Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews 29 (2019) 104–105

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Comment

The reasonable ineffectiveness of biological brains in applying the principles of high-dimensional cybernetics
Comment on "The unreasonable effectiveness of small neural ensembles in high-dimensional brain" by Alexander N. Gorban et al.

Cees van Leeuwen

Laboratory for Perceptual Dynamics, Brain & Cognition Research Unit, KU Leuven, Belgium

Received 28 February 2019; accepted 12 March 2019
Available online 15 March 2019
Communicated by L. Perlovsky

Gorban et al. [1] take the reader on an intellectual journey that is as delightful as it is impressive: we are regaled with quotes from, among others, Einstein and Leibnitz, are offered vistas of the cybernetic tradition and stochastic physics, culminating in the blessing of dimensionality, from which we then descend to applications in artificial and biological neural systems, opening into a topic of immediate translational value, i.e. that of correcting legacy AI systems.

For cognitive neuroscientists, there are several wise lessons to take from this paper: most of us are still inclined to follow Simon [2] in understanding cognitive system as consisting of decomposable subsystems, or modules, of which the function can be studied in isolation. The paper reminds us that the brain is not a linear combination of component functions. Adding up components increases system dimensionality, which can have counter-intuitive consequences we need to get our heads around. In this respect, the article offers some intriguing speculations about the parallels between brain evolution and cascades of error-correcting systems.

Moreover, we can see in a new light the traditional model hierarchy, in which initial convergence for processing is followed by divergence for classification. From the visual system, we are familiar with convergence as signals proceed from the retina through visual hierarchy, where receptive fields are becoming increasingly complex and encoding increasingly sparse. We know much less about what happens in the higher echelons of the hierarchy, but we might understand it better with the blessing of dimensionality in mind. "Grandmother cells" have often been understood as the end point of convergence and, as such were considered highly unlikely. Here we can see that they are the product of subsequent divergence, and are, in fact, highly likely.

Why then, in spite of the wisdom espoused by the paper, does it ultimately feel like a rollercoaster ride? Perhaps, this is because the classical cybernetic understanding of the brain (in particular, the alliance of information theory and the neuron doctrine as sourced in [1] through Barlow) has run its course. Whatever its merits for artificial systems, for the brain it offers, at best, a one sided image. Let me give a few examples:

DOI of original article: https://doi.org/10.1016/j.plrev.2018.09.005.
E-mail address: cees.vanleeuwen@kuleuven.be.

https://doi.org/10.1016/j.plrev.2019.03.005
1571-0645/

- The cybernetic understanding of the brain is based on the notion that neurons compute. Such information processing that is strictly local; effective within neurons or small assemblies. Biological neurons, by contrast, operate in an intricate balance of local and nonlocal activity. They signal locally by sending electric spikes or graded potentials, but also nonlocally, as their individual activity is leaky. Those leaky currents operate like waves in an intra and extracellular field, spreading signals in an all-to-all fashion, at different spatiotemporal scales, from the circuit level [3] to the entire cortex [4,5]. Computations may take place in biological systems with nonlocal wave as well as with local spiking activity.
- Neurons typically don't operate in isolation, but in coordinated clusters of various scales. They have to. The signals they send through their chemical synapses are far too numerous to be individually effective, given the limited dynamic range of the neuron. Strong couplings between neurons, e.g. via gap junctions, as well as Hebbian learning, contribute to these effects [6].
- The systems described by Gorban et al. represent the brain as a feedforward hierarchy: first preprocessing, then classification. But in the brain, the predominant connectivity is in the opposite direction. Possibly, rather than in classification, the brain engages in controlled hallucination. It is predominantly talking to itself. Thus the brain is not inert, but is actively generating predictions about what will happen next, and learning by trial and error.
- Memory in cybernetics is inert. For such systems, which take time to train, the prescriptions in [1] for updating by adding components are, indeed, cost-effective. But the brain, unlike legacy software, is constantly updating itself. The hippocampus is an important mediator of memory, but is not a place where memory is permanently stored. Storage is distributed throughout the brain. But even here, memory may be far from permanent. In an interplay with the hippocampus, memories are continuously reconstructed, rehearsed, and re-encoded, possibly in evanescent circuits [7]. Such a process may well be modeled in terms of long transients [8]. Why is brain activity transient? Because coding, in the brain, is context-sensitive. Grandmother cells (or circuits) do not have fixed meanings, as they operate in a context, proximally given by the ongoing activity of the system in which such a cell is embedded, and distally by the environment the system finds itself embedded in. Accordingly, their response will fluctuate with the context. Consequently, a complex cell may be selectively responding to Jenifer Aniston, but only on average across conditions. This may have no meaning in the specific situation the system finds itself in.

Whereas the cybernetic approach has engrained itself in scientists' understanding of the brain, in fact, it is a rather one-sided and misleading metaphor. Isn't it time we, cognitive neuroscientists, take our deserved share of the blame for this situation? Perhaps we have been lazy in developing a theoretical framework that would appeal to people of the stature like the authors of [1], and that can hold its own against the alliance of neural doctrine and information theory, which has engulfed our understanding of the brain.

Acknowledgement

The author is funded by an Odysseus grant (G.0003.12) from the Flemish Ministry of Science.

References

[1] Gorban AN, Makarov VA, Tyukin IY. The unreasonable effectiveness of small neural ensembles in high-dimensional brain. Phys Life Rev 2019;29:55–88. https://doi.org/10.1016/j.plrev.2018.09.005 [in this issue].
[2] Simon HA. The architecture of complexity. Proc Am Philos Soc 1962;106:467–82.
[3] Sato TK, Nauhaus I, Carandini M. Traveling waves in visual cortex. Neuron 2012;75:218–29.
[4] Ito J, Nikolaev AR, van Leeuwen C. Spatial, temporal structure of phase synchronization of spontaneous EEG alpha activity. Biol Cybern 2005;92:54–60.
[5] Alexander DM, Jurica P, Trengove C, et al. Traveling waves and trial averaging, the nature of single-trial and averaged brain responses in large-scale cortical signals. NeuroImage 2013;73:95–112.
[6] Hellrigel S, Jarman N, van Leeuwen C. Adaptive rewiring of weighted networks. Cogn Syst Res 2019;55:205–18.
[7] Routtenberg A. Lifetime memories from persistently supple synapses. Hippocampus 2013;23:202–6.
[8] van Leeuwen C, Raffone A. Coupled nonlinear maps as models of perceptual pattern and memory trace dynamics. Cogn Process 2001;2:67–116.