Computing local active information storage in input-driven systems

Oliver Obst,[1] Joschka Boedecker,[2] Benedikt Schmidt,[3] and Minoru Asada[2]
[1] *CSIRO Information and Communications Technology Centre PO Box 76, Epping, NSW 1710, Australia*
[2] *Graduate School of Engineering, Osaka University, Suita 565-0871, Osaka, Japan*
[3] *Universität Koblenz-Landau, 56070 Koblenz, Germany*
(Dated: July 16, 2012)

Information theory and the framework of information dynamics [4–6] has been used to provide tools to characterise complex systems. In this case, the characteristic elements of computation are information storage, information modification and information transfer. With help of these information-theoretic tools we hope to understand, and to eventually engineer dynamical systems which requires a proper understanding of their computational properties. In contrast to static measurements of, e.g., entropy of a system at a given time, they focus on dynamical aspects of information processing. Mitchell [1], for example, has been quoted [2] suggesting that "the main challenge is understanding the dynamics of the propagation of information ... in networks, and how these networks process such information."

The information dynamics framework cited above is useful to get a "wholistic" understanding of complex systems. In the case of input-driven systems, this means that, using these tools, we look at the entirety of both the system as such, as well as the input to the system throughout the experiment. This is an important observation, as for example for biological systems perform non-trivial computations and also retain a short-term memory of past inputs [3]: there is no distinction between structure of the input into the system and that of the system itself. In some of our work, we have measured information transfer and active information storage in recurrent neural networks to show peak performance near the edge of chaos [7]. Many other real world systems like cortical networks are also heavily input-driven, and application of these tools may not necessarily lead to intuitively interpretable results. This is particularly true if the environment (i.e., the input to the system) changes over time, or if we are interested in properties of individual components of a complex system.

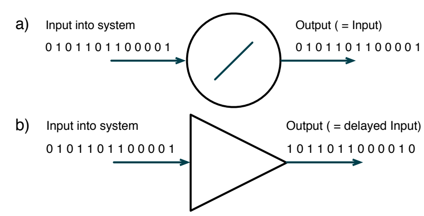

FIG. 1. Simple computational units of artificial neural networks may forward or store information. In a), inputs are just forwarded to the output. b) implements a 1-bit delay line.

Figure 1 illustrates simple cases for quantifying local information storage for input driven systems: We would expect zero local information storage for the unit in a) (since no information is stored in the system), but the computed local active information storage will in fact depend on the particular bias of our input data. Similarly in b), we would expect a local information storage of exactly one bit for a one-bit delay line.

Using the existing framework of information dynamics, we derive a measurement of input-corrected information storage that is suitable for input-driven dynamical systems. Input-corrected information storage quantifies the contribution of the system to information storage alone, i.e., a measure of difference to a pure feedforward unit driven by a particular input.

With input-corrected information storage we hope to better quantify system behaviour, which will be important for heavily input-driven systems like artificial neural networks to abstract from specific benchmarks, or for brain networks where individual components cannot be tested in isolation or with arbitrary input data.

[1] M. Mitchell, Artificial Intelligence **170**, 1194 (2006).
[2] J. T. Lizier, S. Pritam, and M. Prokopenko, Artificial Life **17**, 293 (2011).
[3] S. Ganguli, D. Huh, and H. Sompolinsky, Proceedings of the National Academy of Sciences **105**, 18970 (2008).
[4] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Phys. Rev. E **77**, 026110 (2008).
[5] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Chaos: An Interdisciplinary Journal of Nonlinear Science **20**, 037109 (2010).
[6] J. T. Lizier, *The local information dynamics of distributed computation in complex systems*, Ph.D. thesis, School of Information Technologies, The University of Sydney (2010).
[7] J. Boedecker, O. Obst, J. T. Lizier, N. M. Mayer, and M. Asada, Accepted for publication in Theory In Biosciences (2011).