Exploring the Space of Computational Memory Models

Nate Derbinsky[1] and Nicholas A. Gorski[1]

Abstract. Over their lifetimes, intelligent agents gain knowledge that may be pertinent to their decisions about acting in the world. One goal of memory system research is to develop an optimal set of encoding, storage, and retrieval mechanisms that will harness these experiences to facilitate rational decisions. In this paper, we propose a direction of empirical, computational research that seeks to better understand the behavioural dynamics that arise when an agent endowed with long-term memory is situated in a task, by determining which properties of task and characteristics of memory systems in combination are responsible for which aspects of behaviour. We propose preliminary taxonomies for task and memory spaces, and also propose metrics for systematic evaluation of memory systems.

1 INTRODUCTION

Learning agents gain knowledge that may be pertinent to making intelligent decisions. To harness these experiences while remaining reactive to dynamic environments, cognitive architectures employ one or more *memory systems*: mechanisms that efficiently implement a fixed policy to encode, store, and retrieve agent knowledge [6]. Prior work provides significant psychological [10][14] and computational [3][8] evidence for dissociated memory systems [12][15][16], yet a significant challenge lies in understanding, for a particular task, the functionally optimal set.

One approach to this problem is to empirically explore the functional benefits of individual classes of memory systems in isolation. For instance, Nuxoll [7] has provided theoretical and empirical evidence that endowing an intelligent agent with a task-independent episodic memory affords it a multitude of cognitive capabilities that may be crucial to the efficacy of a robustly intelligent agent. While insightful and focussed, this approach has its limitations. For instance, implementing a memory system may prove to be quite challenging [2], limiting the speed with which one can study the space of design decisions over multiple tasks. As a compounding factor, a memory system's efficacy may also depend on how well it supports an agent's ability to learn to use it [4]: this encompasses not only how an agent might learn to condition behaviour in the environment based on knowledge that is retrieved from memory, but also how an agent learns to perform encoding, storage and retrieval actions to control the memory mechanisms themselves.

Comprehensive empirical study of a large design space has been done before. One example is in reinforcement learning (RL) [13], which has achieved broad and successful application in cognitive and computer science, partially due to its general formulation and dependency only on a scalar reward function. While RL offers insight into adaptive control given this function, it does not specify the source or nature of reward. Much work, for instance [1], has shown the importance and difficulty of specifying an optimal reward signal. Therefore, recent work looks to empirical study over reward function space [11]. Similar approaches are being applied to studying the large design spaces of cognitive architecture [5] and empirical game modeling [9].

In this paper, we propose a similar direction of empirical, computational research over the space of memory models, such as to understand the connection between properties of task, characteristics of memory systems, and an agent's learned behaviour. Our hypothesis is that for different classes of problem spaces (distinguished along the dimensions of task that we propose below), different classes of memory models will result in qualitatively different behaviour. By empirically exploring the interactions that arise between different memory models and tasks, we will improve our understanding of which memory models are appropriate in which classes of problem spaces. While many of the dimensions of memory that we identify are inspired by research into the properties of human memory, we are also interested in exploring computational memory models that differ from human memory along certain dimensions; this may allow us to formulate principled functional arguments for why human memory mechanisms might have certain properties. The ultimate goals of our research, then, are (1) to understand how memory system requirements and efficacy change along with parameterized properties of task and (2) to develop computational structure and constraints for studies of memory system dissociation, ideally resulting in evidenced sets of generally useful memory systems.

To accomplish our goals, we propose a methodological framework of empirical study that will enable us to measure the behavioural implications of the interaction between properties of task and characteristics of architectural memory systems. As conceptually depicted in Figure 1, we assume an adaptive agent, *A*, possessing fixed initial procedural knowledge but endowed with the ability to learn additional control knowledge (over actions both in its environment and its internal memory mechanisms [4]). This agent is endowed with a set, *M*, of one or more memory systems selected from architectural memory system space (Section 2), and situated within a domain, *T*, selected from task space (Section 3). We posit that quantitative and qualitative analysis of the agent's learned behaviour (Section 4), while systematically varying *M* and *T*, will serve to objectively evaluate the efficacy of sets of memory systems under different task conditions. After describing these spaces and metrics, we propose an approach for exploring the space of possible memory models that we believe will yield insights into the relationships between memory and task (Section 5).

We propose that agents in our study use RL to modify their control knowledge over time. We are interested in adaptive agents, rather than those with fixed procedural knowledge, as it is difficult to predict the best strategies for memory usage across

[1] Computing Science and Engineering Div., Elec. Engineering and Computer Science Dept., Univ. of Michigan, 2260 Hayward St., Ann Arbor, MI, 48109-2121. Email: {nlderbin, ngorski}@umich.edu.

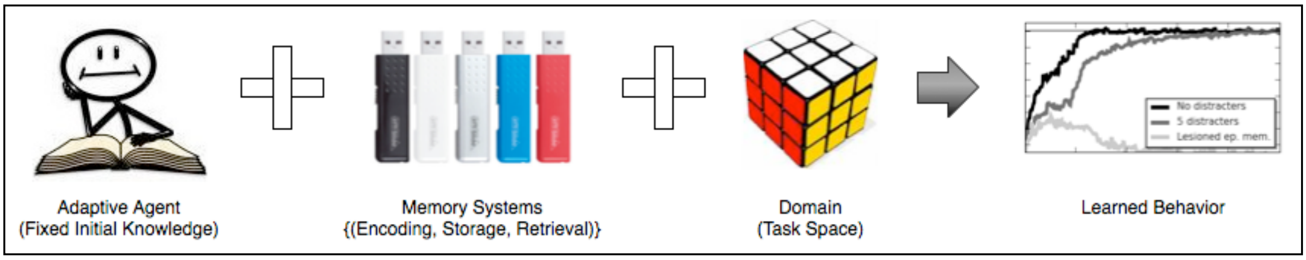

Figure 1. A conceptual depiction of the methodological framework that we will use in our investigations.

arbitrary combinations of memory models and problem spaces. RL offers a framework for optimizing agent control knowledge, without the need for an explicit model of the external constraints of the environment, nor the internal constraints of the agent architecture.

2 MEMORY MODEL SPACE

We have developed the following initial characterization of the space of memory systems, borrowing heavily from and generalizing Nuxoll's breakdown of the space of episodic memory systems [7]. These dimensions of memory mechanisms will be explored in combination with dimensions of task space (section 3) in order to better understand the dynamics that relate memory usage to task behaviour.

Encoding

- *Initiation* – Initiation encompasses the event conditions that trigger the encoding and storage processes. These events may condition upon fixed architectural characteristics of state (such as a temporal frequency) or may be accessible to agent control knowledge.

- *Determination* – Once initiated, the memory mechanism selects features of agent state (or derivation thereof) that compose the knowledge to be stored, as well as any additional context (temporal, spatial, etc) that may also be associated with the knowledge.

Storage

- *Granularity* – Stored experience varies with the grain size at which knowledge can accessed and modified. This may range from minute (such as the symbol level), to moderate (an episode), to coarse (such as the entire knowledge store).

- *Dynamics* – Knowledge in the memory system may change over time, such as to bias retrieval or forget knowledge. The mechanisms that cause this change may be fixed, condition upon agent knowledge, or deliberate agent action.

Retrieval

- *Accessibility* – Experience encoded within the memory system may vary in the degree to which it is exposed to other architectural mechanisms, such as to maintain overall agent reactivity. For instance, a declarative long-term memory may allow for enumeration of all stored memories.

- *Initiation* – Initiation encompasses the event conditions that trigger the retrieval process. As with encoding initiation, these events may condition upon fixed characteristics of state or may be accessible to agent knowledge/control.

- *Cue Determination* – Once initiated, the memory system composes agent state, knowledge, context, and/or [possibly inaccessible] meta-data to select or create a retrieval cue.

- *Selection* – When supplied a cue, the memory system implements a policy for how stored knowledge is matched with respect to the cue, which may be restricted by time, computation, and/or number of results, as well as include bias from agent state, context, and/or meta-data.

- *Result* – When the memory system selects stored experience for retrieval, it may arbitrarily represent the knowledge, associated context, and aspects of the retrieval process, such as match quality, for agent inspection.

We define a memory system implementation as a commitment to features from within the space defined by these dimensions, represented as (encoding, storage, retrieval).

When deliberate agent knowledge can affect architectural memory mechanism function, such as to initiate encoding or retrieval, the agent has a space of actions that it can execute in order to affect this change. These actions are modulated by the same action selection process that controls actions in the external environment, and may vary from the null set (where all encoding, storage and retrieval is architecturally fixed and not affected by the agent's central action selection loop) to where any access to memory (encoding, storage, or retrieval) must be deliberately selected by the agent's internal reasoning system.

3 TASK SPACE

We have identified a preliminary taxonomy of environmental characteristics. These characteristics can be quantitatively parameterized and are independent of each other. This allows for a principled empirical exploration of task space, where an agent with fixed memory models can be situated in a task that varies along one or more dimensions. While these characteristics are not comprehensive, they do offer a preliminary set with which we can begin to explore the circumstances in which particular classes of memory models afford advantages over others.

- *Temporal Distance to Salient Knowledge* – Consider a task in which the agent perceives a sign containing a single element of knowledge (for example, either the symbol "A" or "B"). Later, the agent encounters two doors, and must choose which door to open. If it chooses correctly, then it receives positive reward; incorrectly, negative. Each door is identified by a symbol; the symbol on the correct door also appeared on the sign. The temporal distance between the salient knowledge is a parameter that can be varied, and although in the example only a single element of knowledge was considered, it can be varied across all elements of salient knowledge in a task. How this characteristic of task interacts with a particular memory mechanism will depend on the persistence of knowledge in memory, as well as whether the temporal distance will affect how easily knowledge can be retrieved from memory.

- *Categories of Salient Knowledge* – We refer to the number of distinct elements of knowledge that must be simultaneously maintained in memory while an agent acts as the number of categories of knowledge. If the example above were to be extended, such that there were three signs, each with a unique symbol corresponding to three sets of doors through which the agent must pass, then there would be three categories of salient knowledge vs. one category for the original example. In complex domains, there might be large amounts of knowledge that can be perceived in the domain, but relatively few categories of actual salient knowledge that must be brought to bear on reasoning in order to act well, requiring only very limited capacity in a memory model. Alternatively, there could be a very large number of categories, in which case an effective memory model would require large capacity.

- *Quantity of Salient Knowledge* – We refer to the number of possible values that each category can take on as the quantity of knowledge. This can vary from one to infinity for each category. Intuitively, this is the number of distinct symbols that could appear on a sign in the example task.

- *Quantity of Distracting Observations* – The number of observations that are irrelevant to acting in the world (i.e. are not salient) will have an effect on how easily an agent can retrieve salient knowledge from memory. These observations are distracters across all tasks in a study, not just a single task instance.

- *Sparseness of Reward* – As rewarding events become more sparsely distributed, a learning agent has more difficulty in determining which actions were beneficial and which were not. This will directly affect how well a learning agent is able to adapt to constraints of task and memory.

- *Relative Cost of Acting vs. Reasoning* – Actions that don't lead to immediately rewarding events may be associated with a penalty. The penalty for acting in the world vs. for accessing memory can vary.

- *Size of Action Space* – Particularly pertinent to adaptive agents, the size of the action space will determine how much exploration the agent must undertake, and will affect how easily the agent is able to learn to use memory (if the agent has adaptive control over internal memory actions).

- *Stochasticity of Actions* – As with the size of the action space, this characteristic is particularly relevant to adaptive agents and will in part determine how much exploration in the environment is necessary, and thus how easily an agent can learn to control memory.

4 EVALUATION METRICS

Once the agent, endowed with its memories, has completed the task, we must evaluate its learned behaviour both quantitatively and qualitatively. As to the former, we could consider any/all of the following learning metrics:

- Average reward-per-step
- Maximum reward attained
- Speed of convergence

These measures can be tempered by the amount of memory system storage. We also introduce the concept of *knowledge coverage*. As depicted in Figure 2, at each decision, there is a certain amount of agent experience pertinent to making an optimally rational decision. A retrieval from any given memory system, such as an episodic or semantic memory, may achieve some subset of this knowledge (and multiple retrievals may well overlap). The concept of coverage attempts to get at the role of memory systems in making rational decisions. From a quantitative standpoint, an ideal memory system set on a task will learn quickly, converging to a relatively high reward-per-step, while obtaining maximal pertinent knowledge coverage using the least amount of storage.

Qualitative analysis is also crucial to understanding the role of memory systems, and we have two intended directions. First, our qualitative analysis will involve categorizing behaviours supported by memory in accordance with the cognitive capabilities that agents exhibit [7]. For example, in [4], the authors identified behaviour that was suboptimal but involved using an episodic memory mechanism as a stigmergic memory to condition behaviour. If some memory models support a subset of the cognitive capabilities afforded by other models, then an analysis of how the memories vary along the space of characteristics will improve our understanding of memory. Clustering algorithms applied to agent trajectories can assist in detecting classes of behaviour, which can then be examined by hand and categorized as to how memory was used by the agent.

Second, our qualitative analysis will also take the form of comparing agent actions in a single task across different points in memory space. By doing so, we can gain a conceptual understanding of the functional benefits and drawbacks of a

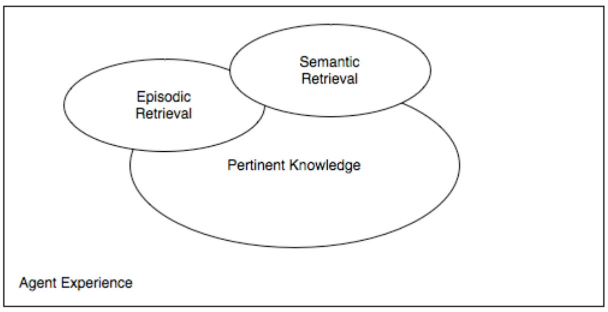

Figure 2. Knowledge coverage

particular memory system. We do not yet have a definitive understanding of how to do this systematically, especially since many points in memory system space may have no direct or even intuitive parallel in humans, but we are confident that a programmatic approach will yield satisfactory results.

5 PROPOSED EXPLORATION

Having discussed characteristics of memory, properties of task, and evaluation metrics, we now discuss our proposed direction for systematically exploring the space of computational memory models in order to understand the functional capabilities that they afford intelligent agents. As the possible spaces of memory models and tasks is infinitely large, our approach is to identify several specific points in the space and initially begin expanding the space of models that we consider starting from those specific points. The initial points in this space will be selected based on biological inspiration, as biological memories are existence proofs of significance, and similarity to existing computational memory models (for relevance).

We plan on pursuing two initial directions. The first, inspired by work performed by Gorski and Laird [4], is focused on exploring the functional characteristics of memory mechanisms situated in tasks but in the absence of adaptive control. In this direction, we will begin with one type of memory mechanism and vary it along several dimensions to better understand how architectural commitments impact the functional capabilities of a memory model.

For example, we have considered systematically sampling the space of episodic memory systems. Given Tulving's functional requirements for an episodic memory implementation, we could fix encoding initiation as automatic (at some architectural frequency), architectural (meaning deliberate agent knowledge cannot modify stored knowledge), and implement a deliberate, bit-vector retrieval policy, using nearest-neighbor (NN) symbolic match and recency as a tie-breaking bias. While holding task properties constant, we could systematically and efficiently vary encoding determination as a function of perception, activation, and other architectural state. Furthermore, we could introduce *forgetting* dynamics (meaning, architectural decay of stored bits) while sampling across a range of temporal distance to salient knowledge. All of these modifications are trivial to implement in a bit memory, and performance is not an issue. In a learning task, we would hypothesize that an agent whose forgetting mechanism decayed faster than the temporal distance would perform quantitatively no better than an agent without a long-term memory, and probably learn not to make use of deliberate memory retrievals.

The second direction focuses primarily on the interaction of learning to control memory. Here, we will begin with several initial memory models: Soar's episodic memory mechanism [2], Soar's semantic memory [3][6], and a simple computational bit memory. We will explore agents that learn to control these memories in several different tasks that vary along the dimensions described above so as to understand their initial functional limitations. When we identify differences in the tasks that agents can learn to use each memory model, we will begin modifying the memory models along dimensions that are most relevant to the functional differences that we have observed in order to determine which characteristics of memory have the biggest functional effects on which types of tasks.

From a computational standpoint, there are many unresolved issues with our approach. For instance, the memory and task taxonomies we have proposed are far from comprehensive and, even if they were, it is unclear how to computationally instantiate an element within these spaces. Given these instantiations, we are also far from realizing an intelligent agent in which we have confidence to optimally learn control over any set of memory systems. Finally, once we overcome these hurdles, the space of memory systems is truly vast, and any comprehensive search therein will be a significant computational and analytical challenge.

6 CONCLUSION

In this paper, we have proposed a principled approach to evaluating memory systems. While there is much analytical and computational work left to implement, we believe this direction of investigation will lead to a better understanding of the role of memory in intelligent agents situated within interesting tasks.

REFERENCES

[1] Abbeel, P., Ng, A.Y.: Apprenticeship Learning via Inverse Reinforcement Learning. In: Proceedings of the Twenty-first International Conference on Machine Learning (ICML), (2004)
[2] Derbinsky, N., Laird, J.E.: Efficiently Implementing Episodic Memory. In: Proceedings of the 8th International Conference on Case-Based Reasoning (ICCBR), pp. 403-417 (2009)
[3] Derbinsky, N., Laird, J.E.: Extending Soar with Dissociated Symbolic Memories. In: Symposium on Remembering Who We Are – Human Memory for Artificial Agents, AISB (2010)
[4] Gorski, N.A., Laird, J.E.: Learning to Use Episodic Memory. In: Proceedings of the 9th International Conference on Cognitive Modeling (2009)
[5] Howes, A., Lewis, R.L., Vera, A.: Rational Adaptation Under Task and Processing Constraints: Implications for Testing Theories of Cognition and Action. In: Psychological Review, Vol. 116 (4), pp. 717-751 (2009)
[6] Laird, J.E.: Extending the Soar Cognitive Architecture. In: Proceedings of the 1st Conference on Artificial General Intelligence (AGI), pp. 224-235 (2008)
[7] Nuxoll, A.: Enhancing Intelligent Agents with Episodic Memory. PhD Dissertation, University of Michigan (2007)
[8] O'Reilly, R.C.: Modeling Integration and Dissocation in Brain and Cognitive Development. In: Processes of Change in Brain and Cognitive Development: Attention and Performance, Munakata, Y. & Johnson, M.H. (Eds.), Vol. 21, pp. 375–402 (2006)
[9] Schvartzman, J.L., Wellman, M.P.: Exploring Large Strategy Spaces in Empirical Game Modeling. In: AAMAS-09 Workshop on Agent-Mediated Electronic Commerce (2009)
[10] Sherry, D.F., Shacter, D.L.: The Evolution of Multiple Memory Systems. In: Psychological Review, Vol. 94 (4), pp. 439-454 (1987)
[11] Singh, S., Lewis, R.L., Barto, A.G.: Where Do Rewards Come From?. In: Proceedings of the 31st Annual Conference of the Cognitive Science Society, pp. 2601-2606 (2009)
[12] Squire, L.R., Knowlton, B., Musen, G.: The Structure and Organization of Memory. In: Annual Review of Psychology, Vol. 44, pp. 453-495 (1993)
[13] Sutton, R.S., Barton, A.J.: Reinforcement Learning: An Introduction (1998)
[14] Tulving, E.: Elements of Episodic Memory (1983)
[15] Tulving, E.: Multiple Memory Systems and Consciousness. In: Human Neurobiology, Vol. 6 (2), pp. 67-80 (1987)
[16] Tulving, E.: Organization of Memory: Quo Vadis? In: The Cognitive Neurosciences, Gazzaniga, M.S. (Ed), pp. 839-853 (1995)