Providing embedded artificial intelligence with a capacity for palimpsest memory storage
14 July 2022, by Thamarasee Jeewandara

Demonstration of a memristive synapse. (A) Microscope images of fabricated RRAM samples. Individual memristive synapses are highlighted (see black dashed squares) in the intersections of their corresponding conductive electrodes. (B) Schematic operation of a volatile memristor. Plasticity events cause a pronounced and rapid change in synaptic efficacy observed on a fast time scale (STM) and a smaller but more stable change on a slow time scale (LTM). The change in synaptic efficacy over time is shown in (B) following the memory pattern applied in (C). (D) Schematic equivalence of the three consolidation stages in (B) with the analogous beaker theory. The temporal evolution of the first beaker’s liquid level is determined by the hidden state of the second beaker. The liquid levels of the first beaker at stages 1 and 3 are deliberately placed at disproportionate distances from the binary threshold to reflect the asymmetric response of the particular memristive synapse to potentiation and depression events. Credit: Science Advances (2022). DOI: 10.1126/sciadv.abn7920

Biological palimpsest consolidation occurs via hidden biochemical processes that govern synaptic efficacy at varying lifetimes. The arrangement can facilitate idle memories to be overwritten without forgetting them, while using previously unseen memories short-term. Embedded artificial intelligence can significantly benefit from such functionality; however, the hardware has yet to be demonstrated in practice.

In a new report, now published in Science Advances, Christos Giotis and a team of scientists in Electronics and Computer Science at the University of Southampton and the University of Edinburgh, U.K., showed how the intrinsic properties of metal-oxide volatile memristors mimicked the process of biological palimpsest consolidation.

Memristors are devices that can regulate the flow of electric current in a circuit to remember the charge flowing through it. Without implementing special instructions, the experimental memristor synapses displayed expanded doubled capacity while protecting a consolidated memory, as up to hundreds of uncorrelated short-term memories temporarily overwrote it. The outcomes highlighted how the emerging memory technologies can effectively expand the capacity of artificial intelligence hardware to conceptualize learning memory.

Biological intelligence vs. artificial intelligence (AI)

Biological synapses are known to store multiple memories on top of each other at different time scales, much like representations of the early techniques of manuscript writing known as palimpsest, where annotations can be superimposed alongside traces of earlier writing.

Neural networks in the cerebral cortex employ an estimated $10^{13}$ to $10^{14}$ synapses to generate a range of cognitive abilities, their re-engineered counterparts require an equal number of trainable parameters for far narrower applications.
To explain this learning capacity difference between biological intelligence vs. artificial intelligence, AI researchers suggest that synapses can consolidate multiple memories to be revealed at different time scales, much like a palimpsest. While synapses can remember long-term plasticity events, they can express altered states in the short-term. As a result, the brain can use the same resource for a range of computational processes. This flexibility can offer neuromorphic hardware a major milestone to integrate AI across a broad range of on-the-edge, continuously on learning systems.

Coexistence of STM and LTM in memristive synapses. (A) A set of six memristive synapses representing weights $w_n, n \in \{0,1,\ldots,5\}$ is fed plasticity instructions $b_n$ as per patterns M1 and M2. Orange and purple stand for potentiation and depression, respectively. (B) Evolution of the memory trace overlap with M1 under the course of the 18 write events. One hundred percent overlap translates to perfect storage of M1, and since M1 and M2 are binary and mutually orthogonal, 0% overlap translates to perfect storage of M2. (C) Probability of recalling at least x% of M1, where x $\in \{50\%, 80\%, 100\%\}$. Credit: Science Advances (2022). DOI: 10.1126/sciadv.abn7920

During the experiments, researchers had previously designed synapses that are largely based on phase change memory materials, and resistive random-access memory (RRAM)-based memristors to implement metaplasticity to tune the learning rate of artificial synapses in neural networks. Giotis and the team built this study based on previous work, to bridge the synaptic plasticity with automatic consolidation and memory protection of learned memories from synaptic modifications—a vital element for efficient online learning. The team explored the characteristics of RRAM volatility to mimic the hidden biochemical processes facilitating palimpsest consolidation in biological intelligence. They realized two consolidated time-scales in a single device to create a technology that can protect a strong memory in its long-term storage with unique characteristics, without special biases or functional complexity.

**Candidate volatile memories for palimpsest consolidation**

The process of biological palimpsest consolidation assumes that hidden variables such as complex biochemical processes can induce changes in synaptic efficacies across different time scales. While these processes remain to be mapped, their phenotypic response can be modeled via fluid diffusion. The researchers used titanium dioxide-based volatile memristive devices to examine the potential of palimpsest consolidation and demonstrated the process on a single memristive synapse. The outcomes indicated how the artificial synapse protected a hidden memory while reserving the flexibility to express another memory atop it to double the memory capacity. Due to the high ratio of volatile and nonvolatile plasticity changes, which provided a key functional parameter of the system, the product provided a good candidate.
Memory performance under continuous stimulation. (A) Evolution of a LTM and following STM signals. Each randomly chosen STM briefly surpasses LTM before the latter reinstates its dominance. Decaying signals retain higher than chance overlaps (they remain above the 50% noise floor). (B) Observation of all random STMs after LTM consolidation. Red regions indicate that the total time STMx signal is stronger than LTM; blue regions suggest the opposite. (C) Histogram of distribution of all STM total lifetimes on the left y axis. The corresponding probability density function (PDF) and cumulative distribution function (CDF) curves are illustrated via the right y axis. (D) Weighted expected STM lifetime for different LTM consolidation strengths “s.” The green s = 2 line corresponds to the memory data presented in (A) to (C). Credit: Science Advances (2022). DOI: 10.1126/sciadv.abn7920

Operating the memory system

Giotis and the team constructed a memristive network composed of six synapses to strengthen, or consolidate, two competing signals. They illustrated the experimental setup where each bit of memories was written on a corresponding memristive synapse and studied how the two memory signals interacted across the synapses. The synapses cycled through some consolidation stages and the scientists continued to macroscopically examine the memory performance. They credited the progressively more successful presentation of a specific memory to its capacity to overwrite a competing memory with the impact of noise as a deciding factor.

They conducted additional experiments and simulations to quantify how noise affected memory performance, and the outcomes showed that the characteristics of the device drove palimpsest consolidation by enabling reversibility from short-term memory to long-term memory states. The network also showed the capacity in familiarity recalls to identify if multiple memories were previously presented. They characterized the lifetime of short-term memories as the period where the short-term memory signal dominated the long-term memory. The work sheds light on the trade-off between capacity and recollection accuracy, alongside metaplastic properties specifically realized via the synapses. The researchers showed how high levels of consolidation appeared to safeguard long-term memory against hundreds of incoming short-term memories.

Consolidation of binary images. (A) Operation timeline: Images are written in memory every 10 s. Random noisy representations of a LTM (NLTM and LTM, respectively) are written three times before two images are written, once each, in the short term (ST1,2). (B) Original signals of image memories. An NLTM depiction is chosen at random. (C) Temporal snapshots of memory states. The initial snapshot corresponds to last observation before ST1 is seen. Time is referenced after that instant. The moments when the two short-term signals are written were noted. In both cases, a noisy trace of each short-term signal can be accessed in the memory immediately after observation. Then, the short-term signal gradually degrades, while LTM is reinstated in memory. (D) Outline of the LTM signal overlap, as it evolves following the timeline in (A). The initial overlap reflects the corresponding noise level. The experiment is repeated for different NLTM noise levels. Even if a noiseless version of LTM is never presented to the system per se, the memory that achieves above the baseline overlaps vis-à-vis LTM after consolidation has finished. This high overlap is retained even after the ST1 and ST2 are seen by the network. Credit: Science Advances (2022). DOI: 10.1126/sciadv.abn7920

Visual working memory
The team further explored the visual memory based on concepts of short-term attention and unsupervised memory construction in the system. During the experiments, they drew inspiration from existing theories of working memory to construct a vision network with short-term attention complementary to its memory capacity. They also studied if the system could identify statistical significance without supervision. When they compared the long-term memory trace with snapshots that followed, they noted the network could automatically denoise consolidated signals.

**Outlook**

In this way, Christos Giotis and colleagues focused on binary synapses known to support learning in mathematical models and deep-learning algorithms. The interplay between intense bidirectional volatility and small non-volatile residues contributed to the palimpsest capability of the device. Future implementations of the research are not limited to the selected material of titanium dioxide technology and can be studied according to application-specific needs. The setup demonstrated transitions from long-term memory to short-term memory, while also resembling short-term attention mechanisms with promise towards more complex AI algorithms. The dual temporal capacity of the devices resembles the bistable switching known to govern synaptic plasticity. The team equate the functionality of core plasticity to the biological process of calcium/calmodulin-dependent protein kinase II mechanism; a primary biological molecular memory mechanism.


Marcus K Benna et al, Computational principles of synaptic memory consolidation, *Nature Neuroscience* (2016). **DOI: 10.1038/nn.4401**

© 2022 Science X Network