Neurocomputing 123 (2014) 140–149

Contents lists available at ScienceDirect

Neurocomputing

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

Reinforcement-conflict based control: An integrative model of error detection in anterior cingulate cortex

Sareh Zendehrouh, Shahriar Gharibzadeh *, Farzad Towhidkhah

Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 30 November 2012
Received in revised form
11 May 2013
Accepted 17 June 2013
Communicated by A. Belatreche
Available online 17 July 2013

Keywords:
Error related negativity
N200
Performance monitoring
Executive control
Computational model

ABSTRACT

The concepts of error detection and compensation play a key role in the studies of human brain behavior. In particular, studies of event-related brain potentials have discovered the presence of a component following errors named the error related negativity (ERN). According to the highly used theories of this field, the ERN can occur due to an error detection in reward prediction (the reinforcement learning theory) or conflict detection (the conflict monitoring theory). Since each of these theories has given numerous experimental validations, there is an urge to reconcile them. Although there have been efforts to integrate these two theories, the research in this field needs to be improved. The present study aims at integrating the above-mentioned theories by proposing a model that simulates the role of anterior cingulate cortex in performance monitoring and reinforcement-guided action selection. In this model, an executive control module uses the conflict signal and the reward prediction error to adjust the timing of the control and to regulate the strength of the control over other structures. The simulation results show that the proposed model can successfully produce the expected experimental results in a modified version of the Eriksen flanker task.

1. Introduction

Anterior cingulate cortex (ACC) has been recognized to be involved in performance monitoring and cognitive control [1]. One of the functions of this region is to detect the occurrence of errors that instigate further control of prefrontal cortex over cognitive or motor processes [2,3]. Error detection and remedial processes are crucial for human cognitive system [4]. Without these abilities, it would be impossible to perform successful behavior.

Studies of the event-related brain potential (ERP) have reported the presence of a negative deflection named the error related negativity (ERN[1]) that begins near the time of the erroneous response and peaks about 100 ms later [11]. Converging evidence from fMRI [12,13], intracranial recording [14], dipole source localization [15,16] suggest that the ERN is generated in the ACC. In addition to the ERN, a negative deflection has also been observed on correct trials prior to the execution of a response in cognitive control experiments. This N200[2] (N2) component peaks between 200 and 400 ms after stimulus onset [10] and appears to have the same neural generator as the ERN [16,19].

Since the discovery of the ERN, several theories such as error detection [11] and processing [6], reinforcement learning (RL) [5,20], response conflict monitoring [19,21], error likelihood [22], and motivational significance [23] have been proposed regarding the mechanism of the ERN generation. Error detection and processing theories state that the ERN is produced either by a process that detects errors [11] or by a process that is engaged following a detected error [6]. Since, ACC neurons are known to react to action's outcome, it is argued that this region is a part of a system dedicated to reinforcement learning. In this view, the reinforcement signals (reward prediction errors (RPEs)) thought to be under the control of dopamine, modulate processing in the ACC and prefrontal cortex [24]. The RL theory of the ERN states that this component indicates an error in the reward prediction [5,20]. The conflict monitoring theory proposes that the ERN can

* Corresponding author. Tel.: +98 21 6454 2378; fax: +98 21 6649 5655.
E-mail address: gharibzadeh@aut.ac.ir (S. Gharibzadeh).

[1] In this article, we emphasize that by an ERN, we mean error-related negativities associated with incorrect responses (also called response ERN (rERN) [5] or error negativity (Ne) [6]), and we do not consider negativities associated with feedback (feedback ERN (fERN), also called feedback related negativity (FRN) [7,8]). FRN (or fERN) is a component observed 230 to 330 ms following incorrect feedback [8] in gambling and trial-and-error learning tasks [9]. In spite of the similarities between these two components, some researchers suggest that they are significantly different [10].

[2] The anterior N2 can be divided into separate mismatch- and control-related subcomponents. The control- related N2 includes no-go N2 and conflict N2 [17]. Several findings from the go/no-go tasks confirm the relation of the no-go N2 to response inhibition, but some data propose that the no-go N2 is predominantly associated with conflict monitoring processes [18]. Here, we report on the conflict-related N2 that appears in tasks such as the Eriksen flanker task.

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.06.020

be explained in terms of coactivation of mutually incompatible response processes (response conflict) [21]. According to this theory, dorsal ACC (dACC or midcingulate cortex (MCC)) detects conflict during response selection and conveys this information to other brain regions for further adjustment and control [25,26]. In the motivational significance theory, the ERN reflects an evaluation of the emotional or motivational significance of an error, and the result of this evaluation is evident in the magnitude of the ERN [23]. However, the error likelihood theory states that the ACC is activated in response to a given task and the magnitude of its activity signals the predicted likelihood of an error [22].

Among these theories, the RL and the response conflict monitoring are in the core of attention [27]. Numerous experiments suggest the plausibility of these two hypotheses. However, none can be excluded entirely [27,28]. The RL theory is based on a class of computational learning models named temporal difference [29] that enables it to be used in reward-guided decision-making. Although the RL theory can account for reward-guided choices, the main weakness of this theory is that it cannot explain conflict-related phenomena. However, the opposite is true for the conflict monitoring theory. Briefly speaking, the RL theory provides explanations for the ERN and reward-guided behaviors, whereas the conflict monitoring theory provides explanations for the ERN and N2.

Several attempts have been made to reconcile the conflict monitoring and the RL theories, particularly using a computational model. Cockburn and colleagues presented a model in which the prediction error signals alter the basal ganglia activity and then generate the conflict in the ACC [27,30]. Although this model unites the RL and the conflict monitoring theories, the detected conflict in the ACC is not used for any further control or any specific function. Putting together the conflict monitoring and decision making accounts, Botvinick [28] proposed a theory that is an extension of the conflict monitoring theory. In this theory, conflict as a teaching signal biases behavioral decision making toward strategies that are more efficient. The initial computational model of Botvinick [28] for demand selection task imitates the empirical data by developing a tendency toward selection of the low-demand task. In this model, using conflict data, the selection of the high-demand task is weakened. Although this theory and its preliminary model are the extensions of the conflict monitoring into the higher level of strategy selection, they do not use the ideas behind the RL theory. Alexander and Brown [31] also proposed the new "predicted response–outcome" (PRO) model. The PRO model is a generalization of RL algorithms that learns and predicts the likely outcomes of actions, detects differences between actual and predicted outcomes, and updates the outcome predictions accordingly. They reported the results concerning the N2 and error effects, but did not provide results regarding the ERN component itself. In our previous work [32], we also proposed a model that is a modified and extended version of RL computational model, which includes the mechanism underlying conflict computation. Although this model produces the N2 and ERN components, the weakness of the model is that the computed conflict signal does not have any controlling effect on other parts of the model.

In the present study, we address this issue by inserting a module that uses: (1) the conflict signals to adjust the timing of the control and (2) the reward prediction errors to regulate the strength of the control over other modules. We then demonstrate that the model can still produce the ERN and N2 components.

1.1. Physiological basis of the model

In both motor and cognitive domains, the process of facilitating the selected action and suppressing other possible actions is learnt via a dopaminergic signal that modulates the activity in the neurons of the striatum [33]. Dopamine neuron activity represents the RPE that signals the difference between the observed and the expected outcome. The reward-motivated behavior stems from activities of dopamine neurons in the ventral tegmental area (VTA) and substantia nigra pars compacta of midbrain [34]. The RPE signal is sent to the corresponding terminal regions as a teaching signal for synaptic modification [35]. One of these regions is ventral striatum that VTA mostly projects to it [36]. One classic interpretation of the ventral striatum function is that it evaluates and learns the current environmental state [37]. Changes in the activity of dopamine neurons drive learning in this area to enhance the accuracy of its predictions in the future [38]. Therefore, prediction error signals are used to learn the values of states, state-action pairs, or both in this area. These values are then used to facilitate the proper action and suppress other competing actions in the cortex. In brief, the cortex first produces candidate actions and then the basal ganglia help to select the proper action [39].

The VTA projections to the ventral striatum activate cholinergic projections to the prefrontal cortex, which are hypothesized to lead to the activation of anterior attention system and related executive functions [40]. Attention can be directed either by external inputs (i.e. exogenous attention) or by internal goals (top-down or endogenous attention). Some overlap exists between exogenous and endogenous attention networks. Parietal lobes, the anterior cingulate cortex, and the prefrontal cortex are proposed to constitute the most important cortical components of these attention networks. By modulating the inputs or response activities in occipital, temporal and motor cortices, the prefrontal and parietal areas of the cortex implement the control of attention [41].

2. Method

Based on the above-mentioned physiological evidences, we propose an extension of our previous model [32]. Here, we simulate human behavior in a modified version of the Eriksen flanker task. We use the behavioral data of Holroyd et al. in which the task is performed by 15 participants [5]. For every participant, there are 24 blocks, each consisting of 200 trials. Thus, 4800 trials are simulated for each participant. Every stimulus is a five-letter array of "H" and "S" letters. In sum, four kinds of stimuli are used (HHHHH, SSHSS, SSSSS, and HHSHH). The central letter of the array is the target stimulus, and the non-target (flanking) letters are either congruent (i.e., HHHHH and SSSSS) or incongruent (i.e., SSHSS and HHSHH) with this target. Moreover, two stimuli with the same target (e.g. HHSHH and SSSSS) were frequent and appeared on 40% of the trials. The remaining two stimuli (SSHSS and HHHHH) were infrequent and appeared only on 10% of the trials. Therefore, four conditions can be determined: infrequent congruent (III), infrequent incongruent (FIF), frequent incongruent (IFI), and frequent congruent (FFF). Participants were requested to respond to the H target with the left hand and to the S target with the right hand.

The proposed model is based on the actor–critic method. Therefore, an actor module (that maps stimulus to response) and a critic module (that criticizes the actions of the actor) constitute two main modules of the model. The executive control module is also added to the model that evaluates the overall performance, detects the need for further control, and consequently implements the control.

2.1. Actor and critic modules

As can be seen in Fig. 1, the three modules of the model are composed of the following blocks: (1) a perception block that