SHALLOW LEARNING IN MATERIO.

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce Shallow Learning In Materio (SLIM) as a resource-efficient method to realize closed-loop higher-order perceptrons. Our SLIM method provides a rebuttal to the Minsky school's disputes with the Rosenblatt school about the efficacy of learning representations in shallow perceptrons. As a proof-of-concept, here we devise a physically-scalable realization of the parity function. Our findings are relevant to artificial intelligence engineers, as well as neuroscientists and biologists.

1 INTRODUCTION

How do we best learn representations? We do not yet fully understand how cognition is manifested in any brain, not even in those of a worm (Rankin, 2004). It is an open question if the shallow brain of a worm is capable of working memory, but if it were then it certainly must depart from the mechanistic models of large-scale brains (Eliasmith et al., 2012). Nevertheless, worm-brain inspired learning combined with "scalable" deep learning architectures have been employed in self-driving cars (Lechner et al., 2020). At present, by scalable we refer to TPU-based architectures (Jouppi et al., 2017; Rumelhart et al., 1986) trained by gradient-descent. However, one could envision a super-scalable future that is less synthetic and based on self-organized nanomaterial systems (Bose et al., 2015; Chen et al., 2020; Mirigliano et al., 2021) that natively realize higher-order (Lawrence, 2022a) and recurrent neural networks. In this short communication, we shall lay yet another brick towards such a future by providing theoretical arguments.

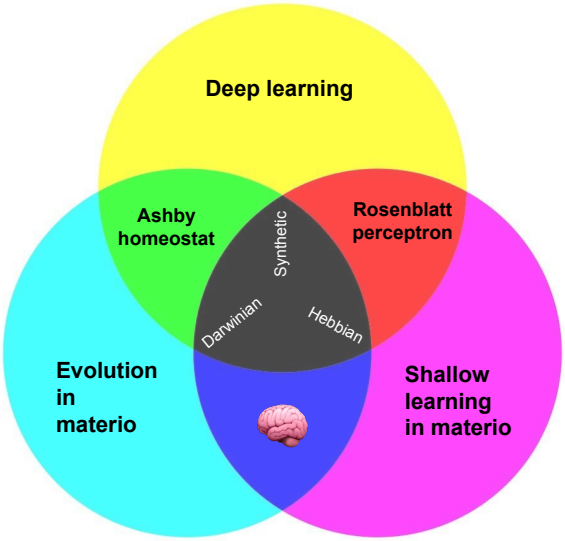

Figure 1: A typology of cognitive material systems.

Our perspective on cognitive material systems is illuminated in Figure 1. Deep learning owes its success to our technological capacity to synthesize massively-parallel and programmable electronic circuits. It is yet to fully exploit Darwinian and Hebbian learning methods that pioneers of the cybernetics movement experimented with by training homeostats (Ashby, 1952) and perceptrons (Rosenblatt, 1961). The spirit of Darwinian (Stanley et al., 2019) and Hebbian (Scellier & Bengio, 2017) learning continues to be alive, though. Here, we add fuel to that fire by advocating for an in-materio approach.

Employing physical systems in their native form for solving computational tasks had gained attention due to the efforts of the 'evolution in materio' community (Miller & Downing, 2002). The earliest result was by Pask (1960) who grew dendritic metallic threads in a ferrous sulphate solution to function as a sound-frequency discriminator (which he called an ear, quite romantically). Now, more recent efforts are under the banner of physical reservoir computing (Tanaka et al., 2019) for realizing sequential functionality. Here, we will commit to combinational functionality by equilibrium-point logic (Lawrence, 2022b) in material systems realizing closed-loop higher-order perceptrons.

2 THEORY

Perceptrons were developed by Rosenblatt and his team, and were trained by a Hebbian learning rule (error-controlled reinforcement) with proven guarantees for convergence. Unfortunately, they started recieving a bad rap after Minsky & Papert (1988) published a proof that 2^N association neurons are required to learn the N-bit parity function. However, this analysis is only applicable if all neurons are threshold logic gates, what Rosenblatt called simple units. Physical neural networks, on the other hand, can natively realize complex units.

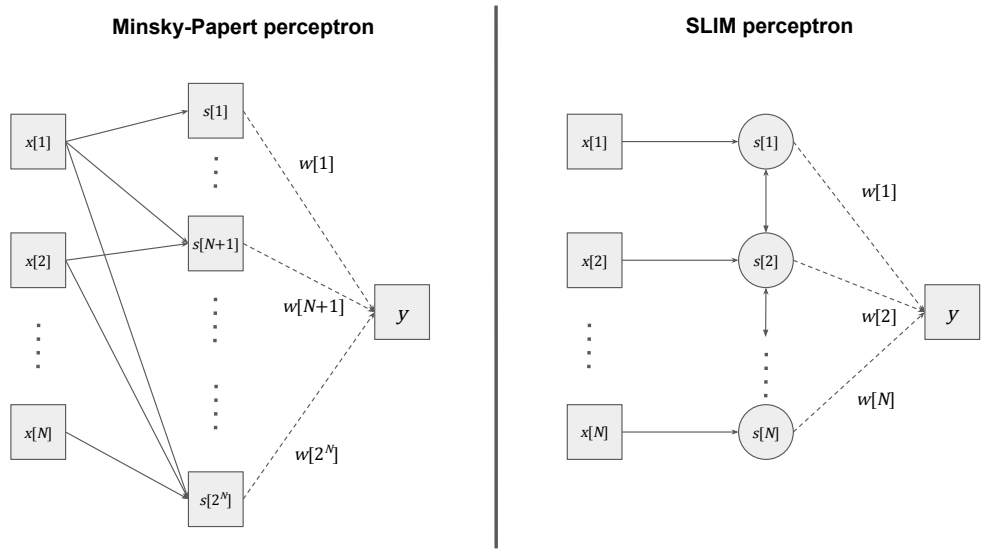

Figure 2: To learn the N-bit parity function, the required number of synaptic weights (depicted as dashed lines) for our SLIM perceptron scales as N instead of the 2^N required for the Minsky-Papert perceptron. This gain in resource-efficiency is possible because the hidden states $s_{1:N}$ of the SLIM perceptron can compute deep-feedforward functionality by equilibrium-point control.

We introduce a shallow learning in materio (SLIM) perceptron as depicted in Figure 2. For our proof-of-concept, we commit to a minimally connected recurrent network with physical states s_i from $i = 1 : N$, yielding a state-space model of the form

$$\dot{s}_i = x_i + F_i(s_{i-1}, s_i, s_{i+1}), \tag{1}$$

where F_i is a nonlinear function. We conjecture that all possible N-bit functions may be realized if arbitrary choices of $F_{1:N}$ are allowed. Imposing conditions of physical realizability on F_i would

affect the neuronal capacity (Baldi & Vershynin, 2018) of our SLIM perceptron, and at present there is no engineering theory to design an optimal F_i (even when $N = 2$). As a step to that, we demonstrate equilibrium-point logic in Figure 3 with $F_{1:2}$ designed as piecewise-constant functions. A promising approach to obtain F_i for higher dimensions is to identify an analogy with cellular automaton that are capable of equilibrium-point parity logic in arbitrary dimensions (Betel et al., 2013). Obtaining scaling laws for the volume of state-space in action during equilibrium-point logic may be another worthy problem to ponder upon.

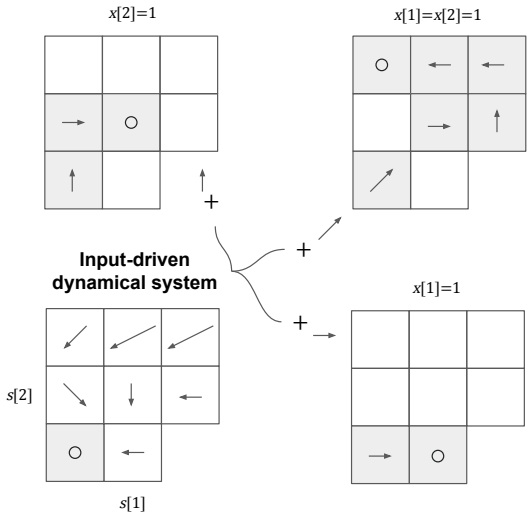

Figure 3: Equilibrium-point logic in a piecewise-constant vector field $\dot{s}_i = x_i + F_i(s_1, s_2)$ designed to realize $s_1 = x_1 \oplus x_2$. The field for $x_1 = x_2 = 0$ is depicted in the bottom-left, with regions that could be visited and the equilibrium-region (in gray). On addition of input(s), the system traverses to a new equilibrium-point as shown for the three possible cases.

3 Conclusion

Our contribution here is threefold: (1) a typology of cognitive material systems that puts a spotlight on yet-to-be-appreciated alternatives to deep learning, (2) arguments to support the investigation of recurrent networks for deep feedforward functionality, (3) framing open problems in equilibrium-point logic. More theory is needed to develop constructive high-dimensional examples, and a statistical analysis of SLIM's performance. Next steps could be to obtain estimates on the learning duration, and check if it is superior to estimates obtained from the principal convergence theorem for perceptrons (Rosenblatt, 1961, Section 5.5, Theorem 4). Given the well established (in silico) deep learning industry, even with a more practical demonstration, business economics would prevent the shift to a SLIM paradigm in the near future. Nevertheless, the SLIM concept may act as a catalyst for gifted mathematicians to make new connections or help neuroscientists in unravelling the mysteries of small-scale brains.

References

W Ross Ashby. *Design for a brain.* Wiley, 1952.

Pierre Baldi and Roman Vershynin. On neuronal capacity. *Advances in Neural Information Processing Systems*, 31, 2018.

Heather Betel, Pedro PB de Oliveira, and Paola Flocchini. Solving the parity problem in one-dimensional cellular automata. *Natural Computing*, 12(3):323–337, 2013.

S. K. Bose, C. P. Lawrence, Z. Liu, K. S. Makarenko, R. M. J. van Damme, H. J. Broersma, and W. G. van der Wiel. Evolution of a designless nanoparticle network into reconfigurable boolean logic. *Nature Nanotechnology*, 10(12):1048–1052, Dec 2015. ISSN 1748-3387. doi: 10.1038/nnano.2015.207. URL http://www.nature.com/articles/nnano.2015.207.

Tao Chen, Jeroen van Gelder, Bram van de Ven, Sergey V. Amitonov, Bram de Wilde, Hans-Christian Ruiz Euler, Hajo Broersma, Peter A. Bobbert, Floris A. Zwanenburg, and Wilfred G. van der Wiel. Classification with a disordered dopant-atom network in silicon. *Nature*, 577 (7790):341–345, Jan 2020. ISSN 0028-0836. doi: 10.1038/s41586-019-1901-0. URL http://www.nature.com/articles/s41586-019-1901-0.

Chris Eliasmith, Terrence C Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *science*, 338(6111):1202–1205, 2012.

Norman P Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, et al. In-datacenter performance analysis of a tensor processing unit. In *Proceedings of the 44th annual international symposium on computer architecture*, pp. 1–12, 2017.

Celestine P. Lawrence. Compact modeling of nanocluster functionality as a higher-order neuron. *IEEE Transactions on Electron Devices*, 69(9):5373–5376, 2022a. doi: 10.1109/TED.2022.3191956.

Celestine P. Lawrence. Deep feedforward functionality by equilibrium-point control in a shallow recurrent network. Technical report, 2022b. URL https://openreview.net/forum?id=401LFvBGIb.

Mathias Lechner, Ramin Hasani, Alexander Amini, Thomas A Henzinger, Daniela Rus, and Radu Grosu. Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, 2(10): 642–652, 2020.

Julian F Miller and Keith Downing. Evolution in materio: Looking beyond the silicon box. In *Proceedings 2002 NASA/DoD Conference on Evolvable Hardware*, pp. 167–176. IEEE, 2002.

Marvin L Minsky and Seymour A Papert. Perceptrons: expanded edition, 1988.

Matteo Mirigliano, Bruno Paroli, Gianluca Martini, Marco Fedrizzi, Andrea Falqui, Alberto Casu, and Paolo Milani. A binary classifier based on a reconfigurable dense network of metallic nano-junctions. *Neuromorphic Computing and Engineering*, 1(2):024007, 2021.

Gordon Pask. The natural history of networks. In *Self-organizing systems (Proceedings of an interdisciplinary conference. 5-6 May 1959)*, pp. 262. Pergamon press, 1960.

Catharine H Rankin. Invertebrate learning: what can't a worm learn? *Current biology*, 14(15): R617–R618, 2004.

Frank Rosenblatt. *Principles of neurodynamics. perceptrons and the theory of brain mechanisms*. Cornell Aeronautical Laboratory, Inc. Report no. 1196-G-8, 1961.

David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

Benjamin Scellier and Yoshua Bengio. Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience*, 11:24, 2017.

Kenneth O Stanley, Jeff Clune, Joel Lehman, and Risto Miikkulainen. Designing neural networks through neuroevolution. *Nature Machine Intelligence*, 1(1):24–35, 2019.

G. Tanaka, T. Yamane, J.B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose. Recent advances in physical reservoir computing: A review. *Neural Networks*, 115:100–123, 2019.