The distributed neocortex: How neuroscience can inspire distributed AI systems

Mia-Katrin Kvalsund[1], Kai Olav Ellefsen[2], Kyrre Glette[2,3], Sidney Pontes-Filho[4] and Mikkel Elle Lepperød[1,4]

[1]Department of Physics, University of Oslo, Norway
[2]Department of Informatics, University of Oslo, Norway
[3]RITMO, University of Oslo, Norway
[4]Department of Numerical Analysis and Scientific Computing, Simula Research Laboratory, Norway
m.k.o.kvalsund@fys.uio.no

Distributed Dynamical Systems and Neural Cellular Automata (NCA) have long been inspired by natural phenomena. Drawing inspiration from morphogenesis [1], metamorphosis [2], and gene regulatory networks [3], among others, the inspired works have showcased the abilities of distributed dynamical systems. In addition, distributed systems of homogeneous computational units are suggested to be quite useful in robot control: In modular robots that reused computational units across the body in locomotion tasks, researchers have suggested that such robot control can lead to robustness to new environments [4], robustness to changes in the body during co-optimization of body and control [5], and an ability to scale up or down the number of modules without further optimizing the controllers [6; 7]. Additionally, as Artificial Intelligence (AI) in the past years has been growing in size and therefore monetary and environmental cost [8], the parameter reduction inherent in homogeneous distributed systems of computational units is becoming very attractive. However, the road ahead from locomotion to more complex tasks might not be immediately apparent, because the control for these mechanically dependent bodies are inspired by natural phenomena or decentralized swarm systems that do not have to function as a whole. How do we expand on distributed systems to come closer to animal intelligence? How much of an intelligent animal's abilities can be achieved with distributed control and processing? Can a distributed system have a unified sense of self?

Many important breakthroughs in AI have been inspired by neuroscience. From the neuron inspiring the perceptron [9], to the visual pathway inspiring convolutional neural networks [10], research on the brain has been integral to AI. However, as the two fields got more specialized, fewer researchers were able to keep a foot in each discipline, and the fields got separated. Recently, many prominent AI researchers put their name to a white paper that suggested that artificial general intelligence could not be reached without a return to being inspired by neuroscience [11]. The emerging field of NeuroAI hopes to reunite neuroscience and AI, to let these two disciplines drive each other's development. It therefore seems apt to consider what role Distributed Dynamical Systems and NCAs can play in NeuroAI.

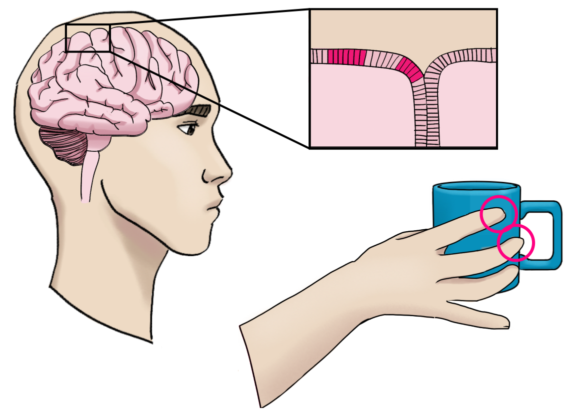

Figure 1: **A simplified explanation.** The neocortex consists of columns of neurons. Certain columns will respond to certain sensors. Here, some columns in the somatosensory cortex receive the fingertips' perceptions of the cup.

The brain is often theorized to be wholly or partially distributed. For example, the PDP movement from the 80s argued for seeing the brain as a distributed system of neurons that self-organizes to perform higher order functions [12]. Recently, Jeff Hawkins popularized the idea of the distributed neocortex in his popular science book "A Thousand Brains" [13]. This was a culmination of his scientific work on the minicolumn hypothesis that arose with Mountcastle in the 70s [14]. The minicolumn hypothesis posits that the neocortex, which is the thin outer layer of the brain, consists of fairly homogeneous columns stacked side-by-side (see Figure 1). The columns only receive and process limited parts of the total sensor input of the animal, and are also involved in the movement of sensors. The columns mostly communicate locally with other columns. The distributed structure of the neocortex is involved in, and might underlie, most higher-order thinking, including language, vision, and problem solving [15] – as well as possibly consciousness [16].

Viewing distributed systems of computational units in

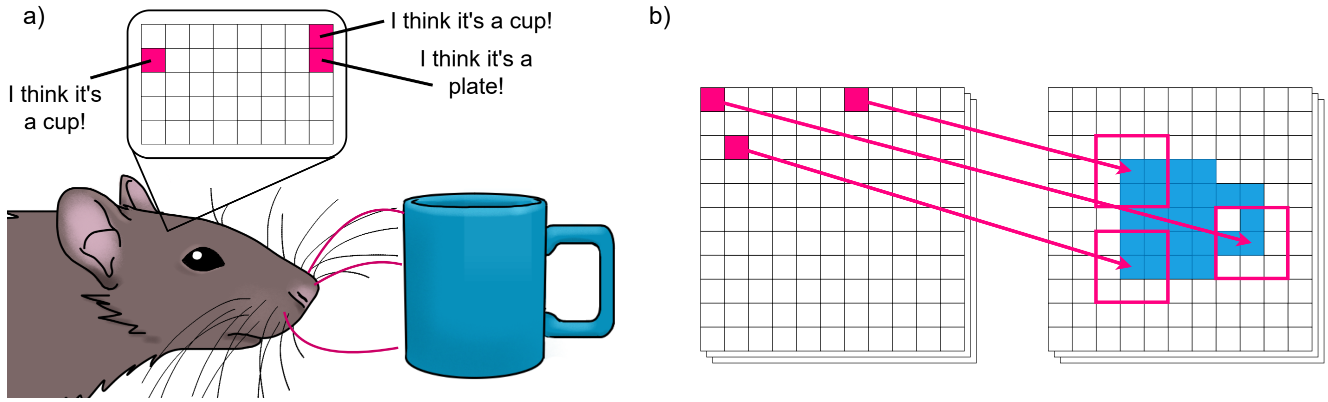

Figure 2: **The project idea.** a) One theory about the function of the cortical columns is that they work together to classify objects. As the rat moves its whiskers across the cup, the corresponding columns in the barrel cortex are activated, and predict what they are sensing. b) Left: The artificial cortex. Right: The sensor input. Just like the whiskers of the rat, each cell can move to receive independent neighborhoods of sensor input.

light of the minicolumn hypothesis provides inspiration for how these systems could be extended to test more of their capabilities in a mechanically dependent body. This gives us a framework for shaping the system, in terms of how computational units should communicate, and how/if the units should be integrated. It also lets us expect what the system should be capable of, judging by what tasks the neocortex is usually involved in. As AI is encouraged to move towards more embodied systems, the minicolumn hypothesis provides a good inspiration for controlling and problem-solving in bodies with a distributed system.

Additionally, modeling the neocortex through distributed systems can inform neuroscience through normative modeling. Normative modeling is the practice of modeling circuits in the brain with optimized artificial neural networks [17; 18]. These normative models are then used to further gain understanding and theorize about the brain circuit. This could potentially help research on the minicolumn hypothesis, and help illuminate how it can function in a distributed manner, while still providing a unified experience.

Proposed future work

As an example of how the minicolumn hypothesis can inform work on distributed systems of computational units, herein NCAs, we present our ongoing project. It builds on the work of Randazzo et al. [19], where an NCA was optimized to classify the MNIST dataset and adapt to perturbations in the input. Likewise, we will work on an image classification task, with the goal of later expanding the system to robot control.

In the work of Randazzo et al., the NCA might be sensitive to elongation of numbers, because the cells get flooded with information that is not representative of the digit. For example, parts of a four resemble a one, but humans know the important parts of a four are not the straight lines. We propose that the system's resilience to such distortions increases if the system only focuses on the relevant sensor input. Such resilience to unimportant information might also later prove to be useful in robots with many sensor inputs.

To construct such a system, we were inspired by the barrel cortex in the neocortex in rodents. Each barrel column in the cortex is directly connected to the independently moving whisker hairs and is involved in their whisking movement [20] (see Figure 2a). We theorize that one benefit of this is that when presented with a lot of similar, and seemingly useless, sensor input, the rodent is able to focus its attention on interesting features in the task space. In this way, a system modeled on the barrel cortex might obtain the ability to ignore irrelevant information in a classification task.

In the system of Randazzo et al., the NCA is applied iteratively to a 20-dimensional image containing the grayscale MNIST datapoint in one channel, 9 channels for communication between cells (called hidden channels), and 10 channels for outputting the believed value of the digit [19]. We will be extending this system with one major alteration: The NCA will be able to choose which part of the image and corresponding hidden channels it sees by an action output for each cell in the image. However, the cell will not be able to change its location in the communication substrate (the artificial barrel cortex, see the left matrix in Figure 2b). This means that just like in the barrel cortex, communication between neighboring whiskers is facilitated, while the actual movement of the whisker, and what it senses, is kept more independent. This ensures information can always flow from every cell to every other cell, given enough time, even as the whiskers focus on different parts of the image. Because this does not enable us to use gradient descent like Randazzo et al., we choose instead to use an evolutionary strategy to optimize the system [21].

References

[1] A. Mordvintsev, E. Randazzo, E. Niklasson, and M. Levin, "Growing neural cellular automata," *Distill*, vol. 5, no. 2, p. e23, 2020.

[2] M. Joachimczak, R. Suzuki, and T. Arita, "Artificial metamorphosis: Evolutionary design of transforming, soft-bodied robots," *Artificial life*, vol. 22, no. 3, pp. 271–298, 2016.

[3] M. Joachimczak, R. Suzuki, and T. Arita, "Fine grained artificial development for body-controller coevolution of soft-bodied animats," in *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pp. 239–246, MIT Press, 2014.

[4] G. Nadizar, E. Medvet, S. Nichele, and S. Pontes-Filho, "Collective control of modular soft robots via embodied spiking neural cellular automata," *arXiv preprint arXiv:2204.02099*, 2022.

[5] M.-K. Kvalsund, K. Glette, and F. Veenstra, "Centralized and decentralized control in modular robots and their effect on morphology," in *ALIFE 2022: The 2022 Conference on Artificial Life*, MIT Press, 2022.

[6] D. Pathak, C. Lu, T. Darrell, P. Isola, and A. A. Efros, "Learning to control self-assembling morphologies: a study of generalization via modularity," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[7] D. J. Christensen, "Evolution of shape-changing and self-repairing control for the atron self-reconfigurable robot," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pp. 2539–2545, IEEE, 2006.

[8] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, "The computational limits of deep learning," *arXiv preprint arXiv:2007.05558*, 2020.

[9] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[10] D. L. Yamins and J. J. DiCarlo, "Using goal-driven deep learning models to understand sensory cortex," *Nature neuroscience*, vol. 19, no. 3, pp. 356–365, 2016.

[11] A. Zador, S. Escola, B. Richards, B. Ölveczky, Y. Bengio, K. Boahen, M. Botvinick, D. Chklovskii, A. Churchland, C. Clopath, *et al.*, "Catalyzing next-generation artificial intelligence through neuroai," *Nature Communications*, vol. 14, no. 1, p. 1597, 2023.

[12] D. E. Rumelhart, J. L. McClelland, P. R. Group, *et al.*, *Parallel distributed processing*, vol. 1. IEEE New York, 1988.

[13] J. Hawkins, *A thousand brains: A new theory of intelligence*. Hachette UK, 2021.

[14] V. Mountcastle, "An organizing principle for cerebral function: the unit module and the distributed system," *The mindful brain*, 1978.

[15] E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum, and A. Hudspeth, *Principles of neural science*, vol. 5. McGraw-hill Medical, 2000.

[16] C. Koch, M. Massimini, M. Boly, and G. Tononi, "Neural correlates of consciousness: progress and problems," *Nature Reviews Neuroscience*, vol. 17, no. 5, pp. 307–321, 2016.

[17] B. A. Richards, T. P. Lillicrap, P. Beaudoin, Y. Bengio, R. Bogacz, A. Christensen, C. Clopath, R. P. Costa, A. de Berker, S. Ganguli, *et al.*, "A deep learning framework for neuroscience," *Nature neuroscience*, vol. 22, no. 11, pp. 1761–1770, 2019.

[18] A. Saxe, S. Nelli, and C. Summerfield, "If deep learning is the answer, what is the question?," *Nature Reviews Neuroscience*, vol. 22, no. 1, pp. 55–67, 2021.

[19] E. Randazzo, A. Mordvintsev, E. Niklasson, M. Levin, and S. Greydanus, "Self-classifying mnist digits," *Distill*, vol. 5, no. 8, pp. e00027–002, 2020.

[20] C. C. Petersen, "Sensorimotor processing in the rodent barrel cortex," *Nature Reviews Neuroscience*, vol. 20, no. 9, pp. 533–546, 2019.

[21] N. Hansen and A. Ostermeier, "Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation," in *Proceedings of IEEE international conference on evolutionary computation*, pp. 312–317, IEEE, 1996.