This is a repository copy of *Robots that imagine – can hippocampal replay be utilized for robotic mnemonics?*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/155149/

Version: Accepted Version

Proceedings Paper:

Whelan, M.T., Vasilaki, E. orcid.org/0000-0003-3705-7070 and Prescott, T.J. orcid.org/0000-0003-4927-5390 (2019) Robots that imagine – can hippocampal replay be utilized for robotic mnemonics? In: Martinez-Hernandez, U., Vouloutsi, V., Mura, A., Mangan, M., Asada, M., Prescott, T.J. and Verschure, P.F.M.J., (eds.) Biomimetic and Biohybrid Systems. 8th International Conference, Living Machines, 09-12 Jul 2019, Nara, Japan. Lecture Notes in Computer Science . Springer International Publishing , pp. 277-286. ISBN 9783030247409

https://doi.org/10.1007/978-3-030-24741-6_24

This is a post-peer-review, pre-copyedit version of an article published in Biomimetic and Biohybrid Systems. The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-24741-6_24

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Robots That Imagine – Can Hippocampal Replay be Utilized for Robotic Mnemonics?

Matthew T. Whelan[1,2], Eleni Vasilaki[1], and Tony J. Prescott[1,2]

[1] Department of Computer Science, University of Sheffield, Sheffield, UK, S10 2TN
[2] Sheffield Robotics, University of Sheffield, Sheffield, UK, S1 3JD
`{mwhelan3,e.vasilaki,t.j.prescott}@sheffield.ac.uk`

Abstract. Neurophysiological studies on hippocampal replay, which was a phenomenon first shown in rodents as the reactivation of previously active hippocampal cells, has shown it to be potentially important for mnemonic functions such as memory consolidation/recall, learning and planning. Since its discovery, a small number of neuronal models have been developed to attempt to describe the workings of this phenomenon. But it may be possible to utilize hippocampal replay to help solve some of the difficult challenges that face robotic cognition, learning and memory, and/or be used for the development of biomimetic robotics. Here we review these models in the hope of learning their workings, and see that their neural network structures may be integrated into current neural network based algorithms for robotic spatial memory, and perhaps are particularly suited for reinforcement learning paradigms.

Keywords: Hippocampal replay · Memory · Planning · Reinforcement learning.

1 Introduction

Hippocampal replay, first discovered in rodents, is a phenomenon in which hippocampal cells that were previously active during awake exploratory behaviours are later replayed in the same temporal order, often during sharp-wave ripple events that compresses the time-scale of the replay event relative to the exploratory activation [?]. Hippocampal replay predominantly occurs whilst the rat is in restful states such as sleep or awake quiescence, and can occur in a 'forward' direction, such that the previously active cells are reinstated with the same temporal order [?]; or it can be in the 'reverse', such that the temporal order of the cells are reversed during a replay event [?] (see [?] for a review on hippocampal replay). But just how or why hippocampal replay occurs is still an ongoing research problem, yet current evidence suggests that it may be important for memory consolidation/recall, planning [?], and (reinforcement) learning [?].

The neural mechanisms of reinforcement learning can be traced back to Schultz's seminal work on dopamine as a reward-predicting error signal [?], and a recent review on the ventral basal ganglia (VBG) – a region heavily innervated with dopaminergic neurons [?] – has shown that the hippocampal region projects to and possibly receives projections from the VBG [?]. Indeed, experimentally there is strong evidence that interactions between hippocampus, VBG, and ventral tegmental area support reward-guided memory and conditioned place preference (CPP) [?,?,?]. Furthermore, recent experimental results have shown that hippocampal replays and sharp-wave ripples coordinate with bursts of activity in the ventral tegmental area [?] and ventral striatum [?], and that changes in reward modulates the rate at which hippocampal reverse replays, but not forward replays, occur [?]. It has even been shown in a recent study on humans that spatial memory is prioritized for rewarding locations "retroactively", suggesting that reward-prioritized spatial memory appears some time after an event has occurred [?]. Perhaps it is hippocampal replay in the interim that modulates the memory?

Hippocampal replay in coordination with dopaminergic activity therefore seems well suited as a potential mechanism for reinforcement learning. A number of models have looked to incorporate dopamine as a neuromodulatory third factor in three-factor learning rules for synaptic plasticity (see [?] for a review), successfully showing, for instance, behavioural changes for conditioned place preference in a simulated Morris water maze task [?]. Traditionally, reinforcement learning algorithms have only partially resembled biology, and there is certainly no mention of hippocampal replay in the main body of reinforcement learning literature (i.e. Sutton and Barto's famous text book [?]). However, some of the reinforcement learning algorithms, such as DynaQ algorithms and the deep Q-network, seem well suited as explanations for the use case of hippocampal replay with their need for offline *sequence replays* [?,?,?,?,?,?].

Memory, reinforcement learning and planning are all active challenges in the field of robotics and AI, and bioinspired models have been and are being developed to tackle these challenges [?]. Given the role hippocampal replay has in mnemonic functions, we review here a selection of the most recent studies that aim to describe the neural dynamics of hippocampal replay through computational modelling. We hope that doing so will help determine how and whether hippocampal replay could be useful for solving some of the mnemonic and learning challenges that face the robotics and AI fields.

2 Computational Models of Hippocampal Replay

Models of hippocampal replay are almost exclusively composed of neural networks with either rate-based or spiking-based neural dynamics, and most, if not all, necessitate the use of recurrent networks in order to store memory traces for later reinstatement. Furthermore, they mostly simplify the problem of place cell activation by assuming evenly distributed place fields, usually overlapping, in an environment for which specific place cells fire as a function of the agent's distance from the centre of the respective place field.

We start with a spiking-based model of leaky-integrate and fire neurons by Jahnke et al. [?]. Here they exploit theta phase precession [?] to generate memory

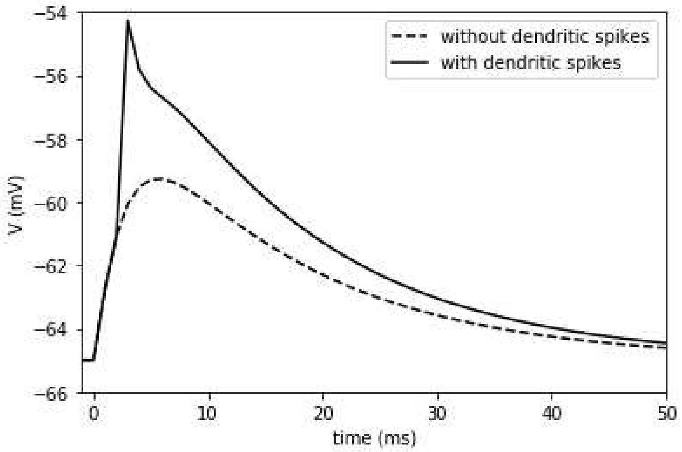


Fig. 1. Dendritic spiking causes supralinear responses to synchronous inputs (solid line) above what would be expected with a simple summation of inputs only (dashed line), important in the model by Jahnke et al. [**?**] for modelling forward/reverse replays and sharp-wave ripples. This plot was modelled using a standard leaky-integrate and fire neuron receiving instantaneous synchronous inputs at t=0, with and without dendritic spiking (see main text).

traces via spike-timing dependent plasticity. But the key inclusion in their model is to use dendritic spiking, which occurs when a high number of synchronous inputs exceed some threshold Θ_b within a time interval of ΔT^s. This then causes a dendritic current impulse which causes an increase in membrane voltage above what would be expected without dendritic spiking (Figure 1).

Once a dendritic spike is initiated, the dendrite enters a refractory period during which time it cannot transmit any spikes. In a linear sequence of place cells with bidirectional connections, this refractory period is important for restricting replays to only travel in a single direction, without reversing back on itself [**?**]. Furthermore, the supralinear nature of the dendritic impulse generates activity pulses that are reminiscent of sharp-wave ripples.

Dendritic spiking, found to occur in CA1 pyramidal cells of the hippocampus [**?**], offers a unique explanation for the occurrence of both sharp wave ripples and replay, and Matheus Gauy et al. [**?**] have extended the use of dendritic spikes, as modelled by Jahnke et al. [**?**], but invented a new cell type termed ‘sequence cells’. The reason for this inclusion is that Jahnke et al.’s model could not account for different trajectories containing the same place cells. Having multiple trajectories emanating from the same place cell would cause replays of multiple trajectories at once. Rather, sequence cells, activated in sequential order as an agent traversed an environment, were paired with place cells via Hebbian learning. As such, one needs only save individual trajectories of sequence cells, and reactivate them in order to reinstate the place cell sequences learned during exploration. It is worth noting that the assumption of sequence cells causes two possible issues: 1) this may necessitate an indefinite number of distinct sequences of sequence cells to

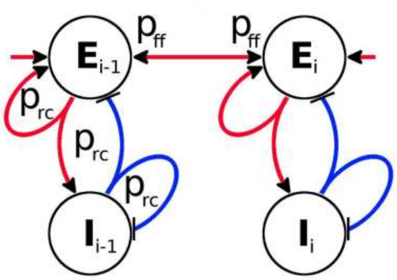


Fig. 2. A model by Chenkov et al. [?] of a synfire-like chain of cell assemblies containing excitatory (**E**) and inhibitory (**I**) cell populations, recurrently connected to each other with probability P_{rc}, and connected feedforwardly with probability P_{ff}. Replay events are characterized as activity propagation from one cell assembly to another and so on, with activity modulated by the inhibitory interneurons. Figure adapted from [?].

account for all trajectories experienced; 2) there is no biological evidence for the existence of sequence cells (though for bioinspiration this may be irrelevant).

Matheus Gauy et al.'s model above had sequences of sequence cells arranged in a similar fashion to synfire chains, and Chenkov et al. [?] designed a similar synfire-like chain of cell assemblies. Within each cell assembly of the synfire chain was a collection of excitatory and inhibitory cells with recurrent connections (Figure 2). By carefully designing each assembly such that there was increased inhibition for accumulating excitatory spikes, the model was able to successfully amplify activity down through the assembly sequences, mimicking replay events, but avoids explosions of activity reminiscent of synfire chain explosions and bursting. This controlled amplification allows weak memory traces, such as those that might be generated during one-shot learning episodes, to successfully re-fire. Furthermore, the increase in inhibition due to accumulating excitatory activity causes replay events to travel in a single direction only.

Refractory periods and inhibitory effects with symmetric bidirectional connections are two methods that allow reverse replays to occur, but Haga and Fukai [?] have shown that the effects of short-term plasticity could also be an explanation for reverse replay. By modelling short-term depression and facilitation at synapses, it is possible to long-term potentiate bidirectional connections in an asymmetric fashion, such that the reverse direction is potentiated more than the forward direction following a forward activation of a sequence. However, it is not clear how this model accounts for forward replay, nor how it prevents continuous reversals in the memory trace strength. For instance, reverse replays cause potentiations to favour the forward direction again, which could be useful, but a forward replay would re-potentiate the reverse direction, which is perhaps not quite clearly as useful.

It seems likely that we'd be interested in models that can support both forward and reverse replays, and perhaps the earliest model of a network incorporating both forward and reverse replay was from Molter et al. [?]. Their original model was more typical in that a traversal through a set of place cells would potentiate that trajectory more in the forward than the reverse direction, but still has non-negligible reverse connections necessary for reverse replays. They also, like in Jahnke et al.'s model [?], employed theta phase precession during memory trace formation. The model was also somewhat simpler and computationally cheaper than the above models, as it was rate-based as apposed to spiking (though both [?] and [?] include rate-based and spiking-based models). But the replays themselves in a 2D environment were similar to a wave-like propagation across the entire environment emanating from the position of replay initiation – as such it does not hold an accurate model of traversal for the environment, though it can provide replays of inexperienced paths.

Following memory trace formation it is then necessary to initiate replay events, and all models suggest that an external stimulus be input to the first (last) cell/cell assembly to initiate forward (reverse) replays. Chenkov et al. [?], however, through control of recurrent and feedforward connection probabilities, show that asynchronous-irregular spiking can spontaneously initiate replay events – whether this is of use is unknown, but a recent study with a DynaQ neural network algorithm suggests 'random' hippocampal replays are not only useful, but necessary, for converging Q-values (see section 3 below).

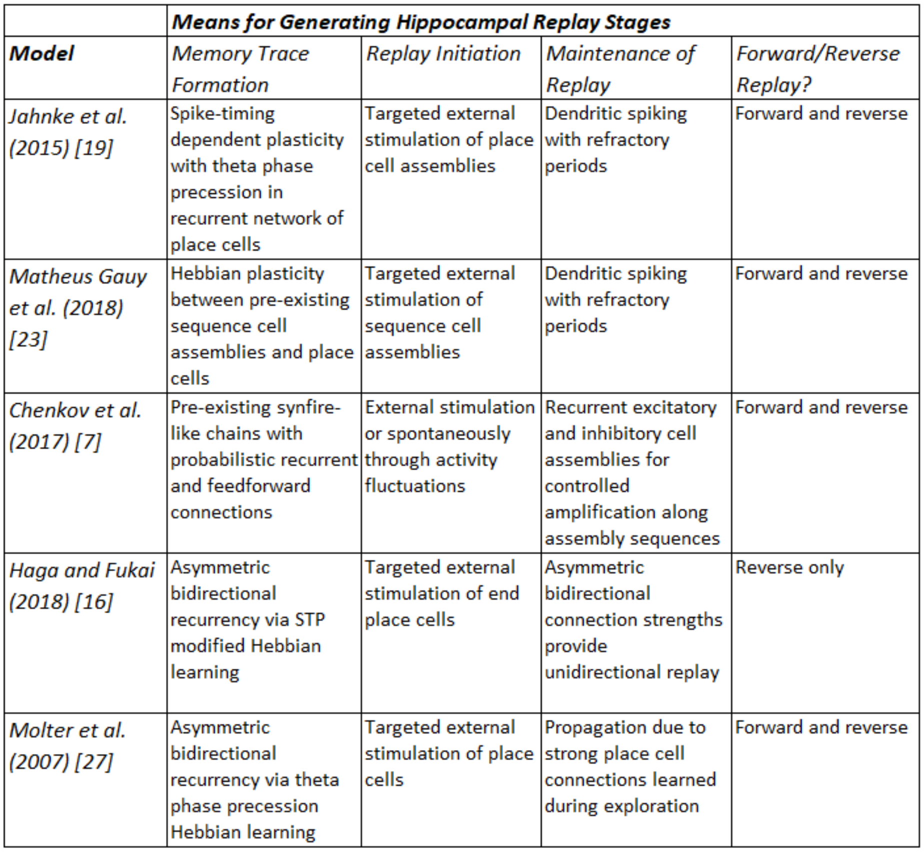
To summarize, there have been a small number of computational models, rate-based and spiking-based, that aim to capture the dynamics of hippocampal replay. Most networks require recurrent connections, either across the whole network or within sub-assemblies that are then connected as synfire-like chains, so that memory traces can be effectively stored and, as a consequence of an external stimulus, reinstated later as a replay event. The mechanisms through which each model forms memory traces and then initiates and maintains replay events is summarized in Table 1.

It is worth noting that a small number of studies have modelled the process of sharp-wave/ripples in the hippocampus, which occurs simultaneously with a replay event [?]. Particularly they model the generation of sharp-wave/ripples via interactions of inhibitory interneurons, extra-hippocampal inputs such as septal inputs, and/or neuromodulators like acetylcholine, and the relationship between sharp-wave/ripples and replay events [?,?,?,?].

3 Hippocampal Replay for Robotics Applications

The models reviewed here are computational models with the sole intent of replicating experimental findings. But they do not prove immediately useful for practical applications, as they all require place cell representations readily available prior to replay, and offer no useful outputs post-replay. What could be missing then is a unified model of place cell, or state, representations at the input end of a hippocampal replay model, and a means for action-selection improvement at the output end. A line a literature on each of these areas is available, but a full review of these is outside the scope of this paper.

Though a few recent studies are worth mentioning here that could integrate well with hippocampal replay. On the place cell representation end, the first



	Means for Generating Hippocampal Replay Stages			
Model	*Memory Trace Formation*	*Replay Initiation*	*Maintenance of Replay*	*Forward/Reverse Replay?*
Jahnke et al. (2015) [19]	Spike-timing dependent plasticity with theta phase precession in recurrent network of place cells	Targeted external stimulation of place cell assemblies	Dendritic spiking with refractory periods	Forward and reverse
Matheus Gauy et al. (2018) [23]	Hebbian plasticity between pre-existing sequence cell assemblies and place cells	Targeted external stimulation of sequence cell assemblies	Dendritic spiking with refractory periods	Forward and reverse
Chenkov et al. (2017) [7]	Pre-existing synfire-like chains with probabilistic recurrent and feedforward connections	External stimulation or spontaneously through activity fluctuations	Recurrent excitatory and inhibitory cell assemblies for controlled amplification along assembly sequences	Forward and reverse
Haga and Fukai (2018) [16]	Asymmetric bidirectional recurrency via STP modified Hebbian learning	Targeted external stimulation of end place cells	Asymmetric bidirectional connection strengths provide unidirectional replay	Reverse only
Molter et al. (2007) [27]	Asymmetric bidirectional recurrency via theta phase precession Hebbian learning	Targeted external stimulation of place cells	Propagation due to strong place cell connections learned during exploration	Forward and reverse

Table 1. Summary of the hippocampal replay models. The means by which each model performs the stages required for hippocampal replay are summarized here. See main text for full details.

is a biologically inspired SLAM algorithm, or RatSLAM, developed by Milford et al. [?], which has proven effective at capturing state representations in the form of 'pose cells.' With an accurate map represented in the form cell values, this offers itself as a candidate for replay models based upon neural networks. Alternatively, Byrne et al. model [?] hippocampal place cells, boundary vector cells and head direction cells, all neurophysiological features of the hippocampal region [?,?], which could provide a more biologically plausible model of place cell representations, whilst Jauffret et al. [?] have recently developed a model of grid cells [?] and place cells that was successfully applied for spatial navigation in a robot.

For action-selection improvement, the first is a DynaQ neural network algorithm developed by Aubin et al. [?], which is a reinforcement learning model using Q-learning and the Dyna algorithm. It too is composed of a neural network that represents states, but pairs the states with (discrete) actions. They indeed integrated a version of hippocampal replay and showed that where Q-values could not converge online due to similarities in state values, they could however converge offline via 'random' hippocampal replays. Mnih et al.'s deep Q-network (DQN) [**?**], in a similar fashion to the DynaQ neural network, utilized *experience replays*, which is conducted by selecting from a random uniform distribution groupings of state, action, reward and next state experiences. A list of experiences could then be denoted by $D_t = \{e_1, ..., e_t\}$ with $e_t = (s_t, a_t, r_t, s'_t)$ being an individual experience, and applying the Q-learning update for each $e_{\text{rand}} \sim U(D)$ where $U(\cdot)$ is the uniform distribution. But both these algorithms suffer from perhaps one minor issue, in that Q-values here are learned for a discrete set of actions. Though perhaps rectifiable, this could prove problematic for states that are represented continuously.

Recent work by Mattar and Daw [**?**] developed a Q-learning based reinforcement learning model that prioritizes Bellman backups. Such a prioritization (for which something similar is found in the model by Aubin et al. [**?**] and termed *prioritized sweeping*) determines whether the agent should prioritize the evaluation of upcoming decisions, or whether to perform updates in order to capture newly learned information about a reward. Prioritization of the former increases the number of forward replays, whilst for the latter reverse replays become more prominent. In this way, the model favours forward replays at the start of a trial, whilst reverse replays are favoured at the end of a trial – effects similar to that found with hippocampal replay [**?**].

Another challenge for robotics is the number of trials required for reinforcement learning algorithms to properly converge. This was a problem addressed by Vasilaki et al. [**?**,**?**] in a spike-based model of hippocampal place cells for a reinforcement learning Morris water maze task. They showed that whereas policy-gradient methods require either a high number of learning trials (due to small learning rates) or cause noisy eligibility traces (when learning rates are high), their model could account for effective learning within a small number of trials, as is found experimentally with rats. Interestingly, they modelled "action cells", which could possibly be found in the basal ganglia as an action selection mechanism [**?**], and further, unlike the models discussed above, they were able to represent actions and states as continuous, rather than discrete. Yet importantly for our discussion here, they did not employ hippocampal replay.

Hippocampal replay could offer another means to achieve the low number of learning trials required – learning is done "offline" as (perhaps noisy) repetitions of previous experience. This could therefore offer an effective and highly efficient mechanism that converges state-action values "offline", which could prove useful for robotic learning, as well as offer bioinspired learning mechanisms for biomimetic robotics. And indeed there can be a symbiotic relationship between neurophysiology and robotics, such that whilst robotics can take inspiration from biology, we may also enhance our understanding of the underlying biology by providing solutions, via behavioural robotics, of some of the functional properties that exists in biological creatures.

4 Conclusion

The discovery of hippocampal replay, via its ability to reinstate previously active place cells, along with dopaminergic activity in the basal ganglia, offers a possibly efficient “offline” mechanism that seems to be in place for post-hoc pairing of state-actions with rewards. We have therefore reviewed a number of biologically-plausible models that aim to describe the neural mechanisms of hippocampal replay in both the forward and reverse directions. Integrating biologically plausible state representations, such as place cells, with a biologically plausible action selection mechanism remains an open challenge, but incorporating hippocampal replay into the process could help understand better the underlying biology, as well perhaps aid in the development of memory and learning algorithms. Whether it can be used for robotic mnemonics and biomimetic robotics remains open for further investigation, but the groundwork seems already to be in place.

References

1. Ambrose, R.E., Pfeiffer, B.E., Foster, D.J.: Reverse replay of hippocampal place cells is uniquely modulated by changing reward. Neuron **91**(5), 1124–1136 (2016)
2. Ariav, G., Polsky, A., Schiller, J.: Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of ca1 pyramidal neurons. Journal of Neuroscience **23**(21), 7750–7758 (2003)
3. Aubin, L., Khamassi, M., Girard, B.: Prioritized sweeping neural dynaq with multiple predecessors, and hippocampal replays. In: Conference on Biomimetic and Biohybrid Systems. pp. 16–27. Springer (2018)
4. Braun, E.K., Wimmer, G.E., Shohamy, D.: Retroactive and graded prioritization of memory by reward. Nature communications **9**(1), 4886 (2018)
5. Byrne, P., Becker, S., Burgess, N.: Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychological review **114**(2), 340 (2007)
6. Cazé, R., Khamassi, M., Aubin, L., Girard, B.: Hippocampal replays under the scrutiny of reinforcement learning models. Journal of neurophysiology **120**(6), 2877–2896 (2018)
7. Chenkov, N., Sprekeler, H., Kempter, R.: Memory replay in balanced recurrent networks. PLoS computational biology **13**(1), e1005359 (2017)
8. Cutsuridis, V., Hasselmo, M.: Spatial memory sequence encoding and replay during modeled theta and ripple oscillations. Cognitive Computation **3**(4), 554–574 (2011)
9. Cutsuridis, V., Taxidis, J.: Deciphering the role of ca1 inhibitory circuits in sharp wave-ripple complexes. Frontiers in systems neuroscience **7**, 13 (2013)
10. Diba, K., Buzsáki, G.: Forward and reverse hippocampal place-cell sequences during ripples. Nature neuroscience **10**(10), 1241 (2007)
11. Foster, D.J.: Replay comes of age. Annual review of neuroscience **40**, 581–602 (2017)
12. Foster, D.J., Wilson, M.A.: Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature **440**(7084), 680 (2006)
13. Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., Brea, J.: Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. Frontiers in neural circuits **12** (2018)

14. Gomperts, S.N., Kloosterman, F., Wilson, M.A.: Vta neurons coordinate with the hippocampal reactivation of spatial experience. Elife **4**, e05360 (2015)
15. Hafting, T., Fyhn, M., Molden, S., Moser, M.B., Moser, E.I.: Microstructure of a spatial map in the entorhinal cortex. Nature **436**(7052), 801 (2005)
16. Haga, T., Fukai, T.: Recurrent network model for learning goal-directed sequences through reverse replay. Elife **7**, e34171 (2018)
17. Humphries, M.D., Prescott, T.J.: The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Progress in neurobiology **90**(4), 385–417 (2010)
18. Ikemoto, S., Yang, C., Tan, A.: Basal ganglia circuit loops, dopamine and motivation: a review and enquiry. Behavioural brain research **290**, 17–31 (2015)
19. Jahnke, S., Timme, M., Memmesheimer, R.M.: A unified dynamic model for learning, replay, and sharp-wave/ripples. Journal of Neuroscience **35**(49), 16236–16258 (2015)
20. Jauffret, A., Cuperlier, N., Gaussier, P.: From grid cells and visual place cells to multimodal place cell: a new robotic architecture. Frontiers in neurorobotics **9**, 1 (2015)
21. Johnson, A., Redish, A.D.: Hippocampal replay contributes to within session learning in a temporal difference reinforcement learning model. Neural Networks **18**(9), 1163–1171 (2005)
22. Martinez-Hernandez, U., Damianou, A., Camilleri, D., Boorman, L.W., Lawrence, N., Prescott, T.J.: An integrated probabilistic framework for robot perception, learning and memory. In: 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO). pp. 1796–1801. IEEE (2016)
23. Matheus Gauy, M., Lengler, J., Einarsson, H., Meier, F., Weissenberger, F., Yanik, M.F., Steger, A.: A hippocampal model for behavioral time acquisition and fast bidirectional replay of spatio-temporal memory sequences. Frontiers in neuroscience **12**, 961 (2018)
24. Mattar, M.G., Daw, N.D.: Prioritized memory access explains planning and hippocampal replay. Nature Neuroscience **21**(11), 1609 (2018)
25. Milford, M.J., Wyeth, G.F., Prasser, D.: Ratslam: a hippocampal model for simultaneous localization and mapping. In: IEEE International Conference on Robotics and Automation, 2004. Proceedings. vol. 1, pp. 403–408. IEEE (2004)
26. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529 (2015)
27. Molter, C., Sato, N., Yamaguchi, Y.: Reactivation of behavioral activity during sharp waves: a computational model for two stage hippocampal dynamics. Hippocampus **17**(3), 201–209 (2007)
28. O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. Brain research (1971)
29. O'Keefe, J., Recce, M.L.: Phase relationship between hippocampal place units and the eeg theta rhythm. Hippocampus **3**(3), 317–330 (1993)
30. Ólafsdóttir, H.F., Bush, D., Barry, C.: The role of hippocampal replay in memory and planning. Current Biology **28**(1), R37–R50 (2018)
31. Pennartz, C., Lee, E., Verheul, J., Lipa, P., Barnes, C.A., McNaughton, B.: The ventral striatum in off-line processing: ensemble reactivation during sleep and modulation by hippocampal ripples. Journal of Neuroscience **24**(29), 6446–6456 (2004)
32. Redgrave, P., Prescott, T.J., Gurney, K.: The basal ganglia: a vertebrate solution to the selection problem? Neuroscience **89**(4), 1009–1023 (1999)

33. Richmond, P., Buesing, L., Giugliano, M., Vasilaki, E.: Democratic population decisions result in robust policy-gradient learning: a parametric study with gpu simulations. PLoS one **6**(5), e18539 (2011)
34. Saravanan, V., Arabali, D., Jochems, A., Cui, A.X., Gootjes-Dreesbach, L., Cutsuridis, V., Yoshida, M.: Transition between encoding and consolidation/replay dynamics via cholinergic modulation of can current: a modeling study. Hippocampus **25**(9), 1052–1070 (2015)
35. Schultz, W.: Predictive reward signal of dopamine neurons. Journal of neurophysiology **80**(1), 1–27 (1998)
36. Skaggs, W.E., McNaughton, B.L.: Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. Science **271**(5257), 1870–1873 (1996)
37. Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. MIT press (2018)
38. Taube, J.S.: Head direction cells and the neurophysiological basis for a sense of direction. Progress in neurobiology **55**(3), 225–256 (1998)
39. Taxidis, J., Coombes, S., Mason, R., Owen, M.R.: Modeling sharp wave-ripple complexes through a ca3-ca1 network model with chemical synapses. Hippocampus **22**(5), 995–1017 (2012)
40. Trouche, S., Koren, V., Doig, N.M., Ellender, T.J., El-Gaby, M., Lopes-dos Santos, V., Reeve, H.M., Perestenko, P.V., Garas, F.N., Magill, P.J., et al.: A hippocampus-accumbens tripartite neuronal motif guides appetitive memory in space. Cell (2019)
41. Vasilaki, E., Frémaux, N., Urbanczik, R., Senn, W., Gerstner, W.: Spike-based reinforcement learning in continuous state and action space: when policy gradient methods fail. PLoS computational biology **5**(12), e1000586 (2009)