ARTICLE IN PRESS

Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Temporal code versus rate code for binary Information Sources

Agnieszka Pregowska, Janusz Szczepanski*, Eligiusz Wajnryb

Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawinskiego 5B, Warsaw, Poland

ARTICLE INFO

Article history:
Received 3 November 2015
Received in revised form
8 June 2016
Accepted 10 August 2016
Communicated by Hu Jun

Keywords:
Information Theory
Information Source
Stochastic process
Information transmission rate
Firing rate

ABSTRACT

Neuroscientists formulate very different hypotheses about the nature of neural coding. At one extreme, it has been argued that neurons encode information through relatively slow changes in the arrival rates of individual spikes (rate codes) and that the irregularity in the spike trains reflects the noise in the system. At the other extreme, this irregularity is the code itself (temporal codes) so that the precise timing of every spike carries additional information about the input. It is well known that in the estimation of Shannon Information Transmission Rate, the patterns and temporal structures are taken into account, while the "rate code" is already determined by the firing rate, i.e. by the spike frequency. In this paper we compare these two types of codes for binary Information Sources, which model encoded spike trains. Assuming that the information transmitted by a neuron is governed by an uncorrelated stochastic process or by a process with a memory, we compare the Information Transmission Rates carried by such spike trains with their firing rates. Here we show that a crucial role in the relation between information transmission and firing rates is played by a factor that we call the "jumping" parameter. This parameter corresponds to the probability of transitions from the no-spike-state to the spike-state and *vice versa*. For low jumping parameter values, the quotient of information and firing rates is a monotonically decreasing function of the firing rate, and there therefore a straightforward, one-to-one, relation between temporal and rate codes. However, it turns out that for large enough values of the jumping parameter this quotient is a non-monotonic function of the firing rate and it exhibits a global maximum, so that in this case there is an optimal firing rate. Moreover, there is no one-to-one relation between information and firing rates, so the temporal and rate codes differ qualitatively. This leads to the observation that the behavior of the quotient of information and firing rates for a large jumping parameter value is especially important in the context of bursting phenomena.

1. Introduction

A fundamental problem in neuroscience is to understand how neurons encode and process information [1–3]. In general, it is not easy to determine the neural code structure. Since Adrian's experiments [4], which established that individual sensory neurons produce action potentials, or spikes, it has been assumed that a single neuron provides information just through spike sequences, i.e. spike trains. Although it is now generally accepted that a spike sequence is the way in which the information is coded by a single neuron, the structure and the mechanisms of code formation remain a mystery. In 1976, Burns and Webb [5] showed for the first time that the total number of emitted spikes arrives in a highly irregular manner. When the same stimulus is applied repeatedly, the number of spikes varies substantially from trial to trial [6]. This has led neuroscientists to formulate very different hypotheses about the nature of the neural code. Two non-mutually exclusive main theories are of special interest. The first theory is based on "temporal code" [3,7–9], which considers the structure of the spike trains while the second, referred to as "rate code" theory [1,3,10–12], assumes that the neural code is embedded in the spike frequency, defined as the number of spikes emitted per second. The temporal coding mechanism, which builds a temporal relationship between the output firing patterns and the inputs of the nervous system, has received significant attention [13–15].

The temporal rules used for processing precise spiking patterns have recently emerged as ways of emulating the brain's computation from its anatomy and physiology, especially in the context of learning and classification problems. In [16], a unified and consistent feedforward system network with a proper encoding scheme and supervised temporal rules was built to solve the pattern recognition task. In this scheme, external stimuli are converted into sparse representations and these temporal patterns are then learned through biologically derived algorithms in the learning layer, followed by the final decision presented through the readout layer. In [17], an integrated computational model with

* Corresponding author.
E-mail address: jszczepa@ippt.pan.pl (J. Szczepanski).

http://dx.doi.org/10.1016/j.neucom.2016.08.034

a latency-phase encoding method and a supervised spike-timing-based learning algorithm was developed. The classification capabilities of such a system that precisely computes timed spikes and realistic stimuli, analogous to cognitive computation in the human brain, have been demonstrated. Recently, efficient pattern classification methods exploiting two layered spiking neural networks have been proposed [18]. In this network, the input layer consists of receptive field neurons, which convert a real-valued input into spikes using the population coding scheme, without any delays. A functional role for precise spike timing has been considered as an alternative hypothesis to rate coding [19], where it was shown that both the synchronous firing code and the population rate code can be used dually in a common framework of a single neural network model. Ref. [20] addressed how efficient stimulus encoding can be carried out within the early stages of the olfactory system. The authors compared the rate-coding scheme with the direct transmission of graded potentials in terms of the accuracy of the estimate that an ideal observer may make about the stimulus. Ref. [21] applied weight limitation constraints to the spike time error-backpropagation algorithm for temporally encoded networks of spiking neurons. They presented a novel solution to the problem raised by non-firing neurons, which lead to a reliable and efficient convergence of the learning algorithm. Ref. [22] proposed a new learning machine method consisting of a recurrent hierarchical neural network of unsupervised processing units, which they called the Clustering Interpreting Probabilistic Associative Memory design. It turned out that this network exploiting temporal spike pattern rules recognize rotated, translated and scaled patterns. Addressing the speed of visual processing, a model that exploits orthogonal wavelet transform was developed [23]. This strategy provides a spike code, thanks to a rank order coding scheme, which offers an alternative to the classical firing rate coding scheme. It can provide efficient real-time applications using an artificial asynchronous neural network that can mimic nature's performance.

This successful classification methodology has also been supported experimentally. Using the MNIST handwritten digit database for machine learning [24], it has been shown that neural net classifiers tend to perform significantly better than other types of classifiers. Specifically, the convolution structure of neural nets accounts for the excellent classification performance.

These results illustrate that the pattern recognition, and consequently the temporal coding, plays an important role in the design of efficient decoding rules based on neuronal networks. Our paper supports these assumptions, especially the alternative hypotheses [19,23] of temporal and rate coding. We show that temporal coding plays an important role when the activity of the neurons is high, i.e."jumping" parameter, which measures transition from state to state is large. Depending on this parameter, temporal coding can be more effective than rate coding.

However, in the transfer of information, the process demanding the most energy is the spiking process [25,26]. Thus, in the first approximation, the firing rate can be treated as the energy marker. Inspired by thermodynamics [27], we also consider the derivative of entropy over energy, which is the analog of the inverse of the temperature.

In this paper we provide theoretical insights into the understanding of the nature of the neural code by studying this problem for two types of binary Information Sources. Assuming that the information transmitted by a neuron is governed by uncorrelated stochastic processes or by processes with a memory, we study the relation between the Information Transmission Rates (*ITR*) carried by such spike trains and their firing rate (F_R). To this end, the Information-Firing-Quotient (*IFQ*), i.e. the ratio of information and firing rate, is introduced in Section 2. For large *IFQ*, the amount of transmitted information is more optimal but it comes at the cost of unit energy. Thus, the value of F_r for which *IFQ* is maximal, can be understood as an optimal value in this sense. Having parameters that characterize the Information Source and the communication channel, one can determine this optimal firing rate numerically. We show that the crucial role in studying *IFQ* properties is played by the "jumping" parameter. This parameter is the sum of transition probabilities from a no-spike-state to a spike-state and *vice versa*. We show that, for low jumping parameter values, the quotient of information and firing rates is a monotonically decreasing function of the firing rate and there is therefore a straightforward, one-to-one, relation between the temporal and the rate codes. However, it turns out that for large enough values of the jumping parameter, this quotient is a non-monotonic function of the firing rate and it exhibits a clear global maximum. Thus, in this case there is an optimal firing rate. Moreover, there is no one-to-one relation between information and firing rate, and the temporal and rate codes differ qualitatively. The behavior of the quotient of information and firing rates for large values of the jumping parameter is especially important in the context of bursting phenomena [28–30].

The paper is organized as follows: in Section 2, we briefly state the basic concepts of Information Theory and formulae concerning Bernoulli and Markov processes; in Section 3, we present the comparison of information transmission and firing rates for these processes; and the last section contains the Discussion and conclusions.

2. Information Theory in neuroscience

In neuroscience, information transfer has been quantified by many authors in terms of Information Theory [3,31]. In general, neuronal communication systems are represented by an Information Source, a communication channel and output signals [32–34]. Both messages coming from the Information Source and the output signals are represented by sequences of symbols [3,30,31,35,36]. These sequences can be understood as trajectories of stationary stochastic processes. In this paper, we study the Information Sources represented by Bernoulli or Markov processes [35,37]. Markov processes are adequate for modelling spike trains, especially when the neuron firing times in individual trials are of interest [38]. When large numbers of neurons are considered, the spike times for the collection of neurons are often modelled as a Poisson process. Sometimes, as an alternative, non-homogeneous Poisson models are employed for interspike intervals [38]. In these cases similar considerations would be performed. In these cases there are not explicit analytical formulas for *ITR*; therefore, the first step in the analysis would be based on numerical simulations.

Entropy. First, we briefly recall the fundamental concepts of Information Theory [32–34] that are adapted to neuroscience issues. Let Z^L be a set of all words (i.e. blocks) of length L, built of symbols (letters) from some finite alphabet Z. Each word z^L can be treated as a message sent by Information Source Z being a stationary stochastic process. If $P(z^L)$ denotes the probability that the word $z^L \in Z^L$ occurs, then the information in the Shannon sense carried by this word is defined as

$$I(z^L) := -\ln P(z^L). \tag{1}$$

In this sense, less probable events carry more information. We use the natural logarithm to obtain a more compact form of the formulas. When logarithm to the base 2 is used, the ln 2 factor has to be included. Expected or average information of Z^L, called Shannon block entropy, reads as follows:

$$H(Z^L) := -\sum_{z^L \in Z^L} P(z^L) \ln P(z^L). \tag{2}$$

Please cite this article as: A. Pregowska, et al., Temporal code versus rate code for binary Information Sources, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.08.034