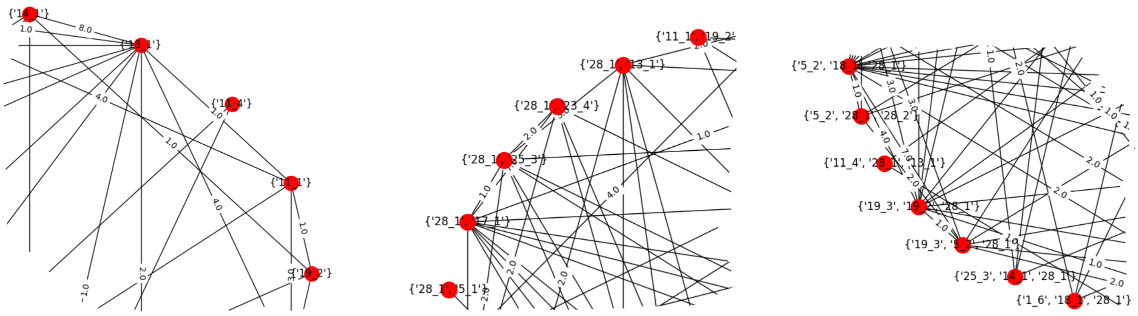


Figure 5.6: *(left) A partial connectivity graph with individual fields. (centre) 2-field intersections. (right) 3-field intersections.*

A combination of all three will also be tried such that connectivity can be derived from areas where only a single cell is firing; where only two cells are cofiring and where only three cells are cofiring. The connectivity graph will contain nodes with lists of 1, 2 and 3 cells together.

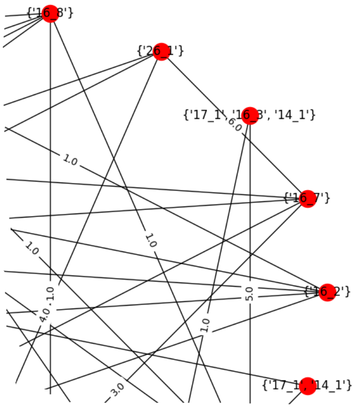


Figure 5.7: *The partial connectivity graph using 1, 2 and 3-field intersections.*

Connectivity Graph Algorithm

There are three options for how the connectivity graph is generated from the time series of spiking cells. These are, the RCC algorithm, the Timing algorithm or a combination of both (outlined in section 4.2). All three will be tested in combination with other parameters.

Connectivity Graph Weight Threshold

The threshold for weights in the connectivity graph, below which arcs should be removed, may reduce error introduced by spurious cofiring. The degree to which this aids performance will be explored with a set of low valued thresholds 0, 1, 2, 3 and 5. Higher thresholds than this are expected to encroach on genuine cofiring.

Branch Cell Threshold for the Topology Graph

To investigate the benefit of removing erroneous branches in the topology graph due to small sets of unconnected fields, the minimum number of cells allowed in a node at the end of a branch will be set to 0, 1 and 2.

5.2 Results for Unprocessed W-maze Data

With no preprocessing of the spike data from the Frank w-maze experiments, that is no smoothing, elimination of MPF cells or interneurons, it was expected that performance would be poor as explained in the previous project. This was confirmed with only 9.71% of topology graphs showing a single 3-junction when the application was run with all parameter combinations. Figure 5.8 shows an example of a successful topology graph. Using the RCC connectivity algorithm with 2-field intersections.

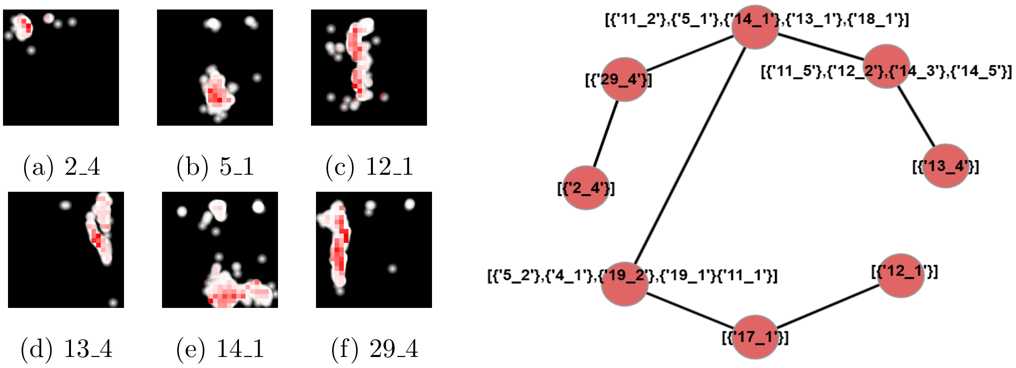

(a) 2_4 (b) 5_1 (c) 12_1
(d) 13_4 (e) 14_1 (f) 29_4

Figure 5.8: *A correctly deduced topology graph. Each node is labelled with the identifiers of cells whose place fields are adjacent. In this case, this includes two identifiers as 2-field intersections have been used.*

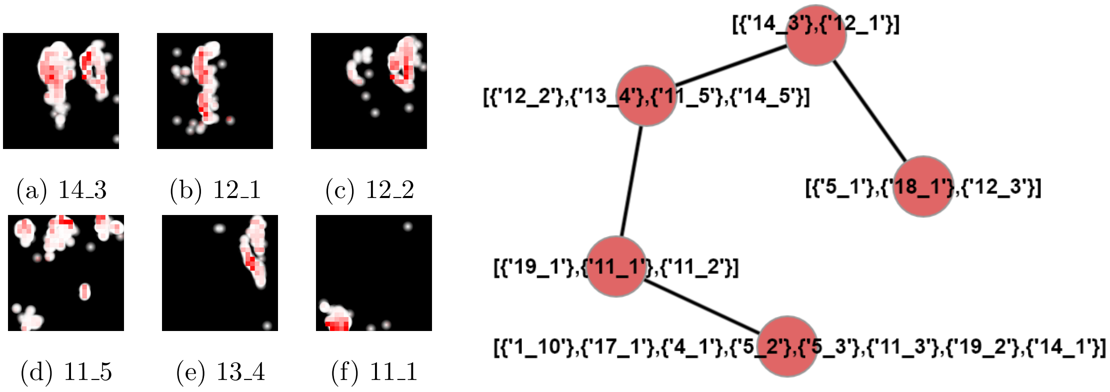
Figure 5.9: *Accuracy Results for Unprocessed W-maze Data*

Connectivity Algorithm	**% correct**	**Field Intersection**	**% correct**
RCC	9.69	1-Intersection	8.21
Timing	8.77	2-Intersection	9.09
RCC&Timing	10.94	3-Intersection	10.98
		1,2,3-Intersection	10.94
Branch Cell Threshold	**% correct**	**Connectivity Weight Threshold**	**% correct**
Branch-0	9.36	Weight-0	9.8
Branch-1	10.92	Weight-1	10.47
Branch-2	9.11	Weight-2	10.35
		Weight-3	9.34
		Weight-5	9.79

The best performing single configuration of parameters with a mean accuracy of 21.8% for the unprocessed data, was using the Timing only connectivity algorithm over individual fields with a connectivity weight threshold of 1 and no branch size threshold. Although the Timing algorithm shows the lowest average accuracy, it also has the highest variance and shows consistently good accuracy with individual fields and 0 branch threshold (mean 17.87%). The timing algorithm benefits from being able to include individual fields in the arms of the maze which do not overlap with those in the rest of the track. This is a common occurrence as there are often fewer fields recorded in the right and left arms of the w-maze. A branch threshold value greater than 0 might discount these extreme fields from the topology graph but this must be balanced with the introduction of erroneous branches elsewhere in the maze. In the case of using the Timing algorithm, it seems, the benefits of a 0 branch threshold outweigh the negatives despite a general improvement in performance for all parameters with the threshold set to 1.

Multiple Place Fields

As expected from data without MPF cells and interneurons excluded, using individual fields for connectivity instead of 2-field or 3-field intersections shows a lower accuracy than average. Figure 5.10 shows an example of how MPF cells can cause an incorrect connectivity graph to be derived from individual fields. Here, the connectivity suffers twice. MPF cell 14_3 is connected to 12_1 in the central arm and 12_2 and 13_4 which are in the right arm of the maze. Cell 11_5 has fields at the top of all arms and in the bottom left corner of the maze and connects 13_4 and 12_2 to 11_1.



(a) 14_3 (b) 12_1 (c) 12_2

(d) 11_5 (e) 13_4 (f) 11_1

Figure 5.10: *The topology graph incorrectly links fields from 11_1, 12_2, 12_1 and 13_4.*

With higher field intersections there is also a greater number of possible combinations of fields and this leads to more nodes in the connectivity graph assuming there is a large number of overlapping fields to begin with. It is possible that this higher number of nodes is another cause for better accuracy.

Large ‘Place Fields’ from Interneurons

Figure 5.11 shows the presence of an interneuron in the spike data causing confusion in the connectivity graph leading to an incorrect topology graph. The connectivity graph shows all place fields overlap with that of cell 5_5 which is most likely an interneuron. A topology graph with a single high degree node is indicative of this problem. When the iterative step of the junction finding algorithm reaches the interneurons field (which should be the second iteration for any field chosen as a starting position is always connected), all remaining fields are considered to be disconnected subsets.

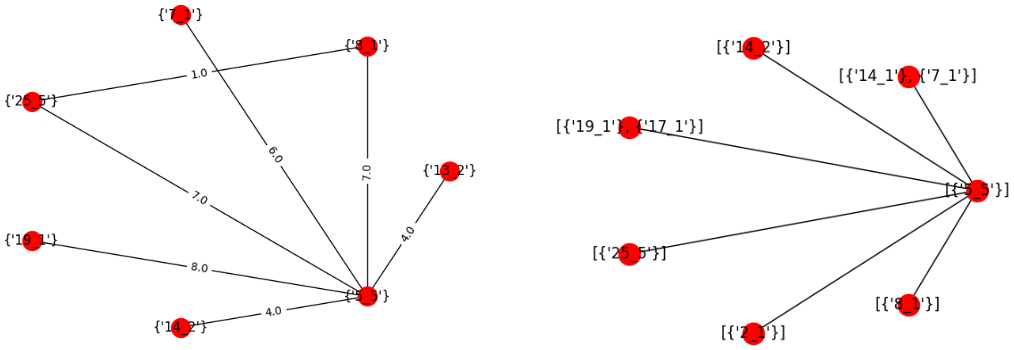


Figure 5.11: *(left) The connectivity graph shows all fields connected to 5_5. (right) The resulting topology graph.*

Provided there is only a single interneuron in the data set, the 2 and 3-field intersection parameters should provide better results than the individual fields. This is because the single 5_5 field will no longer be considered to connect to everything else. Instead, in the 2-field example, all nodes in the connectivity graph will represent intersections of the smaller fields with 5_5. As the smaller fields are entirely subsumed by the interneuron, the intersections will be equivalent to the smaller fields as shown in figure 5.12.

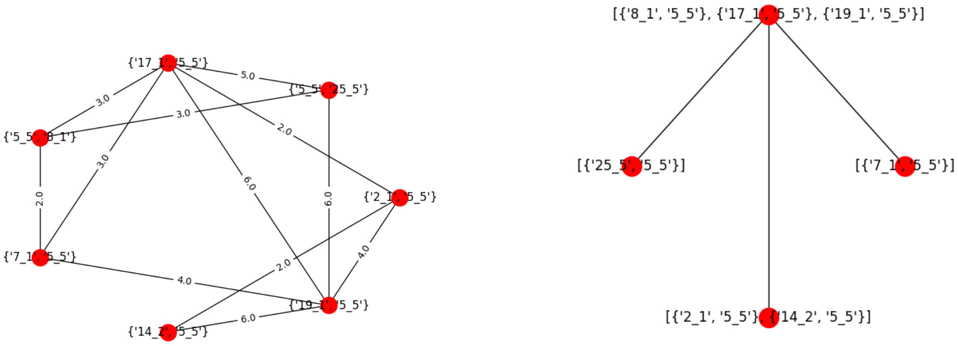


Figure 5.12: *(left) The corrected connectivity graph using 2-field intersections so all notes include 5_5. (right) The resulting topology graph.*

5.3 W-maze Data with HMM Smoothing

With strong HMM smoothing, there is a much lower chance of seeing spurious firing. The fields themselves are also much smaller but consistently form a well defined area. For this result set, there is no check for interneurons or MPFs and so they still appear in the dataset. With the smoothing, the average accuracy across all parameters has increased to 15.9%.

Figure 5.13: *Accuracy Results for HMM Smoothed Data*

Connectivity Algorithm	**% correct**	**Field Intersection**	**% correct**
RCC	14.97	1-Intersection	22.1
Timing	16.61	2-Intersection	12.48
RCC&Timing	16.12	3-Intersection	11.94
		1,2,3-Intersection	18.66

Branch Cell Threshold	**% correct**	**Connectivity Weight Threshold**	**% correct**
Branch-0	16.01	Weight-0	23.5
Branch-1	17.1	Weight-1	17.3
Branch-2	17.1	Weight-2	14.6
		Weight-3	14.7
		Weight-5	13.3

Once again, the Timing only connectivity algorithm with individual place fields performs best giving an average accuracy of 38.27%. Looking at the connectivity weight threshold, a 0 weight greatly out performs higher values. Because of the HMM smoothing, there is far less spurious firing and the need for a weight threshold in the connectivity graph is diminished. At higher values, the threshold begins to remove genuine connectivity information instead of just disregarding noise.

5.4 Exclusion of Interneurons

Using the rat position data to identify and exclude interneurons from the spike data further reduces the error introduced from an overly connected connectivity graph as discussed earlier. Overall, the average accuracy after this additional preprocessing step increases to 21.4%.

Figure 5.14: *Accuracy Results for HMM Smoothed Data with Interneurons Excluded*

Connectivity Algorithm	**% correct**	**Field Intersection**	**% correct**
RCC	11.1	1-Intersection	19.2
Timing	26.4	2-Intersection	3.5
RCC&Timing	26.8	3-Intersection	0.0
		1,2,3-Intersection	24.1

Branch Cell Threshold	**% correct**	**Connectivity Weight Threshold**	**% correct**
Branch-0	26.9	Weight-0	36.8
Branch-1	18.7	Weight-1	22.7
Branch-2	18.7	Weight-2	20.9
		Weight-3	18.7
		Weight-5	8.3

The Timing only connectivity algorithm with individual fields seems to consistently take the top spot for accuracy. In this case, 58.05%. In all experiments including this, the combination of Timing & RCC algorithms performs comparatively well. This is unsurprising as the Timing connectivity is a subgraph of the Timing & RCC graph. As there is no increase in accuracy in the Timing & RCC version over the Timing only version, the additional RCC arcs seem to provide no additional information and may, in fact, introduce some error.

The two graph threshold values, while showing some utility with the unprocessed results, show a clear impedance to accuracy with smoothing and interneuron exclusion. This is encouraging as fewer parameters speaks to the robustness of the algorithm itself and indicates that the problems they were designed to suppress have been eliminated in the preprocessing steps at least for this data set. The 2 and 3-field intersection results and those using the RCC connectivity algorithm are quite poor when compared to the individual field and Timing results. Although it is tempting to eliminate these parameters, it is probable that they would show better performance with larger data sets and more place fields. The timing algorithm was designed as a salvo against gaps in place field overlap. With a larger number of place cells, there would be less chance of these gaps and it is expected that the RCC algorithm would become as performant as Timing.

The accuracy results for the RCC algorithm with combined 1, 2 and 3-field intersections are also comparable to those of the Timing algorithms even though RCC with individual fields gives very poor results. Because the 1, 2 and 3-field intersections lead to a much larger population of nodes in the connectivity graph, this may be evidence that the RCC algorithm requires more data to perform well.

5.5 Reduction of MPF Cells

As well as interneurons, the existence of MPF cells is expected to cause errors as described earlier in section 3.3. This experiment applies HMM smoothing, removal of interneurons and reduction of MPF cells to the dataset before running the application. Although it can be argued that elimination of MPF cells from the dataset is not in violation of the spatial reconstruction experiment, using the position data to identify and record spikes in only one of the multiple fields of a cell certainly is. However, it would be instructive to see the effect on accuracy of having no multiple place fields without reducing the number of recorded place cells.

Figure 5.15: *Accuracy Results for HMM Smoothed Data with Interneurons Excluded and Reduced MPFs*

Connectivity Algorithm	**% correct**	**Field Intersection**	**% correct**
RCC	11.4	1-Intersection	18.4
Timing	25.3	2-Intersection	2.9
RCC&Timing	27.1	3-Intersection	0.0
		1,2,3-Intersection	24.16

Branch Cell Threshold	**% correct**	**Connectivity Weight Threshold**	**% correct**
Branch-0	30.3	Weight-0	31.4
Branch-1	16.7	Weight-1	25.1
Branch-2	16.7	Weight-2	23.1
		Weight-3	18.9
		Weight-5	7.9

There is a slight reduction in accuracy overall with an average of 21.0%. The Timing only connectivity algorithm with individual fields still performs the best at 51.27%. Figure 5.16 shows how multiple place fields could give more opportunities for the junction finding algorithm to produce a graph with a single 3-junction despite inferring an incorrect field ordering. Removing MPFs denies the possibility for these false positives resulting in a lower percentage of topology graphs declared correct despite being more reliable. If there were more time, it would be important to find a way to define correctness based on ordering of fields so this kind of confusion is avoided.

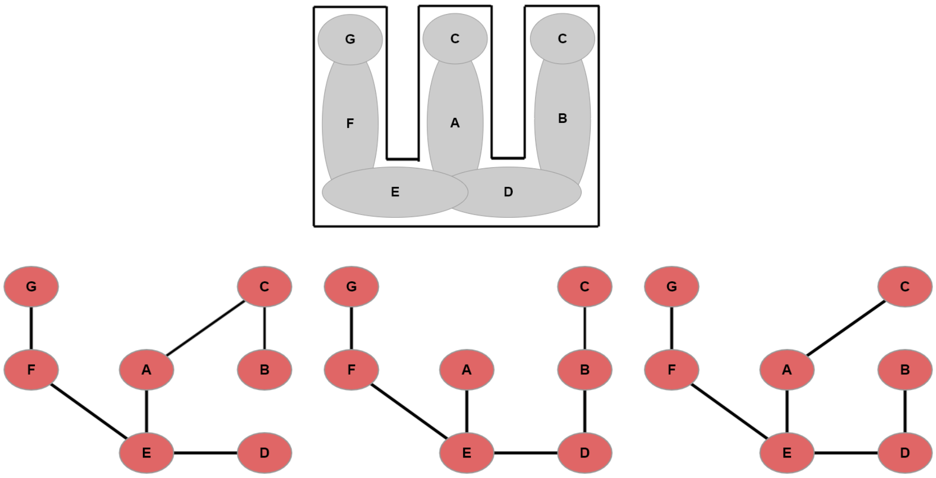


Figure 5.16: *(above) A w-maze with place fields for cells A-G. C is an MPF cell. (below) Three potentially correct topology graphs from the same maze.*

Rat Path Coverage

Figure 5.17 shows an example of how the path which the rat takes through the maze can cause errors in the connectivity and topology graphs even with the previous preprocessing applied. Cells 25_2 and 17_3 have place fields in the same arm of the maze. The field of cell 14_1 spans the width of the arm towards the central junction and so whenever the rat runs up this arm, it passes through 14_1. However, on some occasions, the rat runs up and down the right hand side of the arm through the field of 25_2 and does not pass through the field of 17_3. On other occasions, the rat runs up and down the left side of the arm only passing through 17_3, avoiding 25_2. As a result, these two cells never cofire or fire consecutively. Both connectivity algorithms, therefore, interpret that 17_3 and 25_2 are not connected and the junction finding algorithm derives two separate arms in the maze. This is a highly common feature in many of the experimental sessions and is not remedied by any of the processing techniques explored. If it were the case that another place cell had a field which overlapped or was adjacent to both 17_3 and 25_2 and the rat passed through this on all occasions, the problem would be solved as the connectivity graph would show the two cells connected via this additional third cell. It is probable that such a cell exists and has simply not been recorded, therefore it is safe to say that a greater number of recorded cells would reduce the error caused by uncertainty in the path of the rat.

(a) 14_1 (b) 17_3 (c) 25_2

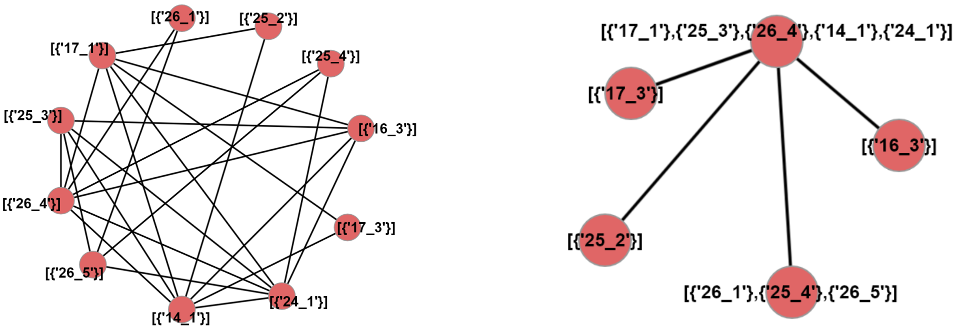


Figure 5.17: *The place fields in (a), (b) and (c) all sit in the same arm of a maze. (lower left) The connectivity graph - 17_3 and 25_2 are not connected. (lower right) The topology graph has a branch for 17_3 and one for 25_2.*

It is also possible to consider a case in which the rat consistently runs up and down one side, or the other, of a wide corridor without travelling through the middle. Regardless of cell population, it would appear to the application as two separate arms. This introduces a caveat to the spatial reconstruction experiment such that the environment must be fully traversed by the rat or that the spatial information inferred describes the movements of the rat itself and not the environment.

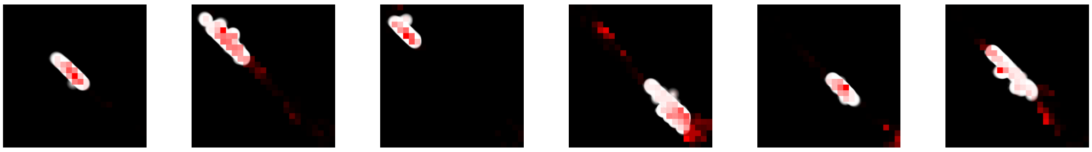
As well as MPF cells, this is another characteristic which regularly introduces false positives into the results. If the left arm of a w-maze is misclassified as two separate arms and the right arm of the maze is not identified due to some other error, counting the central arm and two arms for the left would result in a seemingly correct w-maze topology graph.

5.6 Other Maze Configurations

Having found successful preprocessing techniques to improve the accuracy of the application, they can be tested on different maze configurations. Due to the greater amount of noise than the Frank w-maze data and in some cases fewer available place cells, the average accuracy across all configurations (not including arenas) is only around 8%. However, certain parameter combinations find success with the linear and figure-8 tracks over 35% of the time.

Linear Track

In the trivial case of a straight track in which the rat moves from one end to the other, there is still the possibility for the rat's path to follow the walls of the corridor leading to disconnected fields on either side. Additionally, different recording and experimental techniques may lead to a different level of inconsistent firing to that of the Frank w-maze data which may affect the outcome. However, figure 5.18 shows an example of a successful inference resulting in a topology graph with a single path between two nodes of degree one.



(a) 5_5 (b) 5_8 (c) 5_9 (d) 8_8 (e) 8_10 (f) 8_11

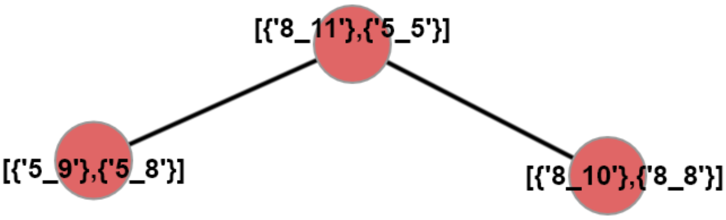
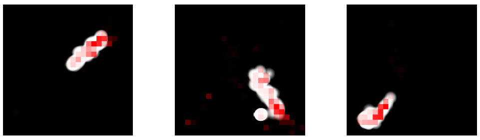


Figure 5.18

Plus Shaped Track

The data from the plus shaped maze experiments provide the fewest recorded cells per session. Figure 5.19 shows one of the better sessions. There is very little activity in the top left arm and only in a very small area. The fourth arm of the topology graph holds cell 6_7 whose field somewhat extends into the top left arm, yielding a correct result.

(a) 5_11 (b) 5_13 (c) 6_7



(d) 6_9 (e) 6_13 (f) 7_12

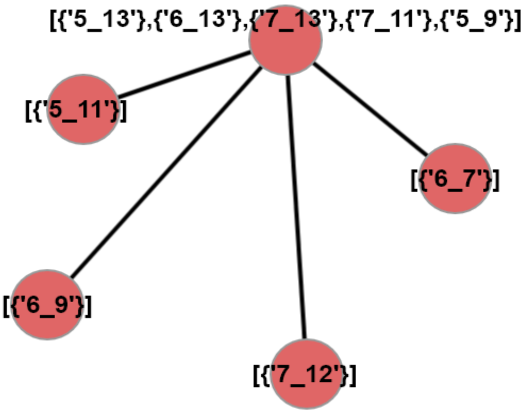
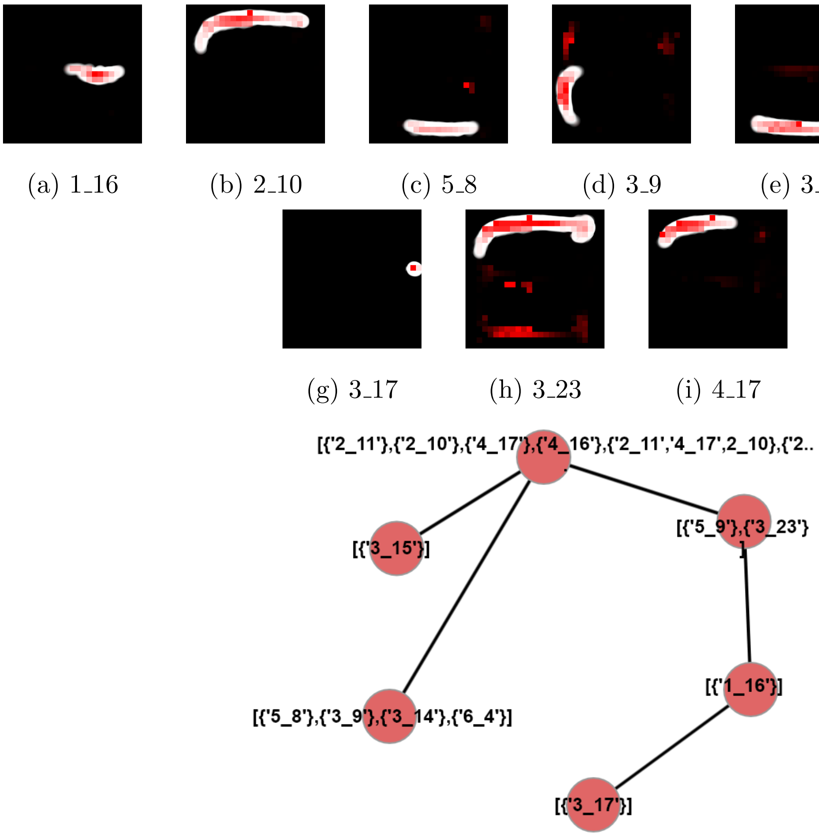


Figure 5.19

Figure-8 Track

The figure-8 track is the only example of a configuration with cycles and so there is a role for the additional step of connecting nodes in the topology graph together based on connected fields. Figure 5.20 presents a good example of data from the continuous alternation task in the Buzsáki experiments. The red arcs have been added by the additional step as potential arcs which complete a cycle.


(a) 1_16 (b) 2_10 (c) 5_8 (d) 3_9 (e) 3_14 (f) 3_15

(g) 3_17 (h) 3_23 (i) 4_17

Figure 5.20

In this example, the application identifies a 3-junction and a 4-junction in the topology graph and two loops. Both loops are small as many fields are collected in a single node due to a lower connectivity weight threshold which produces a more connected graph. If the threshold were higher, gaps would appear in the connectivity graph as there is a great deal of inconsistent firing in the figure-8 experiment. This could be because the place fields have not stabilised (if the rat is new to this environment) or if the place cells are firing according to future alternation decisions as mentioned in the literature review, section 2.3. Cell 3_15 has been incorrectly identified as a separate arm of the graph. However, the two junction nodes do represent fields on either end of the central arm of the maze as would be expected. The node containing cell 3_23 represents the top arm and the node containing 5_8, 3_9 and 3_14 represents the bottom arm.

Arena

There is a great deal of investigation to be made into the topology graphs expressed by this application for large open areas or arenas. Unfortunately there is only time to consider a single example here. Figure 5.21 shows an example of a generated topology graph from an arena. In its current form, if the junction finding algorithm were applied to a fully covered square arena, one might expect the following outcome. Starting from a single field, perhaps somewhere in the centre of the arena, the initial topology node would contain all fields surrounding it. Another node would be defined for all the fields surrounding those and so on producing a single straight path. Finally, as fields in the corners of the square would be disconnected from those in other corners, the path would split into four separate branches. Figure 6.1 in section 6.1 illustrates the outcome of this explanation. To some degree, this can be observed in the results below. The fields in the nodes at the end of the branches (4_5, 2_9, 1_23 and 3_15, 1_15 and 2_14) are all close to the edges of the arena. Fields in the node of degree 2, however, are not fully surrounded by fields in the node of degree 4. This is because the starting place field was not in the absolute centre of the arena or perhaps due to insufficient coverage in some areas.

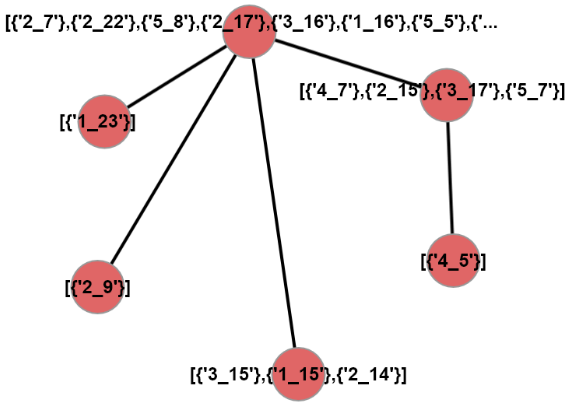


Figure 5.21

It seems incorrect for the concentric circles of fields to form a linear path and surprising that the number of branches may depend on the number of corners in the arena. This will be discussed in section 6.1.

5.7 Rest Period Data

It has been shown [74] that replay is a fairly common occurrence during the rest sessions of the Frank data by comparing the order of firing of pairs of cells. In order to discern topology, however, a stricter requirement was introduced such that a topology graph, generated from a time window during a rest session, must contain at least one path of length 2 or more which is also present in the topology graph of the preceding run session. If this requirement is fulfilled, it is an indicator that not only is there replay of the previous session but that there may be sufficient data to infer the topology.

Sessions from four rats were tested. Only two showed any sufficient replay with only 13 out of 80 sessions passing the requirement. From these 13 sessions, however, the application deduced a w-maze topology 23.4% of the time. Figure 5.22 shows an example of a correct deduction. The graph on the left is the topology graph generated from the previous run session in which place cells were firing as the rat ran around the maze. The graph on the right is the topology graph generated from a window of time during the rest session which followed. Place cells were firing due to replay while the rat was stationary or moving within the small box. The 3-junctions of both graphs share the place cell 6_2 and both have a branch with the single cell 9_1. There are a number of matching cells in the branches with the nodes containing many fields. There is disagreement on the final branch for cells 5_4 and 5_3 however, the fields for each of these cells are in a similar location and a small difference in firing could account for this confusion.

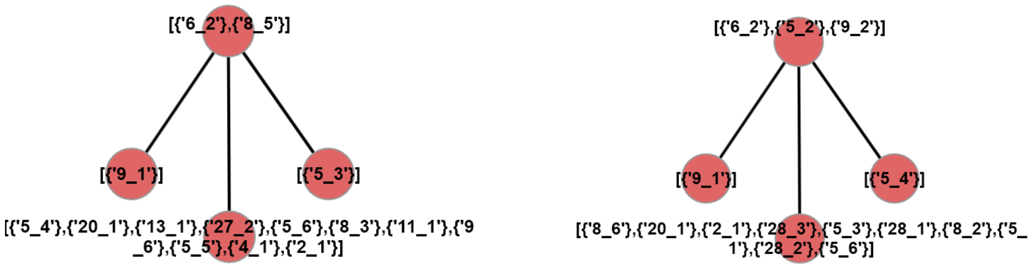


Figure 5.22: *(left) Topology graph from a single run session. (right) Topology graph from a time window in the following rest session.*

It would be beneficial to be able to calculate a significance value for the likelihood of such matching graphs but it is unclear how a null hypothesis would be formulated. This will be discussed in section 6.4.

6. Future Work

6.1 Arenas

In section 5.6, how a square arena might be represented in the graph output from the junction finding algorithm, was discussed. An example from the experimental data supported this explanation. However, two questions were raised about the implications of these results. Firstly, the representation of concentric rings of place fields as a single path of nodes in the topology graph is counter intuitive. Such a structure in the graph can also represent a linear track or corridor. One solution might be, during the junction finding algorithm, in the case where a single group of connected fields has been found, to add those fields to the currently tested node. Currently, a new node is created for the connected fields. In the case of an arena, concentric circles of fields would be reduced to a single node in the topology graph containing all fields excluding those in the corners, as they would form more than one set of connected fields with the outermost circle. This would also have the effect of reducing all corridors to a single arc in the graph regardless of the shape of the maze. It is perhaps more topologically sound and would change the significance of an arena to that of a junction. The ideal topology of a square arena with this adjustment would be a single node of degree 4, similar to the plus shaped maze. Instead of eliminating arenas entirely, bearing in mind that more information about the environment is the aim of this project, it would be prudent to have a separate algorithm for identifying an arena. From the results of the improved junction finding algorithm above, fields that belong to a junction or an arena would have been automatically identified. They are the fields in junction nodes. By performing the *old* version of the junction finding algorithm only on these fields, a single path would be generated as before but the significance of the length of the path is now a measure of the extent of the junction. A path of length 0 or 1 might indicate a small junction between corridors but a path of 2 or greater might indicate a larger open arena. If experimental data can be found with arenas of differing sizes, the study of the effect of arena size on path length might yield interesting results, particularly with respect to the elastic nature of place fields (see section 2.1).

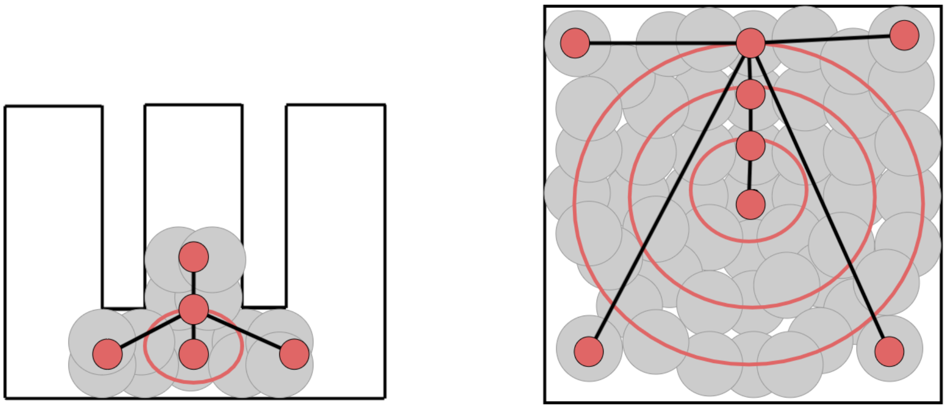


Figure 6.1: *(left) A tight junction with only a single circle of fields in red. (right) An arena with multiple concentric rings of place fields.*

The second question raised from the results of arena testing is that of the significance of disconnected fields at the edges or corners of an arena. In the idealised case of a fully covered square arena, it might be expected that fields in the four corners would not be considered connected by the junction finding algorithm, because fields in the centre and along the middle of the edges would have been collected in previous nodes. A circular arena with four corridors leading away would look the same to the algorithm. Whether these corner branches should be treated as artifacts of the algorithm, or in some way fulfilling the criteria for small corridors leaving the arena, should be investigated with different shapes and with corridors leading away.

6.2 Loops and Holes

The current additional step in the junction finding algorithm to add bridging arcs to the graph when identifying cycles could be designed in a more nuanced way. The paper by Dabaghian et al. [16] outlined a method to generate a simplicial complex (section 2.2) based on the overlap of place fields which would be equivalent to using the connectivity graph from this solution. This method offers an alternative to the junction finding algorithm and has the benefit of more robustly defining ‘holes’. For example, the figure 8 maze would contain two such holes and the w-maze would have no holes. The number of holes would give an indication of the presence of loops in the maze and identify the fields involved. However, given the extent of the coverage of fields in the experimental data, it is possible that incorrect holes could be identified due to missing place fields.

6.3 Stable Place Fields Across Multiple Sessions

It was discussed in section 3.3 that there are certain cells, although by no means all, whose place fields continue to appear in the same place in the maze across multiple sessions and days of experimenting. To improve the performance of the solution, it may be possible to use the spiking data of the same cells across multiple sessions. There are three main obstacles to this. Not all cells fire during all sessions. One session would need to be selected for the initial data set, then when taking cell data from subsequent sessions, only cells which also fired in the initial session would need to be analysed. This leads to the second obstacle in the case that the Timing algorithm is used for connectivity. Let there be three cells A, B and C in the initial session such that cell A fires, then B, then C and the reverse. The connectivity algorithm would infer that A is connected to B, B is connected to C but A is not connected to C. If, in a subsequent session, cells A and C show activity but not B, the Timing connectivity might infer that A is connected to C. The RCC algorithm would not suffer from this problem, however, and may well benefit from the additional data. The final obstacle is that in some cases a cell might still show activity but in an entirely different location which would do nothing but introduce error into the system. The benefit of additional information might be enough to outweigh this problem.

6.4 Rest Data

The examples of replay topology graphs matching those of the previous run session given in section 5.7 are compelling. However, it was unclear how to measure the significance of the findings. The normal way in which this would be done, would be to define a null hypothesis saying that the matching results occurred as often as expected if there were no replay of cells from the previous session. That is to say, the matching topology graphs occurred by chance. A model could be devised to identify the expected number of occurrences were this the case. The proportion of matching graphs in the results could then be compared to this null hypothesis model and a p-value calculated. The difficulty comes from how to define the null hypothesis model. Further discussion with neuroscientists or in-depth review of previous research might yield a plausible model. It may be acceptable to assume that all possible graph configurations among the available place cells are equally likely. If the null hypothesis is valid, the proportion of occurrences of, say, a single 3-junction should match the number of possible configurations containing a 3-junction. This is only one of many ways of defining a potential model.

If the findings do appear to be significant, it would be interesting to use the position data to identify specific periods of sleep as outlined in [74]. Periods of time where the rat does not move for 5 minutes or more could be conservatively defined as sleep periods. Occurrences of matching topology graphs could be compared to those found when the rat was only still for a short time or while moving around the resting box. It may also be possible to compare the topology of replay fields with the graphs of run sessions performed further back in time. It may be that the replay is of an earlier session than just the one immediately before. In particular, there may be occurrences of replay of track A from the w-maze data in rest sessions immediately after run sessions in track B.

Finally, it may be informative to find a quantitative way to describe the degree to which the sequence of place cell firing has changed between a run session and the subsequent rest session. This could be done by analysing the connectivity graph instead of the topology graph and counting matching arcs between those of both sessions.

6.5 Context and Activity Analysis

The main aim of this project was to identify the spatial aspects of the rat's environment. This involved filtering out what was deemed spurious cell firing in areas where a food goal was reached or where a decision was frequently being made (for example at the junction of the w-maze where the rat must choose to go left or right according to its training). These areas are still, however, additional properties of the environment which may be recognisable from the firing data alone. By analysing spikes which do not occur within the place fields, as defined by the HMM smoothing technique, it should be possible to find common areas of 'spurious' firing. Additionally, once the topology graph has been generated, it should be possible to trace the location of the rat during a session as it moves from the set of fields in one node to the next. From this trace it should be clear in the w-maze data that the rat is performing an alternation task. Alternatively, having been told it is an alternation task, it should be possible to identify the times where the rat made a mistake and went to the wrong arm.

6.6 Combination with Grid Cells

Some of the Buzsáki experimental data sets, as well as containing recordings from CA1 and CA3 regions of the hippocampus, also have recordings from the entorhinal cortex. As described in section 2.4, the grid cells in the entorhinal cortex provide a more regular, strictly spatial representation of the environment with less of the 'context sensitivity' of hippocampal place cells. With another source of data describing the spatial location of the rat and the geometric properties of the environment, it may be possible to explore the contextual aspects of place cells more thoroughly. The location of the rat in a maze may be deduced from the firing of a place cell indicating that the rat is most likely in its place field. The position could be confirmed by the combination of cofiring grid cells. With the assumption that the grid cells are a more reliable indicator of location only, if it is the case that these two locations disagree, it it likely because the rat is actually outside the place field and the place cell is firing for some other reason. Investigating the conditions of these mismatched locations may inform on the contextual nature of place cells.

6.7 Additional Physical Information

There is evidence that the size of a place field is linked to the physical location of the place cell in the hippocampus (see 2.1) with place cells in layers having similar sized fields. As seen in the place field images throughout this project, there are many cases where small place fields are entirely subsumed by larger fields and the algorithm itself depends on the existence of overlaps between fields. This implies that there is a great deal of redundancy among place cells if the main function is to represent location. Perhaps it is the case that cells with large place fields provide some method of fast indexing of location with more specificity coming from cells with smaller fields. If this is true, it should be possible given enough data, to split place cells into sets according the place field size and derive spatial characteristics at differing levels of detail. Experiments using mazes with both large features such as corridors and smaller features such as obstacles would be required.

A more immediate use of additional physical data would be to include the spike rates of each cell during a session to estimate the distance at each time point from the centre of the place field [19]. Interneurons may be identified by an extremely high spike rate compared to other place cells which would eliminate the need to use the position data to identify them. Indeed, MPF cells may be identified in the case where maximal firing occurs in more than one location. A non MPF place cell would produce a single maximal firing location in the centre of its place field.

6.8 More Realistic Simulated Data

Having analysed the experimental data, it is clear there are some key features missing in the simulated data used for the previous project. These include, a greater variation in place field size, spurious firing in common locations such as food goals, less random paths taken by the rat and multiple place field cells. These features could be added to the simulation. The aim here is not to produce simulated data which looks like recorded data but should be an estimate of what perfectly recorded data would be like. This means that the number of place cells should not be reduced to match the number available in these experiments and large amounts of spurious firing should not be included to match noisy recording equipment. A more accurate simulator, used with more creative and complex maze shapes would allow for more robust testing of the algorithms used here and in future work.

7. Conclusion and Project Evaluation

7.1 Conclusion

In a set of recordings of cells in the CA1/CA3 regions of the rat hippocampus, it is possible to identify and exclude interneurons and multiple place field cells using the rat's position as it runs around the track. By using a hidden Markov model to smooth the spiking data from the remaining place cells, it is possible to infer, with reasonable accuracy, the topology of the track as described in the Spatial Reconstruction Experiment. Moreover, using this technique, it is possible to infer the topology of the track as the rat is resting or asleep and the place cells are undergoing replay.

Some limitations to the solution still exist. A way to plausibly remove Multiple Place Field cells, to avoid the errors they introduce, has not been found. The number of place cells recorded in each experiment often fails to provide enough coverage of the maze and errors in recording or misunderstood behaviour can be problematic. Finally, the inference made about topology of the maze is dependent on the path taken by the rat which may not be a full traversal of the entire area.

The solution presented allows for further investigation of arenas and analysis of different shaped mazes and opens up the possibility to explore replay periods in much greater detail.

7.2 Project Evaluation

Initially, before the data analysis had been performed, it was difficult to fully lay out a plan which could be strictly followed. However, the intuition that data smoothing would be required and that certain ground truths would need to be identified and taken into account was correct. It was suggested that these ground truths would be specific to the maze or activity of the rat (for example eating or reaching a goal) and while it is possible that these factors still have an influence, identifying neurological features such as interneurons and multiple place field cells were identified, during the data analysis, as greater factors.

Producing an animation visualiser to better understand the data and its context proved to be very successful and the place field images for each cell were instrumental in analysing results. Though incorrectly defined as ‘filters’ early on, the objective to produce a number of processes to be applied to the data was sound and there is clear evidence of the benefits of removing interneurons and to a lesser extent, multiple place field cells. Smoothing of the data was also identified from the beginning. Bayesian inference was suggested as a tool to eliminate spurious firing but later, HMM became a more appropriate choice. Finally, the objective to find further experimental data was successful although there was not enough time to investigate arenas as suggested. Instead, inference from rest data was explored leading to some interesting results. The contingency plan in case of failure of the application or lack of additional data was, thankfully, avoided. However, the number of deliverables listed initially were over promised. Time was not allotted to partial code submissions and only a single code submission was provided at the end of the project. A separate data analysis report was submitted which gave more structure to the implementation stage and facilitated design conversations. As part of the requirements of the project, a status report was completed with an evaluation of the majority of the work completed by that point. In a commercial setting, partial milestones, particularly code drops, would be extremely important to raising confidence from stakeholders of the project. However, they were perhaps not appropriate here given the tight deadline and sufficient, regular contact time with the supervisor.

Bibliography

[1] Eagleman, D. *The Brain: The story of you.* Pantheon, 2015.

[2] McManus S, Bebbington P, J.R. and T., B. Mental health and wellbeing in england: Adult psychiatric morbidity survey 2014. *Leeds: NHS Digital.*, 2016.

[3] M, P., Knapp M, G.M., McCrone P, P.M., Comas-Herrera A, W.R., Adelaja B, H.B., King D, R.A., and D., S. Update to the dementia uk: report to the alzheimers society. *Kings College London and London School of Economics*, 2014.

[4] Jarrard, L.E. On the role of the hippocampus in learning and memory in the rat. *Behavioral and neural biology*, 1993. **60**(1), pp. 9–26.

[5] Squire, L.R. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, 1992. **99**(2), p. 195.

[6] Eriksson, P.S., Perfilieva, E., Björk-Eriksson, T., Alborn, A.M., Nordborg, C., Peterson, D.A., and Gage, F.H. Neurogenesis in the adult human hippocampus. *Nature medicine*, 1998. **4**(11), pp. 1313–1317.

[7] Scoville, W.B. and Milner, B. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 1957. **20**(1), pp. 11–21.

[8] Jensen, O. and Lisman, J.E. Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in neurosciences*, 2005. **28**(2), pp. 67–72.

[9] Mann, D., Yates, P., and Marcyniuk, B. Some morphometric observations on the cerebral cortex and hippocampus in presenile alzheimer's disease, senile dementia of alzheimer type and down's syndrome in middle age. *Journal of the neurological sciences*, 1985. **69**(3), pp. 139–159.

[10] Ferrucci, D.A. Introduction to this is watson. *IBM Journal of Research and Development*, 2012. **56**(3.4), pp. 1–1.

[11] Rutkin, A. Anything you can do, 2016.

[12] Crane, T. et al. *The mechanical mind: A philosophical introduction to minds, machines and mental representation.* Routledge, 2015.

[13] McCorduck, P. Machines who think. 2004.

[14] Dean, J. and Ghemawat, S. Mapreduce: simplified data processing on large clusters. *Communications of the ACM*, 2008. **51**(1), pp. 107–113.

[15] Chu, C.T., Kim, S.K., Lin, Y.A., Yu, Y., Bradski, G., Ng, A.Y., and Olukotun, K. Map-reduce for machine learning on multicore. In: *NIPS*. Vancouver, BC, volume 6, 2006 pp. 281–288.

[16] Dabaghian, Y., Cohn, A.G., and Frank, L. Topological coding of the hippocampus. *Computational modeling and simulation of intellect: Current state and future prospectives*, 2011. pp. 293–320.

[17] Prichard, D.G. Discerning topological maps from hippocampal place cells. *Department of Computing, Leeds University*, 2010.

[18] O'Keefe, J. and Dostrovsky, J. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 1971. **34**(1), pp. 171–175.

[19] O'Keefe, J. Place units in the hippocampus of the freely moving rat. *Experimental neurology*, 1976. **51**(1), pp. 78–109.

[20] O'keefe, J. and Conway, D. Hippocampal place units in the freely moving rat: why they fire where they fire. *Experimental brain research*, 1978. **31**(4), pp. 573–590.

[21] Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C., and Wilson, M.A. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *Journal of Neuroscience*, 1998. **18**(18), pp. 7411–7425.

[22] Gray, C.M., Maldonado, P.E., Wilson, M., and McNaughton, B. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of neuroscience methods*, 1995. **63**(1), pp. 43–54.

[23] Skinner, B.F. Operant behavior. *American Psychologist*, 1963. **18**(8), p. 503.

[24] Morris, R. Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of neuroscience methods*, 1984. **11**(1), pp. 47–60.

[25] 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*, 2003. **425**(6954), pp. 184–188.

[26] Miller, J.F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S.M., Jacobs, J., et al. Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, 2013. **342**(6162), pp. 1111–1114.

[27] O'Keefe, J. and Nadel, L. Précis of o'keefe & nadel's the hippocampus as a cognitive map. *Behavioral and Brain Sciences*, 1979. **2**(04), pp. 487–494.

[28] O'Keefe, J. and Speakman, A. Single unit activity in the rat hippocampus during a spatial memory task. *Experimental Brain Research*, 1987. **68**(1), pp. 1–27.

[29] Keefe, J.O. and Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature*, 1996. **381**(6581), p. 425.

[30] Redish, A.D., Rosenzweig, E.S., Bohanick, J., McNaughton, B., and Barnes, C.A. Dynamics of hippocampal ensemble activity realignment: time versus space. *Journal of Neuroscience*, 2000. **20**(24), pp. 9298–9309.

[31] Gothard, K.M., Skaggs, W.E., and McNaughton, B.L. Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *Journal of Neuroscience*, 1996. **16**(24), pp. 8027–8040.

[32] Jung, M.W., Wiener, S.I., and McNaughton, B.L. Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *Journal of Neuroscience*, 1994. **14**(12), pp. 7347–7356.

[33] Muller, R.U. and Kubie, J.L. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, 1987. **7**(7), pp. 1951–1968.

[34] Markus, E.J., Qin, Y.L., Leonard, B., Skaggs, W.E., McNaughton, B.L., and Barnes, C.A. Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *Journal of Neuroscience*, 1995. **15**(11), pp. 7079–7094.

[35] Moita, M.A., Rosis, S., Zhou, Y., LeDoux, J.E., and Blair, H.T. Putting fear in its place: remapping of hippocampal place cells during fear conditioning. *Journal of Neuroscience*, 2004. **24**(31), pp. 7015–7023.

[36] Fenton, A.A., Csizmadia, G., and Muller, R.U. Conjoint control of hippocampal place cell firing by two visual stimuli. *The Journal of general physiology*, 2000. **116**(2), pp. 191–210.

[37] Muller, R.U., Poucet, B., and Rivard, B. Sensory determinants of hippocampal place cell firing fields. In: *The neural basis of navigation*, Springer, pp. 1–22. 2002.

[38] Chiba, A.A., Kesner, R.P., and Reynolds, A.M. Memory for spatial location as a function of temporal lag in rats: role of hippocampus and medial prefrontal cortex. *Behavioral and neural biology*, 1994. **61**(2), pp. 123–131.

[39] Dusek, J.A. and Eichenbaum, H. The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences*, 1997. **94**(13), pp. 7109–7114.

[40] Fortin, N.J., Agster, K.L., and Eichenbaum, H.B. Critical role of the hippocampus in memory for sequences of events. *Nature neuroscience*, 2002. **5**(5), pp. 458–462.

[41] Randell, D.A., Cui, Z., and Cohn, A.G. A spatial logic based on regions and connection. *KR*, 1992. **92**, pp. 165–176.

[42] Allen, J.F. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 1983. **26**(11), pp. 832–843.

[43] Hatcher, A. Algebraic topology. 2000.

[44] 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. *Journal of Neuroscience*, 2008. **28**(44), pp. 11250–11262.

[45] Skaggs, W.E. and McNaughton, B.L. Spatial firing properties of hippocampal ca1 populations in an environment containing two visually identical regions. *Journal of Neuroscience*, 1998. **18**(20), pp. 8455–8466.

[46] Spiers, H.J., Hayman, R.M., Jovalekic, A., Marozzi, E., and Jeffery, K.J. Place field repetition and purely local remapping in a multicompartment environment. *Cerebral Cortex*, 2015. **25**(1), pp. 10–25.

[47] Frank, L.M., Brown, E.N., and Wilson, M. Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron*, 2000. **27**(1), pp. 169–178.

[48] Wood, E.R., Dudchenko, P.A., Robitsek, R.J., and Eichenbaum, H. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*, 2000. **27**(3), pp. 623–633.

[49] Bostock, E., Muller, R.U., and Kubie, J.L. Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 1991. **1**(2), pp. 193–205.

[50] Gothard, K.M., Skaggs, W.E., Moore, K.M., and McNaughton, B.L. Binding of hippocampal ca1 neural activity to multiple reference frames in a landmark-based navigation task. *Journal of Neuroscience*, 1996. **16**(2), pp. 823–835.

[51] Foster, D.J. and Wilson, M.A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 2006. **440**(7084), pp. 680–683.

[52] Monaco, J.D., Rao, G., Roth, E.D., and Knierim, J.J. Attentive scanning behavior drives one-trial potentiation of hippocampal place fields. *Nature neuroscience*, 2014. **17**(5), pp. 725–731.

[53] Hill, A. First occurrence of hippocampal spatial firing in a new environment. *Experimental neurology*, 1978. **62**(2), pp. 282–297.

[54] Frank, L.M., Stanley, G.B., and Brown, E.N. Hippocampal plasticity across multiple days of exposure to novel environments. *Journal of Neuroscience*, 2004. **24**(35), pp. 7681–7689.

[55] 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. *J Neurosci*, 1989. **9**(8), pp. 2907–2918.

[56] Ranck, J.B. Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats: Part i. behavioral correlates and firing repertoires. *Experimental neurology*, 1973. **41**(2), pp. 462–531.

[57] Wilson, M.A., McNaughton, B.L., et al. Reactivation of hippocampal ensemble memories during sleep. *Science*, 1994. **265**(5172), pp. 676–679.

[58] Skaggs, W.E. and McNaughton, B.L. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 1996. **271**(5257), p. 1870.

[59] Hartley, T., Lever, C., Burgess, N., and O'Keefe, J. Space in the brain: how the hippocampal formation supports spatial cognition. *Phil. Trans. R. Soc. B*, 2014. **369**(1635), p. 20120510.

[60] 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. *Journal of Neuroscience*, 1990. **10**(2), pp. 420–435.

[61] 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*, 2006. **312**(5774), pp. 758–762.

[62] Calton, J.L., Stackman, R.W., Goodridge, J.P., Archey, W.B., Dudchenko, P.A., and Taube, J.S. Hippocampal place cell instability after lesions of the head direction cell network. *Journal of Neuroscience*, 2003. **23**(30), pp. 9719–9731.

[63] Hafting, T., Fyhn, M., Molden, S., Moser, M.B., and Moser, E.I. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 2005. **436**(7052), pp. 801–806.

[64] Moser, E.I., Kropff, E., and Moser, M.B. Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, 2008. **31**, pp. 69–89.

[65] Solstad, T., Boccara, C.N., Kropff, E., Moser, M.B., and Moser, E.I. Representation of geometric borders in the entorhinal cortex. *Science*, 2008. **322**(5909), pp. 1865–1868.

[66] Boccara, C.N., Sargolini, F., Thoresen, V.H., Solstad, T., Witter, M.P., Moser, E.I., and Moser, M.B. Grid cells in pre-and parasubiculum. *Nature neuroscience*, 2010. **13**(8), pp. 987–994.

[67] Jeffery, K. and Burgess12, N. The boundary vector cell model of place cell firing and spatial memory. *Reviews in the Neurosciences*, 2006. **17**, pp. 71–97.

[68] Samsonovich, A. and McNaughton, B.L. Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, 1997. **17**(15), pp. 5900–5920.

[69] Barry, C. and Burgess, N. Neural mechanisms of self-location. *Current Biology*, 2014. **24**(8), pp. R330–R339.

[70] Meshulam, L., Gauthier, J.L., Brody, C.D., Tank, D.W., and Bialek, W. Collective behavior of place and non-place neurons in the hippocampal network. *arXiv preprint arXiv:1612.08935*, 2016.

[71] Singer, A.C., Carr, M.F., Karlsson, M.P., and Frank, L.M. Hippocampal swr activity predicts correct decisions during the initial learning of an alternation task. *Neuron*, 2013. **77**(6), pp. 1163–1173.

[72] Hunter, J.D. Matplotlib, 2017. URL `https://matplotlib.org/`.

[73] Sridhar, M., Cohn, A.G., and Hogg, D.C. From video to rcc8: exploiting a distance based semantics to stabilise the interpretation of mereotopological relations. In: *International Conference on Spatial Information Theory*. Springer, 2011 pp. 110–125.

[74] Karlsson, M.P. and Frank, L.M. Awake replay of remote experiences in the hippocampus. *Nature neuroscience*, 2009. **12**(7), pp. 913–918.

[75] Kilgard, M.J. Opengl utility toolkit (glut), 2017. URL `https://www.opengl.org/resources/libraries/glut/`.

[76] Ikits, M. and Magallon, M. Opengl extension wrangler (glew), 2017. URL `http://glew.sourceforge.net/`.

[77] Creation, G.T. Opengl mathematics (glm), 2005-2017. URL `https://glm.g-truc.net/0.9.8/index.html`.

[78] Partow, A. The c++ bitmap library, 2002. URL `http://partow.net/programming/bitmap/index.html`.

[79] Hugunin, J. Numpy, 2017. URL `http://www.numpy.org/`.

[80] Travis Oliphant, P.P. and Jones, E. Scipy, 2017. URL `https://www.scipy.org/`.

[81] Aric Hagberg, P.S. and Schult, D. Networkx, 2017. URL `https://networkx.github.io/`.

[82] Karlsson, M.P. and Frank, L.M. Network dynamics underlying the formation of sparse, informative representations in the hippocampus. *Journal of Neuroscience*, 2008. **28**(52), pp. 14271–14281.

[83] Mizuseki, K., Sirota, A., Pastalkova, E., and Buzsáki, G. Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop. *Neuron*, 2009. **64**(2), pp. 267–280.

A. External Material

A.1 Experimental Data

As described in section 3.1, the experimental data was collected and processed externally to produce position tracking files, spike timings and various other meta data used by the topology inference application.

A.2 Visualiser and Preprocessing

The experiment visualiser and spike data preprocessing techniques were developed using Microsoft Visual Studio 2017. The graphical elements and window management were facilitated using the OpenGL Utility Toolkit (GLUT) [75] and OpenGL Extension Wrangler (GLEW) [76] libraries. Certain mathematical structures and functions are provided by OpenGL Mathematics (GLM) [77].

File bitmap_image.hpp is provided by Arash Partow [78] to generate bitmap files.

A.3 Data Analysis and Topology Inference

The python scripts produced to analyse the data and generate topology graphs use the traditional scientific libraries NumPy [79] for mathematical features, SciPy [80] for statistical features and MatPlotLib [72] for graph plotting. In addition, the NetworkX library [81] was used to define network graph structures and plot them. It also provided functions for network analysis.

B. Ethical Issues Addressed

B.1 Frank Data

The details of methodology and data collection for the experiments performed at the lab of Loren Frank at the University of California are outlined in [82] and were obtained according to UCSF and National Institutes of Health guidelines.

B.2 Buzsáki Data

The experimental data obtained from the Gyorgy Buzsáki lab at New York University was used in [83]. The supplementary material outlines the protocols used for the experiments, approved by the Institutional Animal Care and Use Committee of Rutgers University.

B.3 Other Concerns

There is a great need for research and new computational tools into the function of the hippocampus, particularly to aid in finding treatments for Alzheimer's and other causes of Dementia.

Although extremely rudimentary, this research examines the degree to which we can make inferences from neural activity. It has been shown that this pertains not only to spatial location in the past and present but also goal planning and future intent. At some point, it may be possible to apply this and subsequent research to human brain activity which could cause privacy concerns. Currently, this is a distant prospect so this implication can be disregarded in favour of the benefits.