UNDER REVIEW

RESCIENCE C

Replication / Machine Learning

[Re] Measures for investigating the contextual modulation of information transmission.

Sepehr Mahmoudian[1,2]

[1]Deptartment of Data-driven Analysis of Biological Networks, Campus Institute for Dynamics of Biological Networks, Georg August University Göttingen, Göttingen, Germany – [2]MEG Unit, Brain Imaging Center, Goethe University Frankfurt, Frankfurt, Germany

Edited by
(Editor)

Reviewed by
(Reviewer 1)
(Reviewer 2)

Received
01 November 2018

Published
–

DOI
–

1 Introduction

Background – Systems such as neurons can have different types of input. The authors[1] propose dividing the input to a neuron into receptive (or driving) and contextual. They further present formulations of how contextual connections can amplify driving input. Finally, they use information theory to study their attributes.

Motivation – This work was published in 1996, but with the advances in computing and information theory, and increased emphasis on the importance of context[2], new studies[3,4] have started building on its foundations. Furthermore, there are many ways to write the activation functions and only a few simple formulations were analyzed. Therefore, the significance of this work is twofold: 1. To reproduce the study in a modern programming language and make it publicly accessible. 2. Enable new activation functions to be readily tested.

Reproduction – A successful attempt is made to qualitatively and quantitatively reproduce the results of the original paper by reproducing the three figures that comprise the results.

2 Methods

Input – The study is limited to information transmission from a neuron which has only one receptive (r) or contextual (c) connection as input. There is input at every time step, +1 denotes firing and -1 being silent, which is then multiplied by the weight of their connection. probability distributions for $\mathbf{r}$ and $\mathbf{c}$ were created in the same way as original authors as described below, where for all figures P(C = 1 | R = 1) was set to 0.889972.

$$P(R = +1, C = +1) = 0.5P(C = 1|R = 1) \tag{1}$$

$$P(R = +1, C = -1) = 0.5[1 - P(C = 1|R = 1)] \tag{2}$$

$$P(R = -1, C = +1) = 0.5[1 - P(C = 1|R = 1)] \tag{3}$$

Correspondence should be addressed to Sepehr Mahmoudian (sepehr.mahmoudian@gmail.com)
The authors have declared that no competing interests exists.
Code is available at https://github.com/rescience-c/template.

$$P(R = -1, C = -1) = 0.5P(C = 1|R = 1). \tag{4}$$

Activation Functions — Three activation functions are studied by the authors. Based on how receptive and contextual information are formulated to interact, they are called A_a (additive), A_m (modulatory), and A_b(both additive and modulatory).

$$A_a(r, c) = r + c \tag{5}$$

$$A_m(r, c) = 0.5r[1 + e^{rc}] \tag{6}$$

$$A_b(r, c) = 0.5r[1 + e^{rc}] + c \tag{7}$$

The output of each activation function is passed into the sigmoidal activation function to generate a probability of firing:

$$p(x = 1|r, c) = \frac{1}{1 + e^{-A(r,c)}} \tag{8}$$

Information Theory — The mutual information between two events (the information shared between them) can be expressed as the difference between the information gained by observing one event alone and the information gained by observing the other event given the first:

$$I(X; R) = H(X)−H(X|R) = H(R)−H(R|X) \tag{9}$$

This can be extended to three variables:

$$I(X; R; C) = I(X; R)−I(X; R|C) = \sum_{x,r,c} p(x, r, c)log\frac{P(x|r)p(x|c)}{p(x)p(x|r, c)} \tag{10}$$

The average can be calculated:
The information transmitted by such a system can be split into three components depending on whether it is shared between the two inputs or is specific to each:

$$I(X; R|C) = H(X|C)−H(X|R, C) \tag{11}$$

Dependencies — The reproduction was done using Ubuntu 19.10 with Intel® Core™ i7-7500U CPU with Anaconda Python-3.8.1, numpy-1.18.1, scipy-1.4.1, and matplotlib-3.1.3. There are no external libraries required for calculating the information theoretic terms as everything is provided in the accompanying code.

3 Results

Activation functions and transmitted information — Figure 1 shows the I(X; R; C) and I(X; C |R) for each of the activation functions and across a range of r input strengths while c remains constant at 1. These results are produced, in a similar way to the paper, by . The results are quantitatively similar to the original paper.

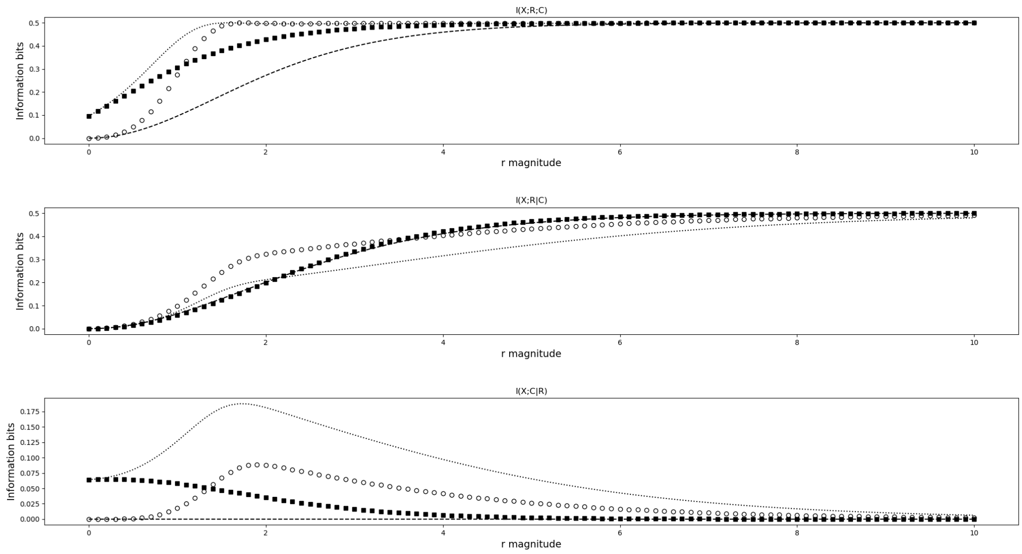

Figure 1. Reproduction of the comparison of the activation functions for three-way and conditional mutual information. The activation functions are zero context (dashed), additive context (squares), modulatory context (circles), and both additive and modulatory combined (dotted).

Figure 2 shows the I(X;R;C) and I(X;C |R) for each of the activation functions and across a range of r input strengths while c remains constant at 1. These results are produced, in a similar way to the paper, by . The results are quantitatively similar to the original paper.

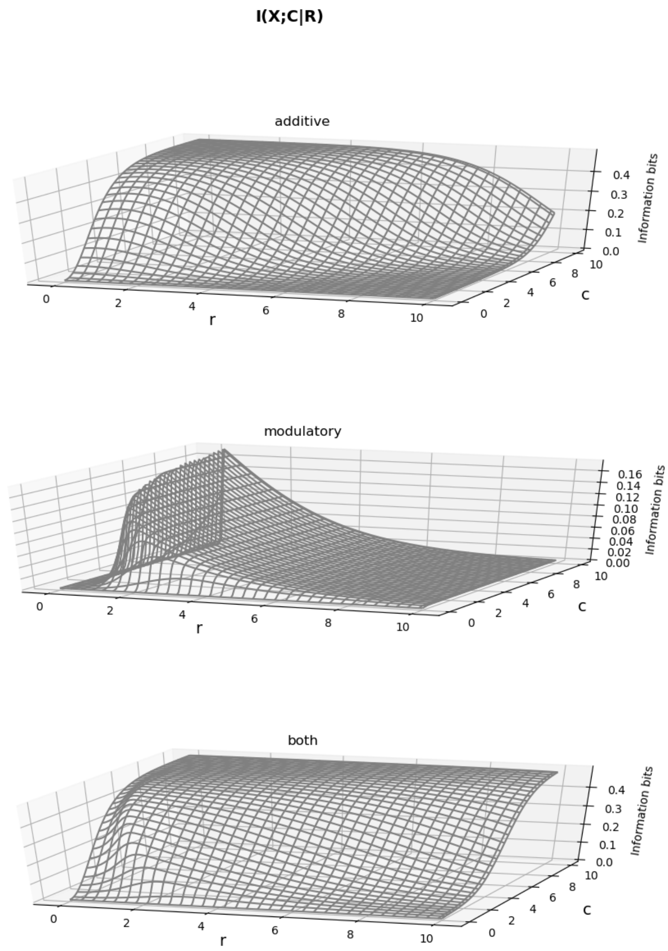

Figure 2. Reproduction of the comparison of the information surfaces of I(X;C |R) for (a) additive, (b) modulatory and (c) both over all RF and CF input strengths. The conditional probability P(C = 1 |R = 1) = 0.889 972. All information measures are analytical.

Sampling biases and variances — Figure 3 compares the mean and standard deviation of the three information measures sampled against the analytical measures for the modulatory activation function. The mean and deviations were computed from 100 trials per sample size and compared with the analytical measure across a range of sample sizes. The sampled and analytical mean values are approximately equivalent for sample sizes above 200 and the standard deviations decrease only very slowly beyond that. This suggests that a sample size of 200 is a good estimate of a lower bound on the sample size to use. The results also show that whereas there is an upward sampling bias on measures of the two-way and three-way mutual information, there is a downward bias on estimates of conditional mutual information. The reason for this is not yet clear.

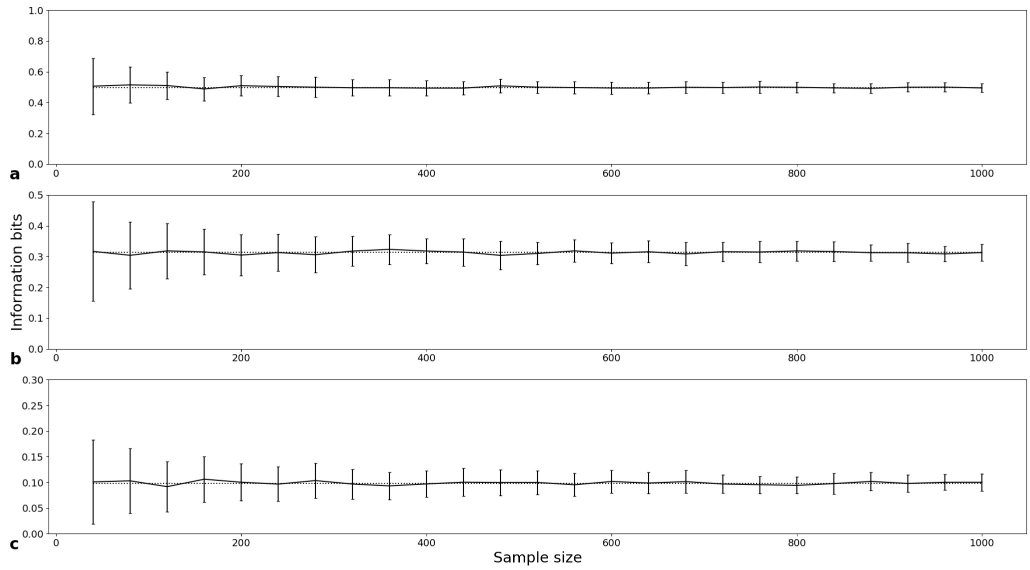

Figure 3. Reproduction of comparison of the analytical (dotted) and sampled (error bars) information measures for (a) I(X;R;C),(b) I(X;R|C) and (c) I(X;C |R). The error bars represent standard deviation about the mean information computed from 100 trials for each sample size. The RF and CF input magnitudes were set at 2.

In this paper we have reported how mutual and conditional information measures between two input sources and the output in a simple system can be used to distinguish various effects of the inputs. In particular, we can say whether or not the effects of some contextual stimuli are additive or modulatory or a combination of both. By comparing sampled against analytical measures of the information we have found that 200 samples in such a binary input system is sufficient to get an unbiased estimate of the information measures. It is valuable to have three measures because they describe the effects of modulation on the different components of the transmitted information. It is also valuable because the different measures have different requirements. For example I(X;R;C) requires some correlation between the two input distributions whereas the conditional measures require that the two inputs are not perfectly correlated. These measures can all be applied to the psychophysical paradigm described in the introduction and results from such experiments will be reported elsewhere.

4 Discussion

The point of the paper was to study contextual amplification and functions of how it could be implemented. Here the main results of their work is reproduced. minor differences in figure 3 at the start which can be attributed to how computers back then handled small numbers.

Acknowledgments — I would like to thank Bill Phillips and Jim Kay for clarifying my questions about the input. This research was supported by the funding initiative Niedersächsisches Vorab of the Volkswagen Foundation and the Ministry of Science and Culture of Lower Saxony.

References

1. D. Smyth, W. A. Phillips, and J. Kay. "Measures for investigating the contextual modulation of information transmission." In: **Network: Computation in Neural Systems** 7.2 (May 1996), pp. 307–316.

2. M. Larkum. "A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex." In: **Trends in Neurosciences** 36.3 (2013), pp. 141–151.
3. J. Kay and W. A. Phillips. "Coherent Infomax as a Computational Goal for Neural Systems." In: **Bulletin of Mathematical Biology** 73.2 (2011), pp. 344–372.
4. M. Wibral, V. Priesemann, J. Kay, J. Lizier, and W. A. Phillips. "Partial information decomposition as a unified approach to the specification of neural goal functions." In: **Brain and Cognition** 112 (2017), pp. 25–38.