To appear in Aronowitz, S. & Nadel, L. (Eds.). (Forthcoming). *Space, Time, and Memory*. Oxford University Press.

Experience Replay Algorithms and the Function of Episodic Memory

Alexandria Boyle[1]

[1] Department of Philosophy, Logic and Scientific Method,
London School of Economics and Political Science

Abstract: Episodic memory is memory for past events. It's characteristically associated with an experience of 'mentally replaying' one's experiences in the mind's eye. This biological phenomenon has inspired the development of several 'experience replay' algorithms in AI. In this chapter, I ask whether experience replay algorithms might shed light on a puzzle about episodic memory's function: what does episodic memory contribute to the cognitive systems in which it is found? I argue that experience replay algorithms can serve as idealized models of episodic memory for the purposes of addressing this question. Taking the DQN algorithm as a case study, I suggest that these algorithms provide some support for mnemonic accounts, on which episodic memory's function lies in the storage, encoding and retrieval of information. By extending and adapting experience replay algorithms, we might gain further insight into episodic memory's operations and contributions to cognition.

Keywords: *episodic memory; experience replay; artificial intelligence; cognitive role functions; models*

1. Introduction

Think back to when you woke up this morning. Maybe you snoozed your alarm several times, or perhaps you jumped out of bed straight away, ready to face the day. Maybe you had a leisurely breakfast and chatted to your family, or maybe you had to tumble out of the house in a hurry with nothing but a coffee to go. Whatever happened, as you think back to this morning you may have the feeling that you're 'replaying' your experiences in your mind's eye. You might mentally smell the coffee, see it foam or hear the hiss as it comes out of the machine. This phenomenon–experience replay–is characteristically associated with *episodic memory:* memory for personally experienced past events.

Episodic memory presents several puzzles for philosophers and scientists of memory. Among them is the question of its function: what does episodic memory contribute to the cognitive system (Cummins, 1975)? To put it another way, what is it that we're able to do because we can remember and 'replay' past events? What makes this puzzle challenging to resolve is the difficulty of evaluating competing theories of episodic memory's function. To do so, we'd ideally need an agent which had episodic memory, whose episodic memory capacity we could 'switch off' whilst leaving its other cognitive functions intact, so that we could isolate the unique contributions episodic memory makes. For reasons I'll outline, it is not easy to find an agent like this in the biological realm. But recent developments in AI suggest an intriguing possibility: perhaps we could make one.

In recent years, a number of algorithms have been developed which exploit an 'experience replay' mechanism. These algorithms record details of their 'experiences' and replay these experiences after the fact. The replay mechanism is often separable from the rest of the cognitive architecture, making it possible to directly assess its contributions to the system's capabilities. In this chapter, my question is whether, by investigating the effects of experience replay on artificial agents, we stand to learn about episodic memory's function in biological systems. For this to be a promising research strategy, there would need to be more than a superficial resemblance between biological and artificial experience replay. The two would need to be meaningfully similar, such that one could hope to draw justified (if defeasible) inductive inferences about biological episodic memory from artificial experience replay. Whilst there are significant differences between these phenomena, my argument here will be that the two are sufficiently similar in relevant ways to ground such inferences–so, by investigating these algorithms we stand to gain defeasible evidence about the function of episodic memory.

I begin in §2 by describing DeepMind's DQN algorithm (Mnih et al., 2015), the experience replay algorithm I take as my case study in this chapter.[1] In §3, I argue that in order to inform our accounts of episodic memory's function, DQN would need to resemble episodic memory at the algorithmic level. I show that there are significant similarities between the two at the algorithmic level: both exploit detailed, iconic representations integrating multidimensional information about specific past events. In §4, I apply this result to the debate about episodic memory's function. After briefly summarising the debate between simulationist and mnemonic accounts of episodic memory's function, I argue that DQN provides some support to the mnemonic view. I also acknowledge several differences between DQN and episodic memory. I argue that for some purposes, these differences will not matter: DQN can fruitfully serve as an idealized model of episodic memory whilst differing from it in various ways. In other contexts, however, these differences will matter. This reveals ways in which we might look to extend or adapt experience replay algorithms to resemble episodic memory more closely. Doing so would facilitate the evaluation of more granular theories about episodic memory's operations and contributions to cognition. §5 concludes.

2. The DQN Algorithm

[1] Given the differences between DQN and other implementations of experience replay, the conclusions I draw here can't be generalized to other architectures without argument. I discuss some other experience replay algorithms briefly in §4.

The Arcade Learning Environment is a platform which uses Atari 2600 games, like Breakout, Pong and Space Invaders, to evaluate AI systems (Bellemare et al., 2013). In 2015, DeepMind reported that their DQN algorithm had achieved a new state of the art in the Arcade Learning Environment, achieving better scores than the previous best performing algorithm in 43 of 49 games tested (Mnih et al., 2015). DQN's performance was comparable to that of a professional human games tester, achieving at least 75% of the human's score on more than half of the games sampled. Key to DQN's success was a 'biologically inspired' experience replay mechanism. Unlike previous systems, which learned from each 'experience' and moved on, this algorithm recorded its experiences and replayed them after the fact.

DQN is a reinforcement learning algorithm. Reinforcement learning is a machine learning paradigm in which an agent learns by taking actions in its environment. Reinforcement learning problems are commonly modelled as Markov decision processes. In a Markov decision process, an agent interacts with its environment at several discrete time stems, t = 1, 2, …, n. At each time-step, the agent observes the state of the environment (S_t), selects an action (A_t), receives a reward (R_t) and observes the resulting state of the environment at the next time-step (S_{t+1}) (Sutton & Barto, 2018). In the Arcade Learning Environment, the reward is the change–positive or negative–in game score. Over time, the idea is that the agent will learn a policy which maximizes its future rewards. Roughly speaking, a policy is a state-action mapping; something that determines how the agent will act in any given state. The agent may initially act randomly, but by observing the effects and rewards produced by its actions, it develops more sophisticated policies, enabling it to respond effectively to the environment (Figure 1).

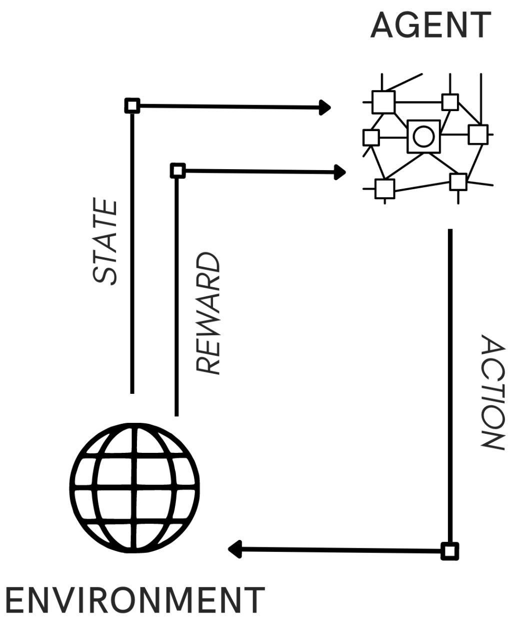

Figure 1: Agent-environment interactions in a Markov decision process.

DQN implements a form of reinforcement learning known as Q-learning. Q-learning involves learning the expected reward (the 'Q-value') of the actions in each scenario. In 'vanilla' Q-learning (i.e., the most basic form), Q-values are represented in a lookup table mapping state/action pairs to Q-values (

Figure 2). When the agent acts, the Q-table is updated to reflect new information about the relevant state/action pairs. DQN is a variation on Q-learning in which the Q-table is replaced by a neural network mapping input states (i.e., observed states of the environment) to action/Q-value pairs.

STATE → ACTION ↓	S1	S2	S3
A1	50	78	4
A2	67	9	33
A3	1	45	70

Figure 2: Q table

At each time-step, the agent selects an action using an 'ε-greedy' policy: it selects a random action with probability ε, and selects the action predicted to have the highest value with probability 1-ε. The value of ε is adjusted during training such that the agent begins by randomly exploring the available actions, later shifting towards exploiting what it has learned. Whenever the agent takes an action, the divergence between the expected and actual reward is used as an error signal to train the Q-learning network.

The important thing about DQN for our purposes is its use of experience replay. In experience replay, the agent's experience at each time-step is recorded. An experience consists of a 4-tuple representation (S_t, A_t, R_t, S_{t+1}), where:

- S_t = the state of the environment at *t*
- A_t = the action taken at *t*
- R_t = the reward obtained at *t*
- S_{t+1} = the state of the environment at *t+1.*

These 4-tuple experiences are pooled in a store called the 'episodic buffer'. The oldest experiences are deleted when the episodic buffer's finite capacity is reached.

Importantly for our purposes, the state representations (S_t) are more complex than this brief sketch suggests. 'S_t' in fact represents a sequence of observations (X_t) and actions (A_t) over the m time-steps leading up to and including t. The number of time-steps per sequence (m) is variable; in DeepMind's original implementation, m=4. S_t is then shorthand for the sequence: X_{t-3}, A_{t-3}, X_{t-2}, A_{t-2}, X_{t-1}, A_{t-1}, X_t. Each X is an observation of the screen of the Atari emulator, represented as pixel vector. Simplifying, this is generated by taking a screenshot from the Atari system, applying some minimal pre-processing to remove artifacts, standardizing the size of the screenshot and converting it to greyscale. Each pixel in the resulting 84x84 pixel greyscale image can now be represented as a single number, the luminance value corresponding to its particular shade of grey. These luminance values are used to convert the image into a two-dimensional numerical array. The two dimensions of the array correspond to image height and width; each number in the array picks out the pixel in the corresponding image location and represents its luminance. The sequence, S_t, is then a temporally ordered sequence of pixel vectors interspersed with the actions taken at the corresponding time-step.

The Q-learning network learns 'off-policy', by training itself on minibatches of experiences selected at random from the episodic buffer, rather than directly on the agent's most recent experience. One advantage of using past experiences for learning in this way is that each experience can be used for learning several times. Another is that it eliminates the correlations which would otherwise hold between consecutive training samples, which are both inefficient for learning and increase the chances of the system getting stuck in local minima–that is, becoming committed to suboptimal strategies. Consistently with experience replay conferring these learning advantages, disabling replay significantly worsens the network's performance, sometimes by an order of magnitude (Mnih et al., 2015, table 3).

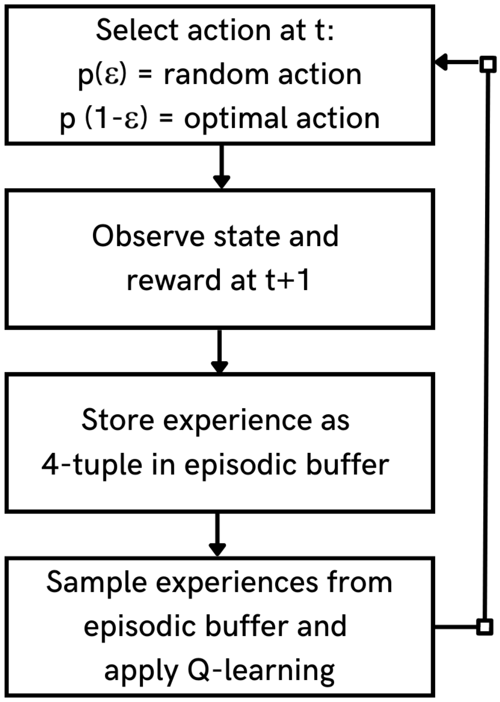

Figure 3: Simplified sketch of the DQN algorithm

3. Representations in Episodic Memory and DQN

As noted in §1, DQN's reliance on experience replay is suggestive because of the characteristic association between experience replay and episodic memory in humans. But to determine whether this is any more than suggestive, we'd need to know whether this is a *meaningful* similarity between episodic memory and DQN.

In general, cognitive systems can be described and explained at various 'levels. David Marr's (1982) prominent account distinguishes the computational, algorithmic and implementation levels. The computational level is the level at which we describe *what* the system is doing and *why*; the algorithmic level describes what representations are used and what algorithms are used to process them; the implementation level describes the physical structures realising this processing. There is some independence between these levels, meaning that two systems might resemble one another at one level whilst being very different at another.

In the current context, we are ultimately interested in understanding the cognitive role function of episodic memory: what it does, and how, by doing that, it contributes to the capacities of the cognitive systems of which it's a part (Cummins, 1975). What makes this puzzling is that, on an intuitive picture of what episodic memory does, it's not obvious that it has a distinctive contribution to make. Its central job is to carry detailed information about specific past events. But we have a general-purpose memory system in the form of semantic memory–a decontextualized memory store carrying information about words, conceptual relations or facts. So, we might wonder, what is the point of having *another* memory system dedicated to the storage, encoding and retrieval of detailed information about specific past events when, after all, much of this information is surely redundant, and the specific events will never come around again?[2] Put in these terms we can see that our question is pitched at the algorithmic level: what is the point of having a memory system that uses *these* kinds of representations in this way? For DQN to aid us in answering this question, we would therefore need to establish that it resembled episodic memory at the algorithmic level.[3] The task of this section is to establish that there are reasons for thinking this is true, by showing that DQN and episodic memory exploit similar kinds of representations.

The first step is to give an account of the representations exploited by biological episodic memory. When Endel Tulving introduced the distinction between episodic and semantic memory (Tulving, 1972), he distinguished the two in terms of their content. Episodic memory was a matter of remembering *what* happened, *where* and *when*. However, Tulving and many others came to see this characterization of episodic memory as inadequate, for two reasons. First, there are examples of episodic memories in which one or more of the 'what-where-when' components is missing. Second, it is not uncommon to semantically remember the 'what-where-when' of events, including events which one could not episodically remember, such as the Battle of Hastings. These seem significant problems: ideally, an account of episodic memory ought at least to capture what *in general* distinguishes it from its most salient contrast class (Boyle, 2020b).

In response to this, Tulving and others came to place greater weight on the characteristic *phenomenology* of episodic recollection: the experience described in terms of 'replaying' past events (Tulving, 2005).[4] To place too much weight on *experience* in characterising episodic memory feels problematic, however. Most significantly for our purposes, this provides little guidance when it comes to evaluating the significance of algorithms like DQN. Since, plausibly, no current AI systems are conscious, no current architecture will exhibit the relevant phenomenology, suggesting that no current experience replay architecture implements episodic memory. To be clear, that may well be the right result. But this does not settle the question of whether architectures like DQN can inform our understanding of biological episodic memory, since they might nevertheless be similar in significant respects. Nevertheless, as I've argued elsewhere (Boyle, 2020b),

[2] See Brown (2023, sec. 3) for a compelling discussion of this puzzle.

[3] This means that we need not be troubled in this context by the obvious fact that DQN and biological episodic memory processes differ at the level of implementation: DQN is implemented in ordinary computer hardware, and episodic memory in brain structures and processes.

[4] Following Tulving (2005), this experience is sometimes described in terms of 'autonoetic consciousness' and 'chronesthesia'. I find this terminology unhelpful for reasons I've outlined elsewhere (Boyle, 2020a), and will avoid it here.

characterising episodic memory in terms of experience replay can be fruitful, since careful reflection on what is involved in 'replaying' past events allows us to characterize the *representations* involved in episodic memory in more granular terms than 'what-where-when'. This reveals an account of episodic memory which both distinguishes it from semantic memory and facilitates comparisons with experience replay algorithms: we can determine whether experience replay algorithms involve representations which carry similar information and are structured in similar ways.

To say that episodic memory characteristically involves experience replay involves taking on some substantial commitments about the content episodic memory carries. First, it suggests that episodic memory carries detailed spatial information about an event. This is not merely to say that it carries information about *where* the event occurred. In fact, one might episodically remember an event without being able to pinpoint its location. But episodically remembering an event typically involves remembering what we might call the event's 'internal' spatial features, that is, the spatial context in which the event occurred. For instance, if you remember having breakfast this morning, you might remember the room you ate in, how the furniture was arranged in that room, where your cereal bowl was relative to the table, where the cereal was relative to the bowl, and so on. That episodic memory captures this kind of contextual spatial information about an event is central to accounts characterising episodic memory in terms of 'scene construction' (Boyle, 2020a; Clayton & Russell, 2009; Rubin & Umanath, 2015).

In addition, to the extent that episodic memory prototypically involves replaying events, it must carry similar information about the event's temporal features. Once again, this is not a matter of remembering *when* the event occurred. In fact, episodic memory need not involve representing events as past at all (Boyle, 2020a). But episodically remembering an event, such that one could in principle mentally replay or re-experience it, must involve remembering its 'internal' temporal features, that is, the order in which its component parts occurred (Boyle, 2020b). So, again, if you remember having breakfast this morning, you might remember that your cereal bowl was full at the beginning and became progressively emptier, that the cereal was crispy at the beginning but became progressively soggier, and so on.

Spatial and temporal information are of course not the only kinds of information episodic memory characteristically carries about an event. Given that episodic memories are of events you personally witnessed or were involved in, we might add that episodic memory can carry self-referential information, such as information about how you were involved, what you were thinking, feeling or perceiving (Boyle, 2020b).

So, we now have a somewhat fuller characterization of the sort of *content* episodic memory prototypically carries, namely, detailed, multidimensional information about an event, including the spatial and temporal organization of the event, and the remembering subject's involvement in, perception of and thoughts and feelings about the event at the time of its occurrence. This is by no means a novel view of episodic memory's content: a number of recent accounts emphasize the fact that episodic memory carries event information across a number of quality dimensions (Brown, 2023; Gershman & Daw, 2017). This picture of episodic memory's content is also reflected in the methodologies used to investigate episodic memory in cognitive science. For instance, the Autobiographical Interview Questionnaire distinguishes the episodic and semantic aspects of an autobiographical memory by coding reported details as either 'internal' (episodic) or 'external' (semantic), where 'internal details' are comments on the event's spatiotemporal structure, and the subject's thoughts, emotions or perceptions during the event (Levine et al., 2002). Several methods for detecting episodic memory in animals also rest on the idea that episodic memory carries detailed multidimensional event information. For instance, the 'what-where-which' protocol investigates whether an animal can discriminate between similar events which occurred in distinct spatial contexts (Eacott et al., 2005), whilst 'source monitoring' studies investigate whether an animal remembers contextual detail about an event besides what happened, where and when (Crystal et al., 2013).

Whilst 'replaying' an event involves recalling detailed multidimensional event information, the various details are not recalled separately from one another. Rather, these various details are presented to us as a package: they seem to be integrated into a single, structured representation of the event as a whole (Boyle, 2021; Rubin & Umanath, 2015). Once again, this idea is reflected in methods used to detect episodic memory. The 'integration' criterion for detecting episodic memory in animals investigates whether the various datapoints an animal remembers about an event are integrated into a single, unified representation (Clayton et al., 2001). One operationalization of this is that memory is integrated when retrieval of one piece of information encoded in the memory predicts retrieval of the rest (Clayton et al., 2003). Another is that a memory is integrated if it is resistant to interference from memories for events composed out of similar informational components–that is, if the subject can remember and distinguish similar but distinct events (Crystal & Smith, 2014). Underlying the idea that episodic memories should be integrated is the idea that episodic memories combine diverse datapoints into a single representational unit.

It is tempting to cash this out as a claim about episodic memory's format: perhaps episodic memory has a (partly) iconic format. Iconic formats are characterized by structural isomorphism. That is to say, the structure of the representation mirrors or 'maps on to' the structure of the thing being represented, and that mapping is semantically significant (Lee et al., 2022; Shea, 2014). Iconic representations are typically informationally rich, integrating detailed, often multidimensional information, into a single representational unit. So, we might think a plausible hypothesis about experience replay is that it involves representations with an iconic format, and this explains why they integrate rich multidimensional information about events in a way that is resistant to interference from similar memories. One thing that renders this intuitively plausible is the way experience replay seems to represent the temporal properties of remembered events. Replayed experiences seem to 'unfold' over time in the mind's eye, in a way that mirrors the represented unfolding of the event (Boyle, 2020b). Here, the temporal features of the representation appear to map onto the temporal features of the represented event.

For these kinds of reasons, Nikola Andonovski (2022) argues that episodic memories are 'structure-preserving models' of past events. On this view, episodic memories and other forms of episodic representation are abstract mental models which mirror the spatiotemporal structure of represented events, and perhaps their structure across other quality dimensions. This mirroring is unlikely to be a strict isomorphism; it may be approximate or simplified. Importantly, this structure-preserving format is unlikely to exhaust the representational features of episodic memory: given the pervasive interactions between episodic and semantic memory (Aronowitz, 2022; Boyle, 2021), instances of episodic remembering will almost always involve conceptual or semantic elements in addition to the structure-preserving model at their core.

On the view of episodic memory we've arrived at, then, it involves retrieving representations which carry detailed multidimensional information about a remembered event. Minimally, this includes its spatial and temporal organization, perhaps along with information about the event's other qualities, the subject's background knowledge, and the subject's involvement in, perception of and thoughts and feelings about the event. These details seem to be retrieved as a package, indicating that the representations involved unite these details into an integrated representational whole. This suggests the representations involved may be at least partly iconic, perhaps taking the form of structure-preserving models in which there are semantically significant approximate isomorphisms between structural properties of the representation and those of the represented event.

We are now in a position to see that experience replay in DQN involves representing past events in a way that is meaningfully similar to biological episodic memory. Like biological episodic memories, the representations used by DQN's 'experience replay' mechanism carry multidimensional information about events and do so in a manner that is at least partly iconic.

First, multidimensionality. Each state representation is an ordered sequence of observations and actions leading up to the time step at which it is recorded. The observations represent the two-dimensional spatial properties of events. When stacked in ordered sequences, they additionally represent an event's internal temporal properties: the way that its spatial properties changed over time. By interspersing action representations between the observations in the sequence, state representations also represent the agent's involvement in the event, and how the environment changed in response to the agent's actions. *Experience* representations (the 4-tuples described in §2) combine state representations and action representations with reward representations, meaning that in total they carry information about the event's spatial and temporal properties, the agent's actions and how they affected the unfolding of the event, and how rewarding the event was for the agent. This seems like a reasonable approximation to the multidimensional information episodic memories prototypically carry: information about a remembered event's spatial and temporal properties, about the subject's involvement in the event, and about their thoughts, feelings or perceptions of the event.

Second, iconicity. State representations combine observations with action representations in an ordered sequence. As such, they are iconic in two ways. First, observations are iconic representations, in the sense that they are structured, their structural properties map onto the structural properties of their representata, and this mapping is semantically significant. Specifically, as noted in §2, each observation is formatted as a two-dimensional array of numbers, where the two dimensions represent the represented Atari screenshot's height and width, and each position in the numerical array corresponds to the pixel in the corresponding location in the screenshot. Second, combining these observations in sequence, together with action representations, produces another structured representation. In this representation, there is a correspondence between sequence position and time, such that items appearing earlier in the sequence are represented as having occurred at earlier points in time. The semantic significance of the structure of these representations means that they are unlikely to be retrievable piecemeal. Without the semantic information provided by the structured representation as a whole, a fragment of the representation would most likely be uninterpretable.

This is not to say that state representations are *wholly* iconic: in particular, luminance values and actions are represented in symbolic form. But this does not vitiate the claim that these representations are iconic in the ways I've outlined, since representations may combine both iconic and symbolic elements (Lee et al., 2022). We might construe a pixel vector as a hybrid map-like format: the numerical symbols stand for luminance values, whilst their location in the array stands for the locations of those values in the image. The use of symbols for luminance values does not negate the fact that there is a semantically significant structural correspondence between the array and the image. More importantly for our purposes, it is unlikely that episodic memories are wholly iconic: they are likely to involve some conceptual or semantic elements. Since our interest in this section is in the similarity between the representations involved in DQN's experience replay and those involved in episodic memory, a partially iconic representation of events with some symbolic components seems to fit the bill.

4. Episodic Memory's Function

In this section, I expand on the idea that by using DQN as a model of episodic memory, we might advance our understanding of episodic memory's function.

I noted in §1 that the function of episodic memory presents a puzzle for philosophers and scientists of memory. We might express the puzzle in terms of a twofold redundancy. First, given that we have other memory faculties including semantic memory, which stores general purpose information about the world, it seems redundant to have a faculty recording information about specific events. Second, much of the information *stored* in episodic memory appears redundant, as it relates to highly specific events which will never be repeated and is unlikely to be directly useful in any other context. This redundancy, together with evidence that episodic memory is subject to systematic patterns of error and neurally overlaps with our faculty for imagining future and hypothetical scenarios, has led *simulationism* to become the dominant view of episodic memory's function (De Brigard, 2014; Schwartz, 2020). Simulationism is the view that episodic memory's function is primarily to support the imaginative construction of future and hypothetical scenarios.

Against this, several philosophers have recently mounted defences of the mnemonic view of episodic memory's function, on which its role is to store, encode and retrieve information. For example, I've argued that episodic memory facilitates retrospective learning, that is, extracting novel information from an event after the event has passed (Boyle, 2019). Simon Brown (2023) offers a related account, arguing that episodic memory supports *unrestricted* learning. The idea is that by capturing multidimensional information about events, episodic memory enables us to continuously revise and expand our models of the world. Elsewhere, I've also argued (Boyle, 2021) that episodic memory plays a significant role in the storage, encoding and retrieval of semantic memory: it is both critically involved in the ordinary process of laying down semantic memories and provides internally generated cues for their retrieval. In a similar vein, Sara Aronowitz (2022) argues that we cannot understand the function of episodic memory in isolation. The process of semanticization, in which information from episodic memory becomes gradually more abstract and is encoded in semantic memory, suggests that episodic memory can be understood only in the context of a broader memory system.

Evaluating these theories is challenging. A natural way to approach the question would be to compare the behavioural repertoires of agents which have episodic memory with those of agents which lack it. We might do this by comparing humans with and without episodic memory deficits. This can be informative but faces some limitations: it can be difficult to know which behavioural differences are attributable to episodic memory. This is both because brain damage is rarely limited to *only* the brain areas involved in episodic memory, and because the brain areas supporting episodic memory may also support other cognitive functions: brain areas can multitask. Alternatively, we might try to compare the behaviours of animals with and without episodic memory. But the distribution of episodic memory in the animal kingdom is another significant puzzle: there is insufficient agreement about this for us to be sure which animals have or lack it.[5]

Given the resemblance between DQN and episodic memory at the algorithmic level, I propose that DQN and similar algorithms provide a testing ground for

[5] For a discussion of this issue, see Boyle (2022).

competing theories of episodic memory's function. In artificial agents, it is possible to ablate experience replay and to know that there has been no other intervention on the agent's cognitive architecture. Any differences in the agent's behavioural repertoire that result from this intervention can be traced directly to the agent's having or lacking the capacity to replay past events. It is also possible to tweak the inner workings of the experience replay algorithm and observe the effects of these changes. Insofar as artificial experience replay resembles biological experience replay at the algorithmic level–i.e., both provide integrated, partially iconic, multidimensional representations of events–this kind of investigation would provide defeasible evidence about the cognitive role function of episodic memory and might differentiate between accounts which are otherwise difficult to empirically evaluate.

Of course, DQN has not been used to directly test rival theories of episodic memory's function, so we should exercise caution in interpreting the results obtained with DQN in this way. Such caution notwithstanding, those results do suggest some support for mnemonic views, particularly those that emphasize episodic memory's role in semantic learning. In DQN, memories of specific events are used to train a network that learns more abstract, relational knowledge structures about the relationships between states of the environment and action/Q-value pairs. By using event memories in this way, DQN was able to learn much faster than rivals which do not make use of event memories in this way. It also attained a new state of the art in the Arcade Learning Environment, a set of complex problems involving high-dimensional sensory input. Disabling the experience replay component of the algorithm significantly impaired its performance. If we view the relationship between experience replay and the Q-learning network as analogous to the relationship between episodic and semantic memory, this provides some support to the theoretical claim that episodic memory supports the rapid acquisition of semantic memory. At the very least, it vindicates the idea that episodic memory *could* carry out a distinctive mnemonic function, even in the presence of a more general, abstract memory system. And given the similarities between DQN and episodic memory, I suggest, this provides some defeasible evidence that episodic memory carries out a similar function in us.

One might worry that the preceding argument overstates the similarity between DQN's experience replay and our episodic memory capacity. One reason for thinking this is that DQN records details accurately and the contents of its stored 'experiences' are not subject to change. By contrast, biological episodic memories are constructive and friable. We do not 'record' events accurately in entirely faithful detail. Our episodic memories are reconstructed at the time of retrieval, often drawing on general knowledge from semantic memory as well as details from the original event. As such, they are subject to change whenever they are retrieved, leading to systematic patterns of error (see De Brigard, 2014 for discussion). A second, converse, issue is that biological episodic memories may include many details DQN's memories do not. Most of us remember perceptual and sensory information that goes beyond the visual: we might remember sounds, smells, sensations and so on. Moreover, there are types of non-sensory content we recall as well, such as information about the emotions we felt or what we were thinking at the time of the event. In summary: DQN's representations *include* a type and level of detail not typical of episodic memories, whilst also *excluding* some kinds of information characteristic of episodic memories.

However, using DQN to learn about episodic memory's function does not require that the two be exactly similar. Our interest is in using DQN as a *model* of episodic memory, such that we can draw justified inductive inferences about episodic memory by observing and manipulating DQN. Models can be useful in this way even if they are *simplified* or *idealized,* that is, even if they omit certain features of the modelling target, or include some features not found in the modelling target. As Catherine Stinson (2020) argues, what matters is whether both the model and the target belong to a common *kind* which licenses inferences from one to the other. In brief, what I have been arguing in this section is that DQN and episodic memory are members of a common kind: a kind of memory system characterized by the storage, encoding and retrieval of partially iconic representations storing detailed multidimensional information about past events. Moreover, our focal question about the function of episodic memory can reasonably be construed as a question about *this kind* of memory system: what is the use of a memory system which processes detailed, multidimensional representations of past events? So, despite the ways in which its representations differ from those involved in episodic memory, DQN seems like a promising model. Of course, it would take empirical work to establish the utility of this model; what I have been arguing is that there are good theoretical grounds for thinking that this empirical work would be fruitful.

Of course, there may be contexts in which these representational differences between DQN and episodic memory really matter. We might be interested in asking a somewhat narrower question about episodic memory's function, such as whether the reconstructive, error-prone processes that characterize our episodic memory confer epistemic or other advantages (Michaelian, 2013; Puddifoot & Bortolotti, 2018). In this context, a useful model would need to belong to a narrower kind: a *constructive* memory system processing similar representations of past events. DQN does not fall into this category. But this does not show that it would not be useful here, since the DQN algorithm could be adapted to incorporate constructive processes.[6] So, rather than vitiating the use of DQN as a model of episodic memory, the concern suggests how refinements to the algorithm might expand the range of theoretical questions with respect to which it is a fruitful model of episodic memory.

A related concern is that, notwithstanding the representational similarity between episodic memory and DQN, there remain significant differences at the

[6] For an example of a (non-DQN) episodic memory algorithm incorporating constructive processes, see Zakharov et al. (2020)

algorithmic level relating to how these representations are processed. For example, in DQN, representations of past episodes are randomly sampled and used to train the Q-learning network. As I've indicated, semanticization in humans is a candidate analogue for this process: episodic memories are gradually consolidated and abstracted into semantic memory. However, it is unlikely that our episodic memories are sampled *randomly* for this purpose: salience, recency and other factors are likely to have a significant impact on which memories are prioritized for semanticization. A related point can be made about forgetting. In DQN, the oldest stored experiences are erased when the episodic buffer reaches capacity. This does not mirror patterns of human forgetting: we forget many things besides our oldest memories and retain some memories for a very long time.

Again, there may be contexts in which these algorithmic level differences may not matter. If we're interested in understanding how a store of detailed, multidimensional event-specific memories can be used to support the acquisition of more abstract, general knowledge, DQN seems a suitable idealized model of this process. And as before, whilst there are contexts in which these differences do matter, the algorithm might be extended to resemble episodic memory more closely in relevant respects, so as to expand our understanding of episodic memory.

For example, Schaul et al. (2016) develop a DQN-based algorithm by adding prioritized experience replay. In this version of the algorithm, memories are not sampled randomly from the episodic buffer. Instead, memories with the highest temporal difference (TD) error are prioritized–that is, experiences which are more 'surprising' because the reward obtained differs significantly from the reward the system would predict. Alternative prioritization criteria could be used; the choice of TD error here is partly motivated by evidence that experiences with TD error are prioritized for replay in the hippocampus. This variant on DQN exhibited faster learning and a new state of the art in the Atari environment.

We might take this to provide defeasible evidence about the function of forgetting: at first blush, Schaul et al.'s results appear to support the view that forgetting is not a design flaw, but a critical design feature on a well-functioning memory system (Fawcett & Hulbert, 2020; Michaelian, 2011). Forgetting in biological systems can take two forms: information either becomes inaccessible or unavailable. Inaccessibility is a matter of information still being encoded in memory but being more difficult to retrieve; unavailability is a matter of the information having been entirely lost. We can see that deprioritization for replay in Schaul et al.'s algorithm provides an approximate analogue for inaccessibility: when information is deprioritized, it's less likely to be retrieved. As such, the algorithm suggests an adaptive role for at least one kind of forgetting: making events with low TD error less accessible leads to quicker learning and improved policies. Investigating the effects of different prioritization criteria might provide further insights into the function of both memory and forgetting.

As Shaul et al. note, one way to extend their work would be to apply prioritization criteria to erasure, for example, by erasing the memories with the lowest TD error, rather than the oldest memories, when it reaches capacity. Extending the algorithm in this way might shed light on the other form of forgetting: unavailability. In this vein, Ruishan Liu and James Zou (2019) investigate the relationship between the size of the memory buffer and learning rates, finding that learning rates slow when the buffer is either too large or too small. They develop an algorithm in which the size of the memory buffer adaptive changes. If the TD error of the oldest memories is increasing, suggesting that these memories are becoming more informative, the buffer size increases and these memories are retained for longer. On the other hand, if the TD error of the oldest memories is decreasing, suggesting that they are becoming less informative, the buffer size decreases and these less informative older memories are more likely to be erased. Again, this suggests an adaptive role for prioritized patterns of forgetting in learning.

Similarly, we might note that whilst DQN only uses recorded experiences to train its Q network, our episodic memories are clearly put to other uses: most obviously, we frequently retrieve salient episodic memories to inform online decision making. This marks another significant difference in how episodic memories are processed in DQN and episodic memory. Once again, this simply shows that DQN is not a useful model in all contexts, as well as highlighting a way in which we might wish to develop experience replay algorithms to suit particular theoretical goals in cognitive science. If our interest is in investigating episodic memory's role in decision making, we would be better off looking at an algorithm in which recorded experiences are used in a similar way. For example, Blundell et al. (2016) develop an alternative experience replay algorithm they call the 'Episodic Controller'. In this architecture, past experiences are stored in a buffer which the agent can query to inform its decision making. When faced with a decision, the agent uses this body of stored knowledge to determine which actions have previously been associated with the highest reward in situations similar to the one it currently faces. The Episodic Controller learns quickly, especially in the early stages of confronting a novel problem, and particularly in sparse reward environments. In these scenarios, it exhibits behaviour 'akin to one-shot learning' (Blundell et al., 2016, p. 7). This architecture could be a fruitful model of episodic memory for the purposes of developing and evaluating accounts of episodic memory's role in decision making. At first blush, the view suggested seems to be that having access to information about individual, salient episodes facilitates fast learning in novel environments when reward is scarce.

5. Conclusion

Episodic memory presents many puzzles for memory theorists. Among them is the question of its cognitive role function: what does episodic memory contribute to the cognitive systems of which it's a part? I've argued that this question properly cast at the algorithmic level: what is the purpose of a memory system that exploits detailed, multidimensional, partially iconic representations of specific past events? It is difficult to gain empirical traction on this question by looking at biological systems. But, I've argued, DQN provides empirical leverage on the question in virtue of its similarity to episodic memory at the algorithmic level. In particular, against increasingly popular simulationist views, the results obtained with DQN suggest that episodic memory plays a distinctive *mnemonic* role in the process of semantic learning, as several theorists have recently argued. Of course, there are significant differences between DQN and episodic memory in biological systems. But these do not undermine the utility of DQN in this context. Rather, they suggest ways in which we might adapt DQN or similar algorithms in future work, in order to evaluate theories about episodic memory's operations and its distinctive contributions to cognition.

Acknowledgments

Earlier versions of this chapter were presented to audiences at the London School of Economics, Kings College London, the London Mind Group, Duke University and the University of Arizona. I'm grateful to all those present for helpful feedback and discussion. Thanks also to Andrea Blomkvist, Julia Haas and an anonymous reviewer for comments on an earlier draft. This research was supported by a UKRI Future Leaders Fellowship (grant number MR/W00741X/1). There are no data associated with this work.

References

Andonovski, N. (2022). Episodic representation: A mental models account. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.899371

Aronowitz, S. (2022). Semanticization Challenges the Episodic–Semantic Distinction. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/721760

Bellemare, M. G., Naddaf, Y., Veness, J., & Bowling, M. (2013). The Arcade Learning Environment: An Evaluation Platform for General Agents. *Journal of Artificial Intelligence Research*, *47*, 253–279. https://doi.org/10.1613/jair.3912

Blundell, C., Uria, B., Pritzel, A., Li, Y., Ruderman, A., Leibo, J. Z., Rae, J., Wierstra, D., & Hassabis, D. (2016). *Model-free episodic control*.

Boyle, A. (2019). Learning from the past: Epistemic generativity and the function of episodic memory. *Journal of Consciousness Studies*, *26*(5–6), 242–251. https://doi.org/10.17863/CAM.35867

Boyle, A. (2020a). Remembering events and representing time. *Synthese*. https://doi.org/10.1007/s11229-020-02896-6

Boyle, A. (2020b). The impure phenomenology of episodic memory. *Mind & Language*, *35*(5), 641–660. https://doi.org/10.1111/mila.12261

Boyle, A. (2021). The mnemonic functions of episodic memory. *Philosophical Psychology*, 1–23. https://doi.org/10.1080/09515089.2021.1980520

Boyle, A. (2022). Episodic memory in animals: optimism, kind scepticism and pluralism. In A. Sant'Anna, C. J. McCarroll, & K. Michaelian (Eds.), *Current Controversies in Philosophy of Memory* (pp. 189–205). Routledge.

Brown, S. A. B. (2023). Episodic Memory and Unrestricted Learning. *Philosophy of Science*, 1–29. https://doi.org/DOI: 10.1017/psa.2023.16

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, *4*(8), 685–691. https://doi.org/10.1038/nrn1180

Clayton, N. S., & Russell, J. (2009). Looking for episodic memory in animals and young children: Prospects for a new minimalism. *Neuropsychologia*, *47*(11), 2330–2340. https://doi.org/10.1016/j.neuropsychologia.2008.10.011

Clayton, N. S., Yu, K. S., & Dickinson, A. (2001). Scrub jays (Aphelocoma coerulescens) form integrated memories of the multiple features of caching episodes. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(1), 17–29. https://doi.org/10.1037/0097-7403.27.1.17

Crystal, J. D., Alford, W. T., Zhou, W., & Hohmann, A. G. (2013). Source memory in the rat. *Current Biology*, *23*(5), 387–391. https://doi.org/10.1016/j.cub.2013.01.023

Crystal, J. D., & Smith, A. E. (2014). Binding of episodic memories in the rat. *Current Biology*, *24*(24), 2957–2961. https://doi.org/10.1016/j.cub.2014.10.074

Cummins, R. (1975). Functional Analysis. *The Journal of Philosophy*, *72*(20), 741–765.

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, *191*(2), 155–185. https://doi.org/10.1007/s11229-013-0247-7

Eacott, M. J., Easton, A., & Zinkivskay, A. (2005). Recollection in an episodic-like memory task in the rat. *Learning and Memory*, *12*(3), 221–223. https://doi.org/10.1101/lm.92505

Fawcett, J. M., & Hulbert, J. C. (2020). The Many Faces of Forgetting: Toward a Constructive View of Forgetting in Everyday Life. *Journal of Applied Research in Memory and Cognition*, *9*(1), 1–18. https://doi.org/10.1016/J.JARMAC.2019.11.002

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*, 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Lee, A. Y., Myers, J., & Rabin, G. O. (2022). The Structure of Analog Representation. *Nous*. https://doi.org/10.1111/nous.12404

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689. https://doi.org/10.1037//0882-7974.17.4.677

Liu, R., & Zou, J. (2019). The Effects of Memory Replay in Reinforcement Learning. *2018 56th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2018*, 478–485. https://doi.org/10.1109/ALLERTON.2018.8636075

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman.

Michaelian, K. (2011). The Epistemology of Forgetting. *Erkenntnis*, *74*(3), 399–424. https://doi.org/10.1007/s10670-010-9232-4

Michaelian, K. (2013). The information effect: Constructive memory, testimony, and epistemic luck. *Synthese*, *190*, 1–28. https://doi.org/10.1007/s11229-011-9992-7

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Puddifoot, K., & Bortolotti, L. (2018). Epistemic innocence and the production of false memory beliefs. *Philosophical Studies*. https://doi.org/10.1007/s11098-018-1038-2

Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*(1), 1–23. https://doi.org/10.1037/a0037907

Schaul, T., Quan, J., Antonoglou, I., Silver, D., & Deepmind, G. (2016). *Prioritized experience replay*. 1–21.

Schwartz, A. (2020). Simulationism and the Function(s) of Episodic Memory. *Review of Philosophy and Psychology*, *11*(2), 487–505. https://doi.org/10.1007/s13164-020-00461-1

Shea, N. (2014). Exploitable isomorphism and structural representation. *Proceedings of the Aristotelian Society*, *114*(2), 123–144. https://doi.org/10.1111/j.1467-9264.2014.00367.x

Stinson, C. (2020). From Implausible Artificial Neurons to Idealized Cognitive Models: Rebooting Philosophy of Artificial Intelligence. *Philosophy of Science*, *87*(4), 590–611. https://doi.org/10.1086/709730

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). The MIT Press.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–402). Academic Press.

Tulving, E. (2005). Episodic memory and autonoesis: uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The Missing Link in Cognition* (pp. 4–56). Oxford University Press.

Zakharov, A., Crosby, M., & Fountas, S. (2020). *Episodic memory for learning subjective-timescale models*. https://arxiv.org/abs/2010.01430v1