Biologically Inspired Cognitive Architectures (2015) 12, 144–154

Available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bica

INVITED ARTICLE

An architecture of the cognitive system with account for emotional component

O.D. Chernavskaya [a,*], D.S. Chernavskii [a], V.P. Karp [b], A.P. Nikitin [c], D.S. Shchepetov [a], Ya.A. Rozhylo [d]

[a] P.N.Lebedev Physical Institute (LPI), Moscow, Russia
[b] MIREA, Moscow, Russia
[c] General Physics Institute (GPI), Moscow, Russia
[d] BICA Labs, Bogomoletz Institute of Physiology, Kiev, Ukraine

Received 11 March 2015; received in revised form 11 April 2015; accepted 12 April 2015

KEYWORDS
Information;
Learning;
Emotions;
Noise amplitude;
Unexpectedness;
Humor

Abstract
This article represents an extension of authors' previous papers (Chernavskaya, Chernavskii, Karp, Nikitin, & Shchepetov, 2012, 2013) in modeling cognitive systems on the base of the Dynamical Theory of Information. The paper focuses on the problem of account for emotions in artificial system. The main hypothesis consists in the assumption that emotions inherent in a living system could be simulated by variation of amplitude of the occasional component (*noise*) inherently embedded into the architecture of artificial system. Within this concept, increasing noise amplitude should correspond to negative emotions (*anxiety*), while its decreasing provides positive emotions (*relaxation, pleasure*). A rapid up-and-down spike in the noise amplitude could imitate a *laugh*. This hypothesis is secured by incorporation of an additional dynamical variable that represents an analogy to the compound of neural mediators in human beings. The system of linked equations in terms of ''noise amplitude – neurotransmitter compound'' is proposed to describe mutual influence of the cognitive process and emotional component. The model permits to reproduce qualitatively certain prominent effects typical for human emotional reactions (like *stress* and *shock*).

Introduction

Recently, we have proposed a version of the cognitive system architecture based on the Dynamical Theory of

* Corresponding author.
E-mail address: olgadmitcher@gmail.com (O.D. Chernavskaya).

http://dx.doi.org/10.1016/j.bica.2015.04.009
2212-683X/

Information (DTI) and neurophysiology data (Chernavskaya, Chernavskii, Karp, Nikitin, & Shchepetov, 2012, 2013). This article represents further advance devoted to analysis of the role and place of *emotions* in the considered scheme. It should be stressed that, in contrast to challenges of Artificial Intelligence (*AI*), that tries to present an artificial system capable of solving some problems *better* than humans do, we are trying to understand *how the very human* cognitive process (including its emotional component) might be designed. The main challenge here is to match neurophysiology level based on the single-neuron state with psychological level realized by human himself. Besides, one should not forget about another challenge: any artificial system pretending to reproduce human cognitive functions should be *individual*, as every human being is.

Regarding human-like cognitive systems, one cannot ignore role of emotions. There have been numerous attempts to consider the emotional component in modeling the cognitive process (see, e.g., Dancy, 2013; Hudlicka, 2014; Izhikevich & Edelman, 2008; Larue, Poirier, & Nkambou, 2013; Shamis, 2006; Yakhno, 1995; Zhdanov, 2009). The majority of researches refers to the *active agent* concept (Laird, 2012; Samsonovich, 2007, 2013; Sellers, 2013) and suggests various principles of organization of the ''emotional space'' that affect the cognitive process. Resembling way is to introduce several discrete emotional states that would affect (with certain weight coefficients) the model calculations for *AI*. Their number may vary — from two (positive and negative ones, Yakhno, 1995) up to 27 (Samsonovich, 2013). However, the exact mechanism of the emotion emergence is not revealed.

It deserves mentioning that some authors (e.g., Sellers, 2013) point out that emotions refer actually not to a *current state*, but to *transition* between different states. In other words, an emotion (in particular, *happiness*) represents not something that *exists*, but something that *changes*, — therefore, any emotional state could not continue for a long time. We completely share and support this viewpoint.

Another approach (Rabinovich & Muezzinoglu, 2010; Treur, 2013) involves two sets of dynamical variables, emotional and rational ones, so that their (nonlinear!) interaction results in various states of the system providing certain nontrivial regimes of transition between those states. However, the neurophysiology interpretation of the emotional, as well as rational, variables within this approach remains somewhat dissatisfactory.

Another approach, called *reverse brain engineering* (see, e.g., Doya, 2000; Koziol & Budding, 2009; Panksepp & Biven, 2012) is based on the analysis of various structures in the brain — *thalamus*, *basal ganglia*, *corpus amygdaloideum*, etc., — directly connected with control of the emotions in cognitive process. This way seems to be the closest to the aim, but the consideration actually is almost verbal: used mathematical apparatus seems rather poor.

Thus the challenge of incorporating the emotions into cognitive architecture still remains actual.

The problem's difficulty is connected with ambivalence of the very notion of emotions. According to the common definition, emotions (Latin *emovere* — excite, worry) are reflections of someone's own subjective attitude towards existing or predicted situation. This definition already implies that the person possesses self-consciousness and ability to appraise own state. Since our goal is to draw up this notion from the features of neuron ensemble, this ability should be formalized at the neural level. However, the emotion formalization itself represents a challenge, since everybody does realize what it is, but could hardly formulate the certain feeling. The most natural way to formalize the emotional component is multi-disciplinary approach.

In this paper we try to collect the interpretations and the main features of emotions provided by different approaches and propose our version of accounting for emotional component in the artificial cognitive system.

Different approaches to formalization of emotions

An approach from the evolution theory

From the evolutionary point of view, emotions represent an ancient mechanism of the analysis of environment to a much bigger extent rather than cognitive activity. Emotional burst, being the product of allocortical and sub-cortical structures, provides synthetic (*integral*) reactions that appear before the analysis of concrete reasons and motives. For humans, the specification of ''*emotio*'' and ''*ratio*'' becomes meaningful after formation of the common ***language*** (that is, the developed system of conventional symbols) within a certain community, since rational thinking implies an ability to pose a problem and argue solution (see, e.g., Deacon, 1997). Let us point out that any language-delivered information (speech) represents a ***successive time set*** of symbols. Hence the reasoning, or rational thinking, represents a *consecutive* method of information processing. Then it seems reasonable to assume that non-rational and emotional reactions correspond to the *parallel* processing of information. Note that the same specialization is widely attributed to the left and right cerebral hemispheres, respectively (e.g., Bianki, 1984; Stirling & Eliott, 2010). Therefore, one can suppose that rational thinking refers to the left hemisphere (LH), while the right hemisphere (RH) is rather connected with emotional component of thinking process. These reasons partly resemble popular (near ''folk'') belief that LH provides a verbal-logical thinking, while RH represents a sensory-imaginary thinking mode (see, e.g., Shamis, 2006). However, such specification seems to be not rather relevant. We shall discuss it below in more detail.

Psychological approach

From the psychological point of view, emotions associated with achieving certain *goal* could be formalized rather simply (Shamis, 2006; Solso, 1998; Zhdanov, 2009): it requires an appraisal of the mission possibility and certain steps to complete the mission. Thus, increasing probability of the goal attainment leads to positive emotions, and vice versa. However, modeling the cognitive system within our approach requires formalization of the notion of ''goal'' at the neurallevel, that itself represents a challenge.

The most credible and common viewpoint is to consider any *new* (*unexpected*) things as *call for emotions*. This concept has been already reflected in robotics (see, e.g.,