Towards Information-Theoretic Limits of the Global Neuronal Workspace Architecture

Lav R. Varshney

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
varshney@illinois.edu

Abstract. The global neuronal workspace architecture has been proposed as a biologically plausible computational model for the operation of human consciousness, essentially arguing that signals flow from several perceptual, memory, and attentional regions to a central broadcast medium where certain signals cause a cascade that enters conscious awareness. Separately, the integrated information theory of consciousness has proposed that multivariate information measures Φ that generalize Shannon's mutual information capture the level of interaction among brain regions and therefore measure consciousness. Here, we ask whether these two theories are in fact two sides of the same coin, by suggesting a mathematical theorem on the operational limits of the global neuronal workspace, similar to Shannon's noisy channel coding theorem, that would naturally be in terms of one particular multivariate information measure Φ.

Keywords: global neuronal workspace theory · integrated information theory · coding theorem.

Over the last several years, a few different kinds of mathematically-oriented theories of consciousness have emerged. One style of theory takes an operational view of consciousness, in terms of the flow of signals and the allocation of attention in the brain. The global neuronal workspace (GNW) model [2] argues that signals from perception, long-term memory, and evaluative systems are modulated by attention (conscious access) mechanisms and combined in order to produce actions through the motor system. As part of conscious processing, signals are also multicast back to the various input systems. The GNW model could be implemented through the massive connectivity arising from long-distance cortico-cortical axons, e.g. in prefrontal cortex. Another style of emerging theory is the integrated information theory (IIT) [8] for measuring consciousness. The basic idea is that measuring a multivariate information-theoretic quantity called integration (Φ) would allow assessment of the extent to which information is interconnected into a unified whole rather than split into disconnected parts. Numerous specific multivariate information measures have been proposed as integration, but there is no consensus on which one makes the most sense; a recent paper by Tegmark suggested 420 different possibilities, of which at least

Table 1. Coding Theorems Link the Operational and the Informational.

	Operational	Informational
Channel capacity (Shannon, 1948)	Maximum rate we can send messages over noisy channel and recover with arbitrarily low error probability	$C(B) = \max_{p_X : E[b(X)] \leq B} I(X;Y)$
Consciousness	The optimal information flow possible in the global neuronal workspace architecture under suitable reliability objectives	A multivariate information measure Φ

20 are efficiently computable [7]. The lack of consensus is perhaps since arguments in favor of various Φ measres are axiomatic, rather than based on a specific connection to an operational interpretation of the information measures.

Dehaene et al. [3] have recently stated: "A more modest proposal is that Φ and related quantities provide one of many possible signatures of the state of consciousness, simply because they reflect the brain's capacity to broadcast information in the global neuronal workspace, and therefore to entertain a ceaseless stream of episodes of conscious access and conscious processing." Here we formally take up this modest proposal to show that GNW (an operational definition) and IIT (an informational definition) are very much intertwined through an understanding of the fundamental limits of information flow in the global neuronal workspace. The approach is analogous to how Shannon's noisy channel coding theorem established equivalence between operational notions of reliable communication and mutual information quantities [6], see Table 1 for a depiction of how coding theorems link the operational and the informational.

Interestingly, the multiinformation among random variables $X_1, \ldots, X_n$:

$$I(X_1, \ldots, X_n) = D_{KL}(p_{X_1,\ldots,X_n} \| p_{X_1} \cdots p_{X_n}) = \sum_{i=1}^{n} H(X_i) - H(X_1, \ldots, X_n)$$

and its extension to partitioning, the minimum partition information

$$I_{MP}(X_1, \ldots, X_n) = \min_{P \in \mathcal{P}} \frac{1}{|P| - 1} \sum_{j=1}^{|P|} H(\{X_i : i \in P_j\}) - H(X_1, \ldots, X_n)$$

defined using KL divergence or entropy emerge naturally in coding theorems for the capacity of multiple-access channels [4], secret key distribution in cryptography [1], and optimal algorithms for unsupervised learning (clustering) [5]. We are in process of a similar mathematization of the GNW architecture and information flow, and believe I_{MP} will emerge as the fundamental limit and the correct measure for Φ.

Information-theoretic limits have driven technological development of communication systems for decades; we similarly believe a characterization of conscious processing will prove inspirational for the design of future AI systems.

References

1. Chan, C., Al-Bashabsheh, A., Ebrahimi, J.B., Kaced, T., Liu, T.: Multivariate mutual information inspired by secret-key agreement. Proc. IEEE **103**(10), 1883–1913 (Oct 2015). https://doi.org/10.1109/JPROC.2015.2458316
2. Dehaene, S., Changeux, J.P.: Experimental and theoretical approaches to conscious processing. Neuron **70**(2), 200–227 (Apr 2011). https://doi.org/10.1016/j.neuron.2011.03.018
3. Dehaene, S., Charles, L., King, J.R., Marti, S.: Toward a computational theory of conscious processing. Curr. Opin. Neurobiol. **25**, 76–84 (Apr 2014). https://doi.org/10.1016/j.conb.2013.12.005
4. Liu, Y.S., Hughes, B.L.: A new universal random bound for the multiple-access channel. IEEE Trans. Inf. Theory **42**(2), 376–386 (Mar 1996). https://doi.org/10.1109/18.485710
5. Raman, R.K., Varshney, L.R.: Universal joint image clustering and registration using multivariate information measures. IEEE J. Sel. Topics Signal Process. **12**(4), 928–943 (Oct 2018). https://doi.org/10.1109/JSTSP.2018.2855057
6. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423, 623–656 (July/Oct 1948)
7. Tegmark, M.: Improved measures of integrated information. PLoS Comput. Biol. **12**(11), e1005123 (Nov 2016). https://doi.org/10.1371/journal.pcbi.1005123
8. Tononi, G., Boly, M., Massimini, M., Koch, C.: Integrated information theory: from consciousness to its physical substrate. Nat. Rev. Neurosci. **17**, 450–461 (2016). https://doi.org/10.1038/nrn.2016.44