The Artificial Memory of Mr. Polly

Memory Simulation in Databases and the Emergence of Knowledge

GISELA SUSANNE BAHR, WILLIAM H. ALLEN, PHILIP J. BERNHARD AND **STEPHEN WOOD**

ABSTRACT

Human memory may be characterized by five dimensions: (1) large capacity; (2) associativity; (3) diversity of memory systems; (4) change over time; and (5) a unified memory experience. The organization and multidimensionality underlying memory can be represented with set theory. This offers a new mathematical perspective, which is the foundation for the cognitive memory architecture Ardemia. The authors present a relational database implementation of Ardemia that supports the creation of the artificial memory of Mr. Polly, the main character in H.G. Wells's novel *The History of Mr. Polly.* In addition to the implementation of Mr. Polly's artificial memory using TimeGlue, his memory is probed with a collection of everyday memory queries that are related to temporal and schema knowledge. The investigation of Mr. Polly's knowledge suggests an alternative representation of schemas; rather than fixed structures or explicit associations, it is possible to model schemas as the results of the interaction between existing knowledge and remembering.

Human long-term memory is an example of naturally occurring big data. It is vast, probably in the order of exabytes, and can last up to a lifetime [1]. Given the rapid development of technology and the capabilities of big data storage and querying, we asked whether we could create a system of artificial memory that resembled the long-term memory of a person. Such a system would be instrumental for conducting research on memory on realistic time and capacity scales. For example, artificial memory could be created, cloned and systematically manipulated to investigate memory phenomena: This may include the study of the forgetting mechanism, simulating how and what we forget, or an investigation of the emergence of time and schema knowledge. In addition to the use of artificial memory as a simulation environment and research tool, future applications may create personal artificial memory in the form of mobile or wearable devices. They may assist users as cognitive multimedia support systems for reviewing and searching memory data, privately and as needed in life-logging or quantified-self applications.

Given that this was a first attempt at creating a human-like memory system with relational database technologies, it seemed advisable to avoid the ethical complications and privacy concerns of working with the memories of a real person. Instead we investigated a fictitious character: Mr. Alfred Polly, the main character in the novel *The History of Mr. Polly* by H.G. Wells (Fig. 1). Over 100 years after the original publication, we brought Mr. Polly's memory to life with the goal of investigating his everyday memory.

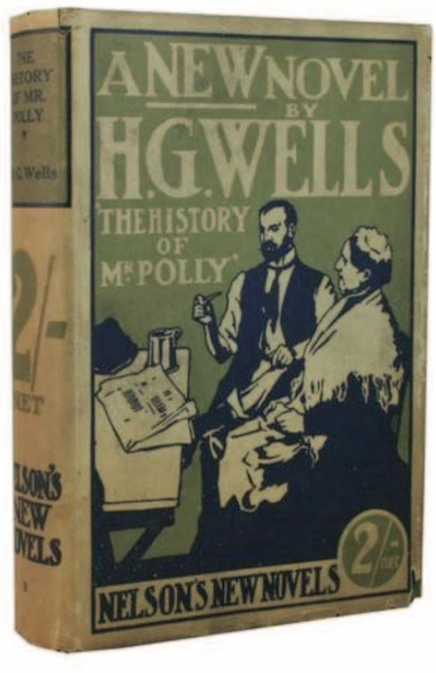

Fig. 1. The first edition of H.G. Wells's *The History of Mr. Polly* (Nelson, 1910). Public domain.

Gisela Susanne Bahr (artist and scientist), DATA314, Melbourne, FL 32901, U.S.A. Email: gsbahr@outlook.com.

William H. Allen (research scientist), Harris Institute for Information Assurance (HIAI), Florida Institute of Technology, Melbourne, FL, U.S.A. Email: wallen@fit.edu.

Philip J. Bernhard (educator, research scientist), Department of Computer Science and Cyber Security, Florida Institute of Technology, Melbourne, FL, U.S.A. Email: pbernhar@cs.fit.edu.

Stephen Wood (ocean engineer), Underwater Technologies and Robotics, Ocean, Marine and Environmental Systems, Florida Institute of Technology, Melbourne, FL, U.S.A. Email: swood@fit.edu.

See www.mitpressjournals.org/toc/leon/52/3 for supplemental files associated with this issue.

 doi:10.1162/LEON_a_01441

This article has two main parts in addition to the introduction and the conclusions. The first presents an overview of the memory simulation architecture and environment Ardemia [2], with focus on two unique aspects related to temporal processing. The second part presents Mr. Polly's artificial memory system and an investigation of his everyday memory for time-related information and schema knowledge.

ARDEMIA

The multidimensionality of long-term memory can be conceptualized with set theory and has been implemented in a long-term memory architecture called Ardemia. This part of the article presents Ardemia's conceptual foundations in set theory, including the Time Universe of Memories, and a relevant implementation aspect, TimeGlue. We chose to highlight time-related features of the architecture because of their importance for the representation for temporal and schema memories presented in the second part.

Five Dimensions of Long-term Memory

Human long-term memory can be described with five dimensions:

- Capacity, which refers to the vast storage volume of human long-term memory
- Associativity, which refers to the intrinsic connectivity of memories and memory systems
- Diversity, which refers to multiple related long-term memory systems
- Unified memory experience, which refers to the seamless integration of memory systems during recall
- Change over time, which refers to the fact that human long-term memory changes constantly as new data are added, and older data or memories are modified or forgotten

This next section presents how this multidimensionality can be conceptualized with set theory and modeled as the Time Universe of Personal Memory.

Human Memories as Sets in the Time Universe

Set theory was first formalized by German mathematician Georg Cantor. It is concerned with the investigation of sets and their properties: "By set we mean any collection M into a whole of definite, distinct objects m (m are the 'elements' of M) of our perception [Anschauung], or of our thought" [3].

To paraphrase, a set represents a *unified whole* that is constructed from individual elements that are conceivable by the human mind. The set concept can be readily adapted to define memory. Like a set, the whole of a person's long-term memory is a collection of conceivable definite, distinct experiences, facts and knowledge.

How memory is organized has been a matter of empirical investigation, and it appears that memory is composed of multiple overlapping systems (Fig. 2). These systems can be viewed as sets of related memories that are implicit, explicit, episodic, semantic, sensory and procedural in nature. The domain of these memory sets is the whole of a person's

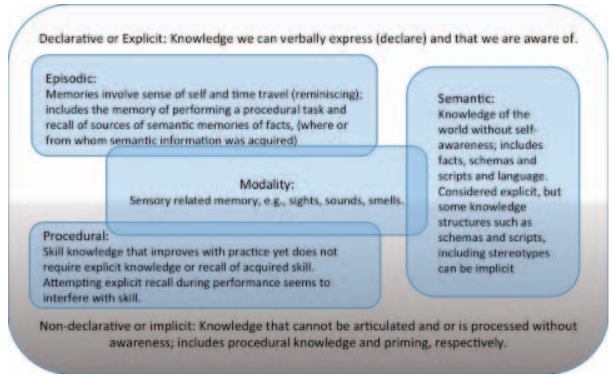

Fig. 2. Long-term memory systems as related sets in the Personal Memory Universe. (© Gisela S. Bahr)

memory in its current state. This can also be called a person's memory universe. In addition to the personal memory universe, Fig. 2 shows that the systems (sets) are related. Set intersections indicate, for example, that skill memory includes sensory (modality) memories such as the unique feel of one's usual keyboard while playing the piano or typing. Within the context of a person's memory universe, the organization of memory systems as sets and their relatedness can be formalized as propositions of their functional associativity. Examples include:

- *implicit and explicit memories*, which are mutually exclusive. Hence, their set representations do not intersect.
- *episodic memories*, remembrances of self that we can remember. Hence, the episodic set intersects with the explicit set.
- *semantic memories* consisting of knowledge that can be recalled implicitly or explicitly. Hence the semantic set intersects with implicit and explicit sets.
- *sensory memories*, overlapping in part with all memories but not necessarily contained by them.

It is evident that the associations between elements and sets are function-based rather than explicitly drawn (as in a graph or a network, for example). Therefore the number of possible combinations, the expressible associativity, is only limited by the functional relationships and operation performed on the data. A full functional analysis is beyond the scope of this article, but a relevant observation is that sets can be related to each other and that based on these relationships, elements can be drawn from different sets to form a new set. Such constructed sets (called relations) are similar to remembering in such a way that recall is the process of reassembling data from multiple systems (sets) into a coherent memory (a new set) without apparent effort or awareness of these processes.

The Time Universe of Memories

Modeling human memory systems as interrelated sets embodies the dimensions large capacity (a person's memory universe), diversity of systems (multiple sets), associativity (functions) and the unified memory experience (relations), but not change over time. Simply put, the current state of the memory universe is mapped to a fixed point in time and cannot be modified as a result of this mapping. If we imagine time as a continuous line, then adding or changing memories requires two things: (1) the expansion of the personal memories to include multiple states and (2) the mapping of these states to different points on the timeline. As such, a time-sensitive memory system may be constructed from multiple versions of a personal universe. These distinct versions exist in parallel and map to corresponding points on a temporal axis; collectively, the versions of a person's universe form a new and greater context: the Time Universe of Memories. In this Time Universe, personal memories are no longer static but express dynamic, temporal phenomena such as forgetting, reorganization and knowledge acquisition. Hence the dynamic nature of memory emerges in the Time Universe by animating the various states of personal memory along the time axis.

Viewing Human Long-Term Memory from a set-theoretical perspective creates a new mathematical approach for the modeling and simulation of memory. This approach can be related to systems designed for storing and managing big data that are based on set theory. The next section presents an implementation aspect using an appropriate, set-theory related database approach based on E.F. Codd's relational model of data [4].

TimeGlue

To create an artificial memory architecture, the majority of memory dimensions can be readily modeled and implemented in a relational database management system such as MySQL or MSSQL. For example, large capacity is an inherent feature of databases, associativity is defined by functions and results from operations on existing data, systems diversity is represented by different datasets (tables) and datatypes, and unified memory experience emerges during query processing as the construction of new relations from existing data. Likewise, change over time is an integral aspect of modern database solutions, and multiple approaches exist to managing data over time. A review of over 20 temporal models led us to the design of a temporal variable called TimeGlue. (For a list of modeling approaches see online Appendix A.)

TimeGlue is instrumental for the modeling of the temporal axis of the time universe. It was inspired by Pavlov's contiguity principle, which states that perceptions and events occurring at the same time are linked in memory [5]. TimeGlue thus represents the time of encoding in natural memory and was operationalized as a timestamp that attaches itself to every memory in every system at the time of encoding (transaction time).

TimeGlue supports the construction of a temporal axis, because timestamps do not repeat and can be ordered to form a linear axis. The relationship between TimeGlue and each memory datum is unique (one to one) within each system and is shared across systems. Sharing a timestamp across systems creates an emergent phenomenon, the linking of systems based on transaction time. This is instrumental during simulated remembering, because TimeGlue permits the synchronization and coordination of simulated memory systems.

In addition to temporal ordering and associativity, TimeGlue supports logical time, which is constructed from inferences based on existing data. Logical time can involve temporal relationships (first, last, before, after or "around some date"), intervals, frequencies, etc.

The next part of this article presents the implications of the temporal modeling for the probing of an artificial memory. The focus is on temporal memory and schema emergence.

MR. POLLY'S ARTIFICIAL MEMORY

This part consists of a brief introduction to Mr. Polly's memory and the first attempt to probe a person's artificial memory for temporal and schema knowledge using a series of everyday memory probes.

The person whose memory was implemented in Ardemia belongs to the main character of a novel by H.G. Wells, Mr. Alfred Polly. We sampled a period during his early twenties while he worked in a small shop in Cambridge, England, selling men's clothing and shoes. Wells describes him as a young man little interested in his job and more concerned with adventure books and the opposite sex.

For the primary data population, memories were extracted from the book. They involved memories of people, events and mundane tasks, such as getting up, eating meals, talking to colleagues, shopping, etc. Data also included plausible emotions that could be reasonably inferred from Mr. Polly's personality. Secondary sources supplied historical and weather data.

Asking Mr. Polly Questions

To ask Mr. Polly questions, it was necessary to probe his artificial memory. Because the number of possible questions is only limited by one's imagination, we arbitrarily selected twelve probes into temporal and schema knowledge.

Unlike temporal knowledge, which seems to be an intuitive property of memory, the concept of cognitive schemas was not discussed until Frederic C. Bartlett first defined the term in a 1932 theory of remembering [6]. His ideas about knowledge abstracted from prior experiences were largely forgotten until Marvin Minsky referred to schemas as the basis for intelligent behavior [7]. Cognitive schemas and scripts are now considered one of the hallmarks of cognition and memory organization [8].

One might think of a schema as a mechanism for data reduction, which allows us to abstract generic information and use it as a blueprint for the past and expectation of the future. Details are forgotten and typical patterns are recalled instead. What remains a topic of investigation is how schemas are represented in memory.

Given the intrinsic link between time and the emergence of typicality, we focused on a selection of memory probes that involve temporal inferences and schema knowledge. As seen in Table 1, Questions 1–4 tap into Mr. Polly's logical time relating to events in episodic memory and Questions 5–12 inquire about facts that are patterns he may have acquired

TABLE 1. Temporal and Schema Queries: Q&A with Mr. Polly's Control Memory

Q1: *What have you been dreaming about recently?*	*sheep, world war 2, India, work*
Q2: *How did you feel when he first met Sam?*	*friendly*
Q3: *Where did you last meet Sam?*	*teashop*
Q4: *Who did you meet first? Sam or Lance?*	*Sam*
Q5: *How many times per week do you usually pick up your laundry from Mrs. Pipps?*	*1*
Q6: *Where did you usually see Ruby?*	*in his bed*
Follow-up question: Sorry to be indiscreet but what exactly did you do?	*daydreams about Ruby, dreams about Ruby, thinks about Ruby*
Q7: *Mr. Polly, What item do you sell the most?*	*Pants*
Q8: *Mr. Polly, What do you sell the least?*	*Cufflinks*
Q9: *Mr. Polly, Do you less gloves or less mufflers?*	*Gloves*
Question10: *Mr. Polly, In the springtime do you sell more hats or handkerchiefs?*	*handkerchief*
Q11: *What you usually dream about at night?*	*Sheep, Ruby, work*
Q12: *What you usually do in the mornings?*	*wakes up, eats, sleeps*

while working or socializing. These questions are deliberately reflective of everyday tasks, events and remembering. They were coded with structured query language (SQL) [9].

The replies to the questions (query output) are also given in Table 1. Mr. Polly's answers are brief, factual and honest. It is apparent that the response language is choppy and in third rather than first person. Future versions of Mr. Polly's artificial memory will address these limitations.

The results of the memory probes indicate that Mr. Polly's artificial memory is capable of producing answers that are not explicitly coded into artificial memory but instead emerge as the result of query processing. It appears that logical time reveals temporal knowledge and schema-like knowledge that can be inferred from instances that have occurred over time. Hence Mr. Polly's memory is not encumbered with the explicit storage of derived patterns but is capable of generating inferences about time and schemas on the fly, as the result of query processing.

CONCLUSION

Schank makes the point that intelligence and higher-order cognition are often recall and memory in disguise [10]. For instance, when answering questions and in conversation, humans rarely create innovative answers but instead seem to rely on existing knowledge of previous events, general facts and personal preferences, as well as typical patterns and expectations. To store such knowledge, it has been suggested that schemas could be represented by data containers, similar to frames [11] or numbers of connections [12] that are considered part of memory; the simulations of Mr. Polly's recall suggest an alternative model: Schemas could be the results of a data-process interaction.

Schemas as abstractions from accumulated experience, created by processes that retrieve and organize existing memories into new knowledge, have a provocative implication for human cognition: Similar to artificial memories in Ardemia, schema-like abstractions, such as stereotypes, may be cognitively constructed on the fly. Therefore stereotypes and what appear to be rigid preconceptions are potentially malleable cognitive phenomena. As such, they can be influenced and possibly changed by acquiring new memories.

The investigation of Mr. Polly's artificial memory reveals that artificial memory systems are capable of providing intelligent answers, demonstrating schema and temporal knowledge that is not explicitly stored or coded into memory. Therefore simulated remembering is not limited to retrieval of prior knowledge or stored data but integrates retrieval and computation in response to a specific question. The artificial memory presented here was implemented in Ardemia, a set-theoretically based, relational memory architecture that embodies the five dimensions of human long-term memory: capacity, associativity, diversity, change over time and unified memory experience.

References and Notes

1 G.S. Bahr and S.L. Wood, "The Big Data between Your Ears: Human Inspired Heuristics for Forgetting in Databases," *IEEE International Conference on Multimedia and Expo Workshops* (2015) pp. 1–6.

2 The name "Ardemia" is pronounced (Ar-Dee-Me-Ah) and an approximation of the phonetic spelling of the acronym RDMA (Relational Data Memory Architecture). The data model and implementation details can be found at www.artificialmemory.org or www.gsbahr.com.

3 J.W. Dauben, *Georg Cantor* (Princeton Univ. Press, 1970) p. 170, translated by G.S. Bahr.

4 E.F. Codd, "A Relational Model of Data for Large Shared Data Banks," *Communications of the ACM* **13**, No. 6, 377–387 (1970).

5 The simultaneity of events is an inherently subjective experience and based on point of view. This observation is in line with Einstein's special relativity (A. Einstein, "Zur Elektrodynamik bewegter Körper," *Annalen der Physik* **322**, No. 10, 891–921 [1905]), as discussed in P.J. Nahin, *Time Machines: Time Travel in Physics, Metaphysics, and Sci. Fiction* (American Institute of Physics, 1993). As a result of subjective simultaneity, every memory formed by association through TimeGlue is unique and compatible only with its owner and his or her viewpoint.

6 F.C. Bartlett, *Remembering: A Study in Experimental and Social Psychology, 2nd Ed.* (Cambridge Univ. Press, 1995 [1932]).

7 M. Minsky, "A Framework for Representing Knowledge," P. Winston, ed., *The Psychology of Computer Vision* (New York: McGraw-Hill, 1975) pp. 211–277.

8 S.P. Marshall, *Schemas in Problem Solving* (Cambridge Univ. Press, 1995). See also Robert J. Sternberg, *Textbook Cognitive Psychology*,

6th Ed. (Wadsworth, 2011) Chapter 8; Robert W. Weisberg and Lauretta M. Reeves, *Cognition: From Memory to Creativity*, 1st Ed. (Wiley, 2013).

9 The query code is available in text format at artificialmemory.org or www.gsbahr.com.

10 R.C. Schank, *Tell Me a Story: A New Look at Real and Artificial Memory* (Scribner, 1991).

11 Minsky [7].

12 J.L. McClelland, "Retrieving General and Specific Information from Stored Knowledge of Specifics," in *Proceedings of the Third Annual Conference of the Cognitive Science Society* (1981) pp. 170–172.

Manuscript received 4 February 2016.

GISELA SUSANNE BAHR *is a computer scientist and experimental cognitive psychologist by training (PhD Florida Institute of Technology; PhD Texas Christian University) and an artist by nature. When she is not computing over data and modeling memory, she is busy writing science fiction stories about artificial memory and composing music.*

WILLIAM H. ALLEN *is an associate professor of computer science at the Florida Institute of Technology. His research interests include computer and network security, useable security, modeling and simulation of network-based attacks, and computer forensics. He earned his PhD in computer science from the University of Central Florida.*

PHILLIP J. BERNHARD, *PhD, is currently chair of the Department of Computer Engineering and Sciences at Florida Tech. His research interests include database systems and theoretical computer science.*

STEPHEN WOOD *is an ocean engineer and robotics expert who is committed to cleaning up our oceans and finding solutions to the devasting effect of global warming on our seas. He is program chair and associate professor of ocean engineering at Florida Tech and writes action-, science- and engineering-packed STEM novels for teenagers.*