Available online at www.sciencedirect.com

SCIENCE DIRECT

NEUROCOMPUTING

Neurocomputing 69 (2006) 1175–1178

www.elsevier.com/locate/neucom

Reward-biased probabilistic decision-making: Mean-field predictions and spiking simulations

Daniel Martí[a,*], Gustavo Deco[a,b], Paolo Del Giudice[c], Maurizio Mattia[c]

[a]Computational Neuroscience Unit, Universitat Pompeu Fabra, Pg. Circumval-lació 8, E-08003 Barcelona, Spain
[b]Institució Catalana d'Estudis Avançats (ICREA)
[c]Complex Systems Unit, Department of Technologies and Health, Istituto Superiore di Sanità, V.le Regina Elena 299, 00161 Roma, Italy

Available online 3 February 2006

Abstract

In this work we study the basic competitive and cooperative mechanisms of neural activity in the context of a two-alternative free-choice eye-movement task, as a function of the expectation of reward. We use a simplified version of the protocol followed by Platt and Glimcher [Neural correlates of decision variables in parietal cortex, Nature 400 (1999) 233–238], in which each choice is associated with independent underlying reward schedules, and explicitly model it using a biophysically realistic network of integrate-and-fire neurons that forms a categorical choice from the expected gain contingencies, via a simple bias mechanism. The model accounts for several experimental findings, such as the gain-modulated firing activity observed by Platt and Glimcher and the matching law.

Keywords: Computational neuroscience; Decision-making; Network model; Lateral intraparietal area

1. Introduction

In recent years the neural signatures that encode behavioral value have been identified, opening the possibility to investigate decision-making at the physiological level. In this work we study the basic competitive and cooperative mechanisms that underlie the neural activity correlated with the decision-making process in primates. In particular, we model the dependence of neural activity on the expectation of reward associated with the eye-movement response performed by the monkey. To achieve this, we explicitly model the processes occurring at the level of AMPA, NMDA and GABA synapses using a cortical recurrent network of integrate-and-fire neurons. Due to the rich phenomenology of the spiking dynamics, a preliminary analysis of the dynamical regimes accessible to the system is done. This analysis consists in exploring the stationary attractors in the relevant parameter space via a mean-field reduction consistent with the underlying synaptic and spiking dynamics [4]. Once the regimes of operation of the network are characterized and the corresponding parameter ranges revealed, both the non-stationary dynamical behavior, as measured in neuronal recording experiments, and the asymptotic stationary regimes are studied via the full simulation of the spiking network.

2. Behavioral task

In our simulations we have used a simplified version of one of the protocols used by Platt and Glimcher [5]. In the task, while the subject is keeping his gaze aligned to the fixation point, two eccentric stimuli are illuminated. After some time, the extinction of the central stimulus instructs the subject to look at either of the two eccentric stimuli. The expected gain associated to each stimulus is manipulated by delivering different amounts of juice to the monkey. In this sense the estimation of value made by the subject can be controlled externally. The expected gain for each response is then varied across blocks of trials to test whether neural activity in lateral intraparietal (LIP) area is correlated with subjective value. The frequency with which the animal chooses each response is used as a

*Corresponding author.
E-mail addresses: daniel.marti@upf.edu (D. Martí), gustavo.deco@upf.edu (G. Deco), paolo.delgiudice@iss.infn.it (P.D. Giudice), mattia@iss.infn.it (M. Mattia).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.12.069

behavioral readout of the subjective value of each option. Platt and Glimcher explicitly proved that subjective value was represented by the firing activity of LIP neurons. They also observed the 'matching law' in action, giving a linear dependence of the probability of a given choice on the reward bias.

3. Computational model

Our network model is composed of a pair of neural excitatory 'selective' populations (or pools, labeled A and B) with strong recurrent synaptic self-couplings and weak mutual excitation. In addition, A and B are reciprocally connected to an inhibitory population and to an 'unselective' excitatory population; all populations receive external excitatory synaptic inputs coding for stimuli and other external influences, including background spontaneous activity (see Fig. 1 and [1,6] for the general theoretical setting). A and B can be 'selective' in that they react to stimuli and can engage in competition due to shared inhibition, such that even for equal or very similar inputs to A and B the network can exhibit high A firing activity with suppressed B activity (which will be taken to encode 'decision A') or the reverse ('decision B') [3,7].

In absence of stimuli, every cell in the module receives external input modeled as a Poisson train with rate

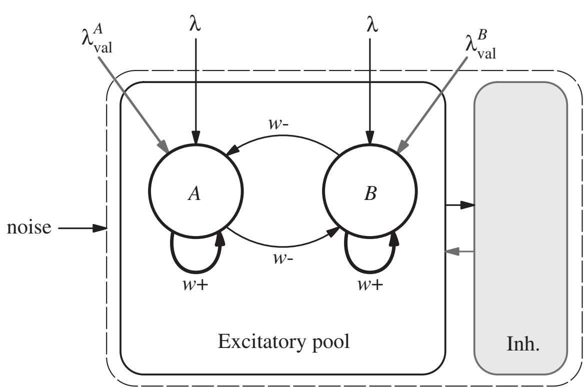

Fig. 1. Architecture of the network. The module consists in N neurons and is divided in two major groups: one inhibitory and one excitatory population, with N_{I} and N_{E} neurons each ($N = N_{\mathrm{E}} + N_{\mathrm{I}}$). Within the excitatory pool there are two types of populations: two selective pools A and B, each constituted by fN_{E} neurons ($f = 0.15$), and one non-selective population, formed by all excitatory neurons not belonging to a selective pool ($(1-2f)N_{\mathrm{E}}$). The firing rate activity of the two selective pools encode the decision to make. w_+ are the synaptic weights connecting neurons within the same selective pool, whereas w_- denotes the connection weight between neurons in different selective pools and from non-selective to selective neurons. All other possible connections have weight 1 (baseline strength). To assure that the overall recurrent excitatory synaptic drive in the spontaneous state remains constant as w_+ is modified, w_- is set to $1 - f(w_+ - 1)/(1-f)$. The signals associated to the stimuli are denoted by λ and the signal carrying value information is $\lambda_{\mathrm{val}}^{\mathrm{A,B}}$ (one per pool). Every neuron in the network receives a background Poisson spike train of rate 2.4 kHz.

$\nu_{\mathrm{noise}} = \nu_{\mathrm{out}} N_{\mathrm{ext}} \sim 3\,\mathrm{Hz} \times 800 = 2.4\,\mathrm{kHz}$, where ν_{out} is the average firing rate of any neuron outside the module and N_{ext} is the number of external synapses. The presence of a stimulus is implemented by an increase of the external input perceived by every selective neuron. So, during stimulus presentation a selective neuron, either in A or in B, receives a Poisson spike train of rate $\nu_{\mathrm{ext}} = \nu_{\mathrm{noise}} + \lambda$, where λ represents the intensity of the stimulus. The expectation of reward is implemented extrinsically; we assume that the decision-making process is triggered by an external signal coming from a module that stores the representation of value. The value signal is added to the total background noise perceived by each neural population. Even though this is an over-simplified model, it can shed light on how basic reward-biased decision-making mechanisms work in a network model.

3.1. Mean-field parameter exploration

Spiking simulations are too computationally expensive for an extensive search in the parameter space. Mean-field approximations allow to compute the attractors to which the network would converge in the limit of an infinite number of neurons, and require much less computational load. The mean-field approximation we used was that derived by Brunel and Wang [2]. The goal of these explorations was to find the number of stable network states (attractors) that coexist for a given set of parameters. We can distinguish four different stable network states: in the *spontaneous* state (S) the firing rates of the two populations are comparable and low ($\nu_{\mathrm{A}} \simeq \nu_{\mathrm{B}} \sim 3\,\mathrm{Hz}$); in the *mixed state* (M) the activity of the two populations is also comparable, but substantially higher than typical spontaneous activity. The other two states are the *selective* states (A and B), in which one of the two populations shows elevated activity while the other population fires at a very low (suppressed) rate. In a selective state the ratio of the high firing rate over the suppressed firing rate, $\nu_{\mathrm{high}}/\nu_{\mathrm{low}}$, is typically higher than 10. If the stable states A and B represent the two categorical options the network has to choose from, the mixed state M could be an interesting dynamic option for describing an 'undecided' state, possibly corresponding to unusually long decision times.

The network shows multistable behavior: there may coexist several stable states given a fixed set of parameters. The set of stable states that the network can sustain for some values of the parameters determines the phase or regime of operation of the system. In this system, the relevant parameters are the recurrent potentiation weight w_+, and the amplitude of the external signal received by each neuron in a pool: $\nu_{\mathrm{ext}}^{\mathrm{A}} = \nu_{\mathrm{noise}} + \lambda + \lambda_{\mathrm{val}}^{\mathrm{A}}$, $\nu_{\mathrm{ext}}^{\mathrm{B}} = \nu_{\mathrm{noise}} + \lambda + \lambda_{\mathrm{val}}^{\mathrm{B}}$. It is convenient to define $\bar{\lambda} \equiv \lambda + (\lambda_{\mathrm{val}}^{\mathrm{A}} + \lambda_{\mathrm{val}}^{\mathrm{B}})/2$ and $\Delta\lambda \equiv (\lambda_{\mathrm{val}}^{\mathrm{A}} - \lambda_{\mathrm{val}}^{\mathrm{B}})/2$. Fig. 2 summarizes the regimes of operation of the network found at each point in the parameter space.