On Possibility to Imitate Emotions and a "Sense of Humor" in an Artificial Cognitive System

Olga Chernavskaya
P.N.Lebedev Physical Institute (LPI)
Moscow, Russia
E-mail: olgadmitcher@gmail.com

Yaroslav Rozhylo
NGO "Ukrainian Center for Social Data"
Kyev, Ukraine
E-mail: yarikas@gmail.com

***Abstract*—The problem of modeling and simulation of emotions and a sense of humor in an artificial cognitive system is considered within Natural-Constructive Approach (NCA) to modeling the human thinking process. The main constructive feature of this approach consists in splitting up the cognitive system into two linked subsystems: one responsible for the generation of information (with required presence of an occasional component, "noise"), the other one – for reception of well-known information. It is shown that human emotions could be imitated and displayed by variation of the noise amplitude; this very variation does control the switching on the subsystems activity. The *sense of humor* is proposed to be treated as an ability of quick adaptation to unexpected information (incorrect and/or undone prognosis) with getting positive emotions. It is shown that specific human emotional response to the humor (the laugh) could be imitated by abrupt changing ("spike") in the noise amplitude.**

Keywords- neuroprocessor; noise; information generation; switching.

I. INTRODUCTION

The problem of modeling the cognitive process is actual and very popular now (e.g., [1]-[5]). The majority of imitation models proposed are aimed to construct the artificial cognitive systems (Artificial Intelligence, AI), for solving certain problems *better* than human beings. Hence, those systems have to be *efficient*, *reliable* and *fast-acting*. However, it becomes more and more popular to incorporate emotions into AI systems [2]-[6]. In our works [7], [8], we focus on modeling just the human-like cognitive systems, thus, on the features inherent to the *human* cognition, such as *individuality, intuitive* and *logical* thinking, *emotional impact* to cognitive process, etc. Although the ultimate goals are different, several results obtained within our approach could be applied to design an AI endowed with human-like reactions.

We use so called Natural-Constructive Approach (NCA), which is based on the Dynamical Theory of Information (DTI, [9],[10]), neurophysiology [11], and neural computing [12]-[14]. DTI itself is relatively new theory elaborated in the post-middle of XX*th* century as a subfield of Synergetics [9],[15]. This theory provides clear definition of cognition as the *self-organized process of perception (recording), memorizing (storage), coding, processing, generation and propagation of the information.* Thus, any cognitive architecture is to perform these functions.

Let us stress an important inference of DTI. Since information is defined by Quastler [16] as a ***memorized choice of one version among several possible (and similar) ones,*** it might emerge from just two processes. The first one is the ***generation*** of information, that is, free (occasional) choice. It could appear only in the presence of *occasional component* (the "*noise*"). The second one is ***reception*** of information, which represents a forced (supervised) choice. According to DTI, these modes are *complementary* ones (one possibility excludes the other one), so these functions should be shared between *two different subsystems.*

It should be noted that similar ideas were put forward by psychologist E. Goldberg concerning the role of two cerebral hemispheres [17]: the right one is responsible for learning the new information (generation of information), the left one is dealing with the well-known information (reception). This very specialization of two subsystems is realized in the model presented below.

Recently, these ideas become popular in robotics as well [4]. However, the two-subsystem architecture is not used widely, because the mechanism of regulation of switching-on the subsystem activity has not been revealed yet.

In this paper, we present (schematically) the version of the human-like cognitive architecture elaborated within NCA [7][8]. According to this model, the emotional manifestation in an artificial system could be imitated by the derivative of the noise amplitude. Moreover, this very derivative is shown to be a tool to control the activity of two functional subsystems. A particular case of the noise-amplitude behavior, namely –– the abrupt up-and-down change ("spike"), –– is proposed to be treated as an analogue to human *laugh*.

It is worth noting that, as compared to [8], this paper represents an attempt to apply the results of our analysis of human cognitive process to specific goals of AI design. So, the paper is aimed to attract attention to possible advantages of AI, based on the human-like cognitive architecture.

 ISBN: 978-1-61208-462-6

The paper is organized as follows. Section II presents the description of the cognitive architecture designed within NCA. In Section III, we discuss the role and place of emotions in the architecture proposed. In Section IV, we present the example of application of the model proposed to describe the effects of stress/shock. In Section V, we discuss possible manifestations of the sense of humor in AI. Further working perspectives are discussed in Section VI.

II. ARCHITECTURE OF COGITIVE SYSTEM

The scheme of cognitive architecture designed within *NCA* in our works [7][8] is presented in Fig. 1. This system represents a composition of several neural processors of Hopfield (***H***) and Grossberg (***G***) type, with each processor being a plate populated with *n dynamical* formal neurons. Those processors differ by their functions: ***H***-type one serves for recording the *images* (distributed memory), while ***G***-type plates contain the encoded information (*symbols*). The number of symbolic (***G***) plates is neither fixed nor limited since they appear "as required" in course of system's evolution.

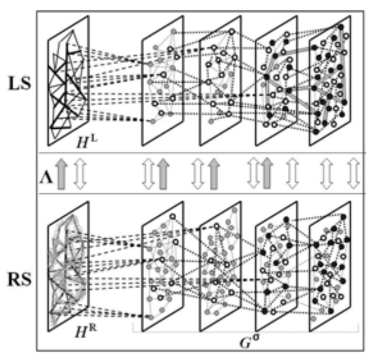

Figure 1. Schematic representation of the cognitive architecture.

A. Constructive Peculiarities

The main constructive feature of this architecture is splitting up the whole cognitive system into two subsystems, with one of them being responsible for perception of *new* information and *learning* (***generation*** of information), while the other one is dealing with well-known information (***reception)***. These subsystems are named "right subsystem" (**RS**) and the "left subsystem" (**LS**) since they represent an analogue to the right and left cerebral hemispheres, respectively. The fact that this sub-system specialization coincides with that put forward in [16] represents a pleasant surprise and indirect indication of NCA relevance.

The equations describing interactions between neurons of various types could be written in the form:

$$\frac{dG_k^{R,\sigma}}{dt} = \frac{1}{\tau_G}[\hat{\Upsilon}\{\alpha_k, G_k^{R,\sigma}, G_l^{R,(\sigma+\nu)}\} + Z(t)\cdot\xi(t) + \Lambda^{L\to R}\cdot G_k^{L,\sigma}, \qquad (1)$$

$$\frac{dG_k^{L,\sigma}}{dt} = \frac{1}{\tau_G}[\hat{\Upsilon}\{\alpha_k, G_k^{L,\sigma}, G_l^{L,(\sigma+\nu)}\}] + \Lambda^{R\to L}\cdot G_k^{R,\sigma}, \qquad (2)$$

where $G^{R,\sigma}{}_k$, $G^{R,\sigma}{}_k$ are dynamical variables referring to the **RS** and **LS** respectively, σ is the number of symbol's level (for the sake of brevity, the imaginary plate ***H*** is treated as ***G^0)***. The functional $\Upsilon\{\alpha_k, G_k^{\sigma}, G_k^{\sigma+\nu}\}$ describes intra- and inter-plate interactions between neurons (for details, see [7]); α_k and τ_G are model parameters. The term Z(t)ξ(t) in (1) corresponds to the occasional component ("*noise*"): Z(t) is the noise amplitude, 0<ξ(t) <1 is random function (obtained, e.g., by the Monte-Carlo method). It is presented in **RS** only, thus securing the ability to generate information. Besides, all connections in **RS** are trained according to Hebbian rule [18]: initially weak, the links become stronger ("blacker") in course of the learning process. When the connections become strong ("black") enough, the image is transferred to **LS.** Such mechanism of learning has been called in [7][8] as *the principle of "connection blackening"*. In **LS**, all connections are trained according to original Hopfield mechanism [12] "excess cut-off". This implies that all connections are initially equal and strong; in the learning process, the connections with neurons that do *not* belong to the given image diminish gradually. Thus, learning in **LS** represents not the *choice*, but *selection*, with **RS** acting as a Supervisor for **LS**.

Connections Λ(t) between those subsystems play the role of *corpus callosum* and provide the "dialog" between the subsystems. They should not be trained, but have to switch on depending on the current goals. At the stage of learning $\Lambda^{R\to L}$ have to switch on accordingly to the "connection blackening" principle. At the stage of solving the problems, the role and mechanism of Λ are to be specified (see below).

B. Solving the Problems

Let us discuss how the problems of ***recognition*** and ***prediction*** could be solved in the already trained system (that has sufficiently developed symbolic structure).

The incoming information is perceived by both subsystems. If it is well known, these problems are solved in the **LS** by means of Hopfield-type mechanism of *refinement*: all the images are treated as already known ones by fitting them to coincide with already stored patterns. In the case of insufficient recognition (when the fitting procedure fails) the participation of **RS** becomes necessary. An unrecognized image is treated as a new one and undergoes the common procedure of a new symbol formation.

The ***prognosis (prediction)*** can be treated as "recognition of time-depending process". It proceeds in **LS** after the symbol of the given *process* is formed. This symbol collects all the information about the "process pattern" in a compressed form. Then, the information on initial stage of the given process activates its symbol, providing the activation of the entire chain of symbols

 ISBN: 978-1-61208-462-6

enclosed in this process.

III. ROLE OF EMOTIONS

Incorporating the emotions into artificial cognitive system represents really the challenge, since emotions have dual nature. On the one hand, they represent *subjective self-appraisal* of the current/future state. On the other hand, emotions are associated with *objective and experimentally measured* compound of neural transmitters in the human organism. The latter is controlled by more ancient brain structures (so called "old cerebrum"), than the neocortex, namely – thalamus, basal ganglia, corpus *amygdaloideum*, etc. [19]. Since the cognitive process is commonly attributed to the activity of neocortex, the realization of mutual influence of these structures requires special efforts. It concerns AI specially, since the notions of *"feeling"*, "*hormone splash"*, "*instinct"*, etc. are absent here. The emotional self-appraisal could be in principle formalized in AI, but this requires definite criteria of the system's state. So, the question of emotion classification is far from trivial.

A. The Problem of Emotion Formalization in AI

In psychology, the self-appraisal (emotion) is ordinarily associated with achieving a certain *goal*. Commonly, they are divided into positive and negative ones, with increasing probability of the goal attainment leading to positive emotions, and vice-versa. Furthermore, it is generally known that any *new* (*unexpected*) thing/situation calls for *negative* emotions [17], since it requires additional efforts to hit the new goal (in the given case, to adapt to unexpected situation). Our representation of emotions relies on this concept as well.

In neurophysiology, emotions are controlled by the level and compound of the *neurotransmitters* inside the organism [11], [19]. The entire variety of neurotransmitters can be sorted into two groups: the *stimulants* (like *adrenalin*, *caffeine*, etc.) and the *inhibitors* (*opiates*, *endorphins*, etc.). Note that this fact indicates indirectly that the binary classification – positive and negative emotions – seems bearable despite its primitiveness. However, there is no direct correspondence between, e.g., positive self-appraisal and the excess of either inhibitors, or stimulants.

According to DTI, emotions could be divided into two types: *impulsive* (useful for generation of information) and *fixing* (effective of reception). Since the generating process requires the noise, it seems natural to associate impulsive emotions (*anxiety, nervousness*) with the *growth of noise amplitude.* Vice-versa, fixing emotions could be associated with *decreasing* noise amplitude (*relief*, *delight*). By defining the goal of the living organism as the maintenance of *homeostasis*, (i.e., calm, undisturbed, stable state), one may infer that, speaking very roughly, this classification could correlate with negative and positive emotions, respectively.

B. The Main Hypothesis on Emotion Representation in AI

We propose the following hypothesis on the nature of emotions: *The occasional component (noise) in artificial systems does correspond to the emotional background of living systems, as well as free (occasional) choice imitates the human emotional choice.*

Within this concept, we get at once three tools directly connected with emotions, with all of them being individual for any given artificial system:

Z_0 – stationary-state background, i.e., the value that characterizes the state "at rest";

$\Delta Z(t) = Z(t) – Z_0$ is the excess of the noise level over the background, which reflects the *measure* of cognitive activity;

dZ/dt – time derivative of the noise amplitude, which apparently is the most promising candidate to the analogue to emotional reaction of human being. The absolute value of derivative dZ/dt corresponds to the *degree* of emotional manifestation: drastic change of noise amplitude imitates either *panic* ($dZ/dt>0$), or *euphoria* ($dZ/dt<0$), and so on.

Various combinations of these values reveal a wide field for speculations and interpretations. For example, the value Z_0, being graduated, could serve as the indicator of *individual temperament.* The states with $Z(t) < Z_0$ could be interpreted as *depression*, etc.

These parameters could be applied to construct artificial cognitive systems (*robots*) of various "psychology" types.

C. Sources of the Noise-Amplitude Variation

In human organism, emotional bursts are actually produced in certain structures of so called *allocortex* ("old cerebrum") [19]. Within our main concept, their influence on the *cognitive* process (commonly attributed to the activity of *neocortex)* could be accounted for by liking the value of *dZ/dt* with an *aggregate* variable μ representing the compound of neural transmitters (i.e., the difference between the *stimulants* and *inhibitors*), as it was done in [8].

In artificial cognitive system (AI), such structures are absent. However, even here we can input an additional variable μ as an external factor to control the "emotional" state of the system. Then, we can write a system of equations describing mutual interaction of μ and Z(t) variation in course of cognitive process:

$$\frac{dZ(t)}{dt} = \frac{1}{\tau^Z} \cdot \{a_{Z\mu} \cdot \mu + a_{ZZ} \cdot (\mathrm{Z} - \mathrm{Z}_0) + F_Z(\mu, Z) + \mathrm{X}\{\mu, G_k^{R,\sigma}\} + [\chi(\mu) \cdot D - \eta(\mu) \cdot \delta(t - t_{D=0})]\}, \quad (3)$$

$$\frac{d\mu}{dt} = \frac{1}{\tau^\mu} \cdot \{a_{\mu\mu} \cdot \mu + a_{\mu Z} \cdot (Z - \mathrm{Z}_0) + F_\mu(\mu, Z)\}, \quad (4)$$

where a, χ,η,τ are model parameters, the functional $\mathrm{X}\{\mu, \mathrm{G}_k^{R,\sigma}\}$ refers to the process of new symbol formation (which decreases $Z(t)$ value, see details in [8]). Linear in Z and μ part in (3), (4) provides the system's homeostasis: stationary stable state corresponds to $\{Z=Z_0,\ \mu=0\}$. The functions $F_Z(\mu,Z)$ in (3) and $F_\mu(\mu,Z)$ in (4) are written to account for possible nonlinear effects, which could emerge from mutual influence of “emotional” (neurophysiology) and “cognitive” (referring to the neocortex ensemble) variables (see below).

The last term in (3) refers to processing the incoming information. D stays for the *discrepancy* between the *incoming* and *internal* (learned and stored) information, which provokes *Z increasing.* This very situation refers to the “effect of unexpectedness”, that should give rise to human’s negative emotions. Vise versa, finding the solution to the problem (D=0) results in momentary decrease of Z, which corresponds to positive emotional splash. Thus, the model (3), (4) seems quite reasonable.

Besides, regulating the ratios of parameters η, χ, and τ^Z in (3), (4) one could provide a desired *temp* of emotional reactions (the analogue of “alertness of cognition» in a living system). This problem deserves further analysis.

D. Specifying the Inter-Subsystem Connections $\Lambda(t)$

Summarizing the previous arguments on correlation between the required activity of specific subsystem (**RS** or **LS)** and the appraisal of the system state, we can set: $\Lambda^{R\to L} = -\Lambda^{L\to R} = \Lambda$ and propose the final hypothesis:

$$\Lambda(t) = -\Lambda_0 \cdot th\left(\gamma \cdot \frac{dZ(t)}{dt}\right), \qquad (5)$$

where Λ_0 being characteristic value of the inter-subsystem connections, γ is the model parameter, which specifies the Λ dynamics.

Note that *hyperbolic tangent* function in (5) corresponds to the step-wise θ-function at γ>>1. This implies that $\Lambda = \Lambda_0 = \Lambda^{R\to L}$ at $dZ(t)/dt$<<0 and $\Lambda = -\Lambda_0 = \Lambda^{L\to R}$ at $dZ(t)/dt$>>0, with Λ being zero at $dZ(t)/dt$=0. Small/moderate variations of dZ/dt around zero provide corresponding oscillations of Λ(t) that represent permanent (normal) “dialog” between subsystems. Besides, the solution to standard problems can be found in **LS** only and commonly does not provide any emotional reaction — here, $\Lambda \sim dZ/dt = 0$ (any inter-subsystem connections are not activated). Thus, this equation fits completely our previous consideration on the psychological role of unexpectedness.

IV. APPLYING THE MODEL TO DESCRIBE THE EFFECT OF STRESS/SHOCK

Let us consider an example of applying this model to reproduce certain observable effect. The effect of “stress and shock”, that emerges when people find themselves in a stressful situation, was investigated for several years by the group of neurophysiologists [20]. Two specific characteristics of electrocardiogram were measured, one of them being an appraisal of vegetative imbalance, another one being the measure of heart-rate variability. It was observed that under small or moderate external impact, people gradually calm down after several oscillations of measured characteristics. But in the case of strong impact, initial excitation changes for *depression* and only after sufficiently long time the person can return to ordinary (regular) reactions. This type of behavior is identified as “*stress*”. Moreover, there are situations called a “*shock*”, when the probationer, after too strong initial excitation, falls down to *deep depression* (*stupor),* and cannot relax independently without medical assistance. In the latter case, the vegetative balance is controlled by the *opiates* (pronounced inhibitors) only, with the variability index comes to zero. It deserves mentioning that the levels of initial excitation resulting in “irregular” regimes of behavior were just individual.

All these regimes could be reproduced within the proposed model by choosing an appropriate parameter set. Let us note that the first attempt to describe these effects was done in [8], where we have used *two different* sets of parameters to reproduce the “normal\stress” and “shock” regimes, respectively. This means that the transition between the stress and shock states was treated as *parametric* modification of the system. Here, we present another version of this model (another choice of parameters), where *all the regimes* could be reproduced within *single* combination of parameters by means of varying the initial conditions. Besides here, the description of the stress-to-shock transition seems to be more interesting and relevant (see below).

In Fig. 2, presented is the phase portrait for the model (3)-(4) where the parameters were chosen to provide the N-shape isoclinic curve dZ/dt=0 with just *two* stationary states. The normal state {$Z=Z_0$, $\mu=0$} corresponds to normal system homeostasis. The second one {$Z=Z^*$, $\mu=\mu^*$} corresponds to anomalous state where the noise is deeply suppressed ($Z^*<0$), and the transmitter imbalance is shifted to deep inhibitor region ($\mu^*<<0$). This state just corresponds to that of the “shock” – this implies deep depression (*stupor*) transient to a *coma*.

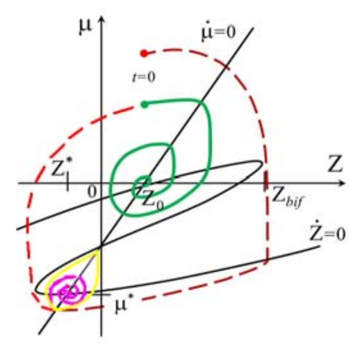

Figure 2. Model phase portrait in terms of “noise amplitude Z *versus* an aggregated transmitter compound μ”.

Normally, the dynamical regime represents *damping oscillations* around the homeostasis point {Z_0,0}. Initial excitation μ(t=0) (imitating an external impact) provokes growth of Z supplied by following decrease of μ down to negative values, which then changes for decreasing Z with μ growth, and so on. Thus, the values of Z and μ gradually (over several cycles) trend to their stable points (solid green curve). But if the trajectory, starting from somewhat larger initial value of μ, would pass beyond some *bifurcation* value Z_{bif}, the dynamical regime changes (dashed red curve). The trajectory falls down to negative μ (inhibitor) values where spends a long time. Then it slowly, over the *depression* zone $Z<0$, returns to regular (oscillatory) mode. This regime qualitatively corresponds to the "*stress*" behavior.

The yellow curve in Fig.2 *separates* its attraction zone from the "normal" behavior mode. It should be stressed that the trajectory could cross the separatrix only occasionally (due to small external impact), thus commonly, the *stress* regime returns to a normal mode and should not result in the shock state. But since at certain stage of the process, the trajectory comes very close to the separatrix, the least impact could result in hitting the shock zone. Thus, this model version enables us to infer that the stress regime is *dangerous* for human beings, since this process includes the stage (just before the stress mode turns to increasing μ values, i.e., to rather normal behavior) when the least external excitation could provoke momentary stress-to-shock conversion. This is the novel model prediction, which could be tested experimentally; certain evidences in favor of this effect were already detected [20].

Since the stationary state $\{Z^*, \mu^*\}$ is stable focus, the trajectory cannot leave the zone of its attraction without certain external (medical) assistance. Thereby, this model could be applied to analyze possible results of use of different medical impacts, such as the adding certain *stimulants* at different stages of the stress process. These researches could lead to pronounced applied results.

The described effects are in good qualitative agreement with the experimentally observed ones [20]. Quantitative correspondence is intricate, since the characteristics that are measured experimentally are close *per se* to $Z(t)$ as a measure of irregularity, and $\mu(t)$ as a measure of mediator imbalance. However, the question of exact correspondence between measured and model variables requires additional analysis.

V. INTERPRETATION OF A SENSE OF HUMOR

Within the presented concept, the sense of humor is interpreted as an ability to adapt quickly to unexpected information with getting positive emotions. This process is illustrated in Fig.3.

Let the incoming information represent a time sequence of symbols that is perceived *consequently* by **LS,** as it is shown in Fig.3. At initial stages, the information perceived is usually not concrete enough to correspond to one *symbol of process* G^2, thus the system makes no predictions. A prognosis could be done when accumulated information enables the subsystem to choose one symbol among others (in Fig.3, "black" symbol at G^2 plate, which has more strong connections than the "green" one, i.e., it corresponds to more "common" process). Then the system *waits* for further detailing the predicted process (this means activation of the "black"-symbol chain at G^1 plate). Up to certain moment t^*, the incoming information ("violet" chain in Fig.3) fits these expectations. At the moment t^*, the prognosis on further information could appear to be *incorrect*, — the next symbol at G^1 plate belonging to "violet" chain, actually is not involved into the "black"-symbol chain, and thus unexpected. Then the system has to appeal to **RS** (down Λ arrow in Fig.3); in this process, the emotions are negative: $dZ/dt > 0$. However, the system may rapidly find a new solution — this implies that there *already exists* the symbol of another process that matches completely both, former and next information ("green" symbol at G^2 plate in Fig.3). This leads to positive emotions ("aha" moment) and hence, switching on the $\Lambda^{R\to L}$ connections (up arrow in Fig.3).

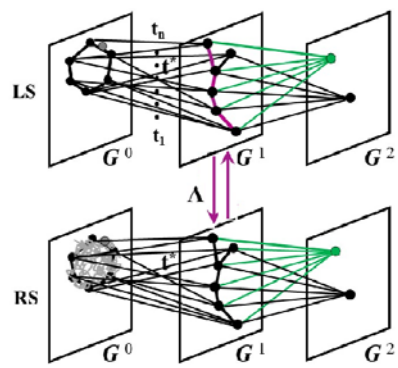

Figure 3. Illustration for the process of perception of incoming information in the well-trained system.

According to this concept, a good anecdote should be a story that, up to certain moment t*, permits a well-known interpretation. The next information block should *not deny* the previous version, but suggest another, also well-known solution. In this case, the system has to return to the turning point t^* and then choose the "true" chain of symbols fitting all the incoming information. The very process of returning and jumping to the true trajectory requires definite specific efforts —so again it leads to the spike of noise amplitude that corresponds to laugh.

Let us stress that this is possible, if the system is reach enough with symbols of processes, i.e., has large enough "repertoire" of various symbols and images. Then this process is rapid, both trends appear to be superimposed: the value $Z(t)$ undergoes abrupt increase-and-decrease ("spike"), that could be interpreted as an analogy to human *laugh* (abrupt involuntary reaction). Thus, we infer that a sense of humor could be inherent to the well-learned system only, just as it is for human beings.

VI. CONCLUSION AND FUTURE WORK

In summary, we can infer that NCA inherently contains the possibility to imitate various human emotions due to involvement of an occasional component (noise) into the cognitive process. Human emotions could be imitated in AI by the noise-amplitude *derivative dZ/dt*. The noise, being presented in the generating subsystem (**RS**) only, provides also *regulating* the activity of two subsystems, what represents an analogue to emotional manifestation. Negative emotions, which are imitated by $Z(t)$ increasing ($dZ/dt>0$) correspond to unexpected incoming information (incorrect and\or undone prognosis); in this process, **RS** should be activated. Vice-versa, a found solution to any problem results in positive emotions and, correspondingly, decrease of noise amplitude ($dZ/dt<0$) – then, only **LS** remains active, while **RS** gets an opportunity to be "at rest". Specific case of abrupt up-and-down jump of $Z(t)$ could be associated with specific emotion (the *laugh*).

Realization of this program in AI could be accompanied by certain sound effects, such as *laugh* in the case of abrupt spike in $Z(t)$ dependence. In addition, variation of the noise amplitude during the process of problem solving could be accompanied by the display of visual "symbols", such as cheery or sorrowful "faces", etc.

This approach opens a wide field for imitation and model analysis of various human peculiar features. This implies that various types of temperament could be associated with certain values of the rest-state noise amplitude Z_0. Also, the model enables us to analyze the dependence of the reaction rate on the ratios of model parameters in (3), (4), etc. Furthermore, the model described the stress\shock effect could be employed for working up new medical-treatment techniques for specific (neural) diseases. All these tasks require further study.

It should be stressed that all of these possibilities emerge from just the human-like architecture of the cognitive system proposed. This implies two combined subsystems in analogy with two cerebral hemispheres, with the interaction between them being controlled by the proposed (original) mechanism of emotional manifestations (noise-amplitude derivative). It is important to accentuate that within NCA, the noise is treated not as annoying and unavoidable obstacle, but as full and required member of each process relating to the generation of new information. In this connection, AI constructed according to NCA provides a unique possibility to study the process of problem solving since here, it is possible to vary the noise amplitude "by hands", thus testing various working regimes. These possibilities are absent in other approaches to AI constructing.

Thus, it is shown that NCA provides a possibility to imitate emotional responses in an artificial cognitive system. The main constructive feature of this approach consists in splitting up the cognitive system into two linked subsystems, one (**RS**) for generating information (with required presence of an occasional component, "noise"), another one (**LS**) for reception of well-known information. It is shown that human emotions could be imitated and displayed by variation of the noise amplitude; this very variation does control $\Lambda(t)$, i.e., switching the subsystems activity. The *sense of humor* is treated as an ability of quick adaptation to unexpected information (incorrect and/or undone prognosis) with getting positive emotions. It is shown that specific human emotional response to the humor (the laugh) could be imitated by abrupt changing ("spike") in the noise amplitude. These ideas require further research.

References

[1] J. E. Laird, "The Soar cognitive architecture", MIT Press, 2012.

[2] E. Hudlyka, "Affective BICA: Challenges and open questions" Biologically Inspired Cognitive Architectures, vol. 7, pp. 98-125, 2014.

[3] A. Samsonovich, "Bringing consciousness to cognitive neuroscience: a computational perspective". Journal of Integrated Design and Process Science, vol. 1, pp. 19-30, 2007.

[4] K. Kushiro, Y. Harada, and J. Takeno, "Robot uses emotions to detect and learn the unknown", Biologically Inspired Cognitive Architectures , vol. 4, pp. 69-78, 2014.

[5] M. I. Rabinovich and M. K. Muezzinoglu, "Nonlinear dynamics of the brain: emotions and cognition" Physics-Uspehi, vol. 53, 357-372.

[6] J. Schmidhuber, "Simple algorithmic theory of subjective beauty. novelty, surprise, interestingness, attention, curiosity, creativity, science, music, jokes". Journal of Science, vol. 48 (1), pp. 21-32, 2009.

[7] O. D. Chernavskaya, D. S. Chernavskii, V. P. Karp, A. P. Nikitin, and D. S. Shchepetov, "An architecture of thinking system within the Dynamical Theory of Information". BICA, vol. 6, pp. 147-158, 2013.

[8] O. D. Chernavskaya, D. S. Chernavskii, V. P. Karp, A. P. Nikitin, D. S. Shchepetov, and Ya.A.Rozhylo, "An architecture of the cognitive system with account for emotional component" BICA, vol.12, pp. 144-154, 2015.

[9] H. Haken, "Information and Self-Organization: A macro-scopic approach to complex systems", Springer, 2000.

[10] D. S. Chernavskii, "Synergetics and Information. Dynamical Theory of Information". Moscow, URSS, 2004 (in Russian).

[11] J. Stirling and R. Eliott, "Introducing Neuropsychology": 2nd Edition (Psychology Focus), Psychology Press, 2010.

[12] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", PNAS, vol. 79, p. 2554, 1982.

[13] S. Grossberg, "Studies of Mind and Brain". Boston: Riedel, 1982.

[14] T. Kohonen, "Self-Organizing Maps". Springer, 2001.

[15] I. Prigogine, "End of Certainty". The Free Press, 1997. ISBN 0684837056.

[16] H. Quastler, "The emergence of biological organization". New Haven: Yale University Press, 1964.

[17] E. Goldberg, "The new executive brain". Oxford University Press, 2009.

[18] D. O. Hebb, "The organization of behavior". John Wiley & Sons, 1949.

[19] L. F. Koziol and D. E. Budding, "Subcortical Structures and Cognition. Implications for Neurophysiological Assessment", Springer, 2009 .

[20] S. B. Parin, A. V. Tsverlov, and V. G. Yakhno. "Models of neurochemistry mechanism of stress and shock based on neuron-like network" Proc. of Int. Simp. "Topical Problems of Biotonics", Aug. 2007, pp. 245-246.