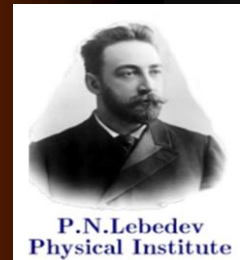


On Modeling the Role of Negative Emotions and the Effect of Panic in an Artificial Cognitive System

Olga Chernavskaya

P.N.Lebedev Physical Institute, Moscow, Russia

olgadmitcher@gmail.com



Nice, France, Oct 10-15, 2020



Abstract

- We continue a series of works on modeling especially the human aspects of cognitive process, such as intuition, the influence of emotions, the role of personality, etc. The Natural Constructive Cognitive Architecture proposed and analyzed in our previous works has an important design feature: the entire system consists of two connected subsystems conventionally corresponding to the cerebral hemispheres. One is responsible for the processing of well-known information, the other is aimed at learning new and creative work. This paper is focused on analyzing the **extreme mode** of thinking process: the effect of panic in creative work (“**throes of creativity**”). It is shown that the regime of panic in an artificial cognitive system could be imitated by **chaotic fluctuations in the amplitude of self-excitation (*noise*) around an *abnormally high level***. It could result in ***insight***, i.e. suddenly finding a solution (**Eureka! moment**), which is accompanied by **emotional burst**. Otherwise, it could lead to a decrease in efficiency (deep long *depression*).

OUTLINE

- Introduction and goal settings
- Main features of NCCA
- Representation of Emotions (E)
- Typical emotional patterns for recognition, forecast and Aesthetic Emotions (AE) = contemplation of Art patterns
- Simulation of *panic* = “throes of creativity”
- Conclusion and discussion

Introduction

- NCCA is used to analyze the extreme mode of the cognitive process – the effect of **intellectual panic** caused by the need to solve urgently (before a *deadline*) certain **creative** cognitive problem (“*throes of creativity*”).
- **Creativity** is tightly connected with the “Explanatory Gap” problem: **Brain vs Mind**.
 - **Brain (B)** = **objective** inf provided by Nature
 - **Mind (M)** = **subjective** inf created inside the cognitive system *itself*
 - **Creativity** = effort to bring a **piece of personal B** into the M and world
- Creative solution is *hidden in the B* (**sub-consciousness**) and could be realized only *occasionally*, due to the neuron’s random self-excitation (= **noise**).
- **Panic** could result in extreme *noise* behavior → increasing probability of finding the hidden solution (**insight**)

Levine J. "Materialism and Qualia: The Explanatory Gap". 1983



Explanatory Gap

Psychology (**MIND**) Neurophysiology (**BRAIN**)

● Consciousness

● Ensemble of Neurons

Emotions:

● Subjective (!) self-appraisal
of current/future state

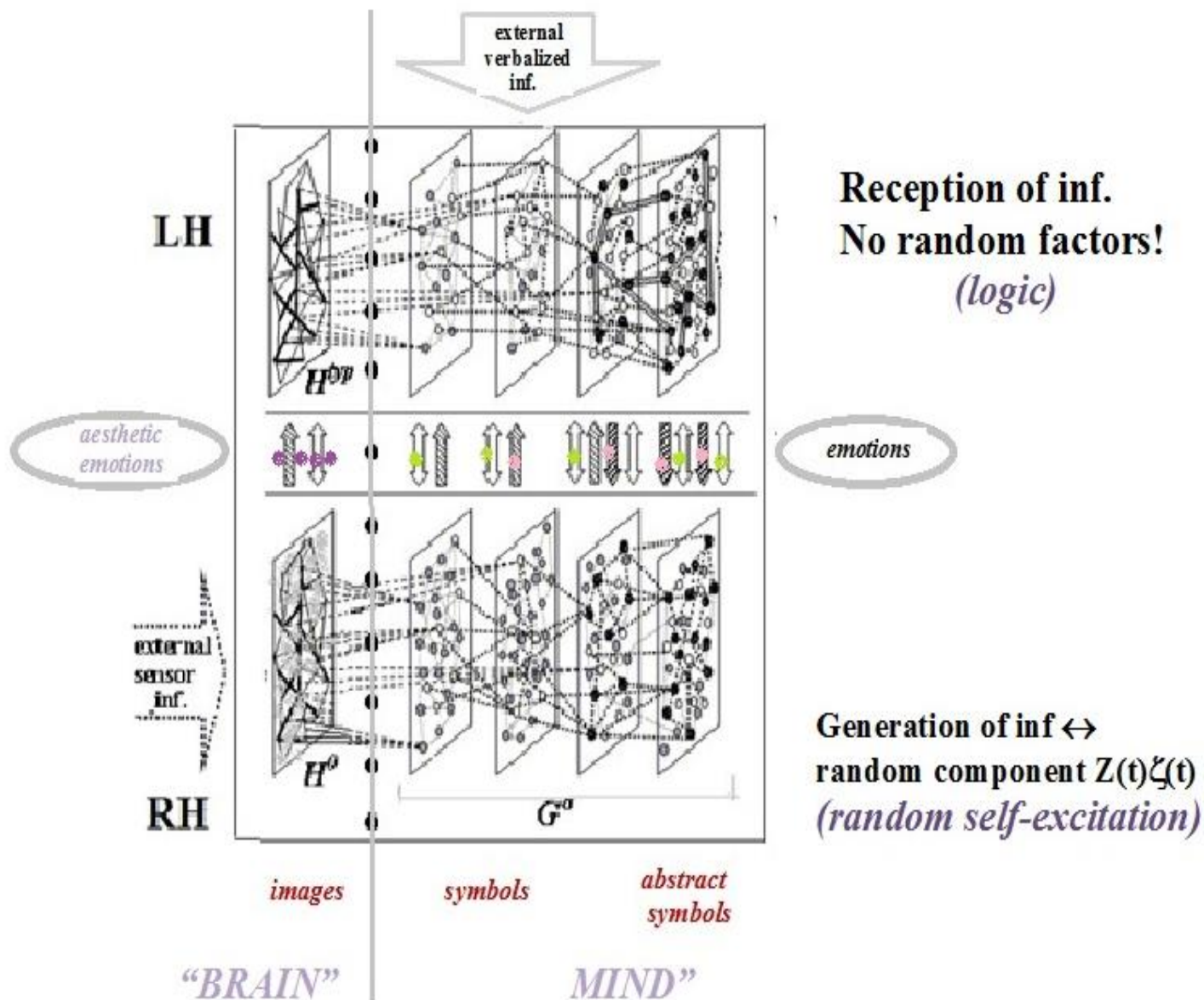
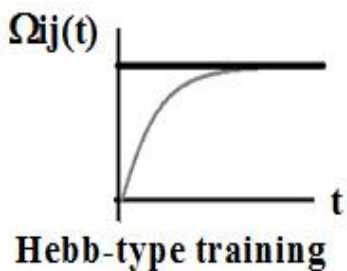
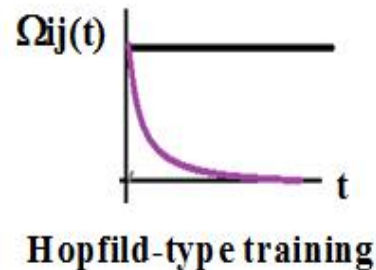
● neural transmitters
(**objective! Measurable!!**)

Natural-Constructive Cognitive Architecture (NCCA)

- Neurophysiology & psychology data
 - *E. Goldberg*: RH \leftrightarrow learning (new inf.)
LH \leftrightarrow processing well-known inf. (recog.)
- Dynamical Theory of Information (DTI)
(Haken, 2000, Prigogine, 1997, Chernavskii, 2000)
 - *generation* of *new* inf. and *reception* of inf. are dual functions:
should be implemented by *2 different subsystems*
 - *generating inf requires noise (random self-excitation)*
 $Z(t)\zeta(t)$ (Z is the noise amplitude, $\zeta(t)$ = random function)
- Neural computing
 - Set of Hopfield-type (*distributed* memory) and *localization=WTA* (Grossberg-type) processors
 - Combined with nonlinear differential equation technique \leftrightarrow
dynamical formal neuron concept

The Natural-Constructive Cognitive Architecture

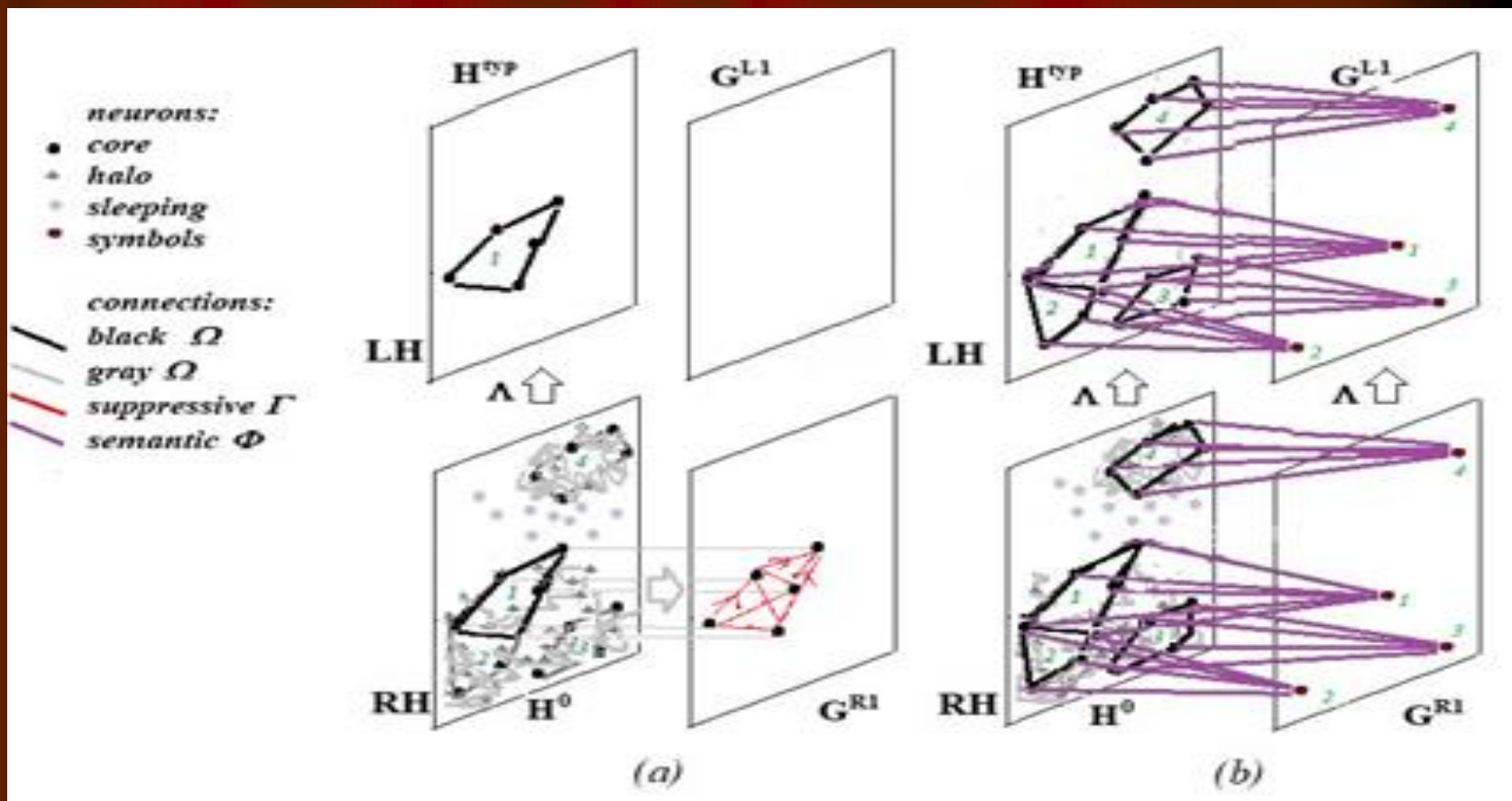
(NCCA) Chernavskaya et al, BICA 2013, 2015



Comments to the NCCA scheme:

- *Splitting up into 2 subsystems (as cerebral hemispheres)*
 - RH \leftrightarrow generation of new inf. (=learning , creating)
 - **Noise is required!**
 - LH \leftrightarrow processing well-known inf = recognition, prognosis, etc.
- Different laws of cons. training:
 - RH \leftrightarrow Hebbian (connection *amplification*)
 - LH \leftrightarrow Hopfield (“**redundant cut-off**”)
- Connection blackening principle : replication RH \rightarrow LH if $\Omega = \Omega_0$
- **Hierarchical structure:**
 - $\sigma=0$: images =sensory inf (objective!) \leftrightarrow **BRAIN !**
 - $\sigma=1, \dots, N =$ *symbolic* inf = convension \leftrightarrow **MIND !**
 - $\sigma > 1$: symbols \leftrightarrow WORDS = **verbalization** \rightarrow consciousness
 - $\sigma \gg 1$: **abstract** (not-sensory) inf. = *symbol-concepts*

Mechanism of NCCA scheme formation: small fragment of basic levels $\sigma=0,1$ (Fig. 2)



- (a): **early** stage: only 1 image in H^0 (RH) is well-learned (“typical”), it is translated to H_{typ} (LH) and G_{1R} , where the symbol is chosen (by **competition**)
- (b): **final** stage: all 4 images became typical and obtain their symbols that form inter-level (**semantic**) connections with image neurons

Comments to the Fig. 2

- **Connection blackening principle**: images are forming in H0 (RH) by Hebbian learning mechanism up to strong (“black”) connections (typical images) and then are translated (replicated) to Htyp (LH) and to G1R for symbol creation (winner-choosing procedure)
- **Core** neurons ↔ typical attributes
 - Provide the base for **symbol** formation
- **Halo** neurons ↔ atypical (inessential) attributes \ rare representations
 - provide **implicit (indirect) associations that are lost at the transition RH→LH**
 - Are hidden in H0 (BRAIN) only
- **“Sleeping” neurons** = never been excited in any cognitive process
 - **not belong** even to BRAIN experience
- **Sub-consciousness** = manifold (variety) of **halo-neurons** along with their weak (“gray”) **connections** = **seemingly unimportant unrealized and non-verbalized hidden personal (Brain) experience**
 - not connected with any symbol = “**out of control**”
 - **something that BRAIN does know, while MIND can't realize**
 - source for **creative** solutions that could be excited only by **noise (=insight!)¹⁰**

Master equations (details in Chernavskaya, *BICA*, 2015)

$$\frac{dH_i^0(t)}{dt} = \frac{1}{\tau_i^H} [\mathfrak{S}_H\{H, \beta_i\} + \sum_{i \neq j}^n \Omega_{ij}^{Hebb} H_j^0 + \sum_k \Phi_{ik} G_k^{R,1} - \Lambda(t) \cdot H_i^{typ} + Z(t) \cdot \xi_i(t)]$$

$$\frac{dH_i^{typ}(t)}{dt} = \frac{1}{\tau_i^H} [\mathfrak{S}_H\{H, \beta_i\} + \sum_{i \neq j}^n \Omega_{ij}^{Hopf} \cdot H_j^{typ} + \sum_k \Phi_{ik} \cdot G_k^{L,1} + \Lambda(t) \cdot H_i^0]$$

$$\frac{dG_k^{R,\sigma}}{dt} = \frac{1}{\tau_G} [\mathfrak{S}_G\{G_k^R, \alpha_k^\sigma\} + \hat{Y}\{G_k^{R,\sigma}, G_l^{R,(\sigma+\nu)}\} - \Lambda(t) \cdot G_k^{L,\sigma} + Z(t) \cdot \xi(t)]$$

$$\frac{dG_k^{L,\sigma}}{dt} = \frac{1}{\tau_G} [\mathfrak{S}_G\{G_k^L, \alpha_k^\sigma\} + \hat{Y}\{G_k^{L,\sigma}, G_l^{L,(\sigma+\nu)}\} + \Lambda(t) \cdot G_k^{R,\sigma}]$$

$$\frac{dZ(t)}{dt} = \frac{1}{\tau_Z} [a_{Z\mu} \mu + a_{ZZ} Z + F_Z(\mu, Z) + \Theta(Z, H, G_k^q)]$$

$$\frac{d\mu}{dt} = \frac{1}{\tau^\mu} [a_{\mu\mu} \mu + a_{\mu Z} (Z - Z_0) + F_\mu(\mu, Z)]$$

$$\Lambda(t) = -\Lambda_0 \cdot th\left(\gamma \cdot \tau^Z \cdot \frac{dZ}{dt}\right)$$

Comments to the sys. of Equation : 1

- H_i, G_i – variables representing i -th neocortex neurons (Fig. 1)
 - $H \leftrightarrow H$ -type proc. = *distributed* memory (*images*)
 - *Objective individual* inf (sensor signals) on real objects ever presented
 - $G \leftrightarrow G$ -type proc. = *symbols* of previous-level images + *generalized images* (=image-of-symbols)
 - *Subjective* inf created inside the system itself
 - *Scaling* = the same formation principle at any σ level
- σ = level of hierarchy
 - $\sigma = 0 \dots 1 \leftrightarrow$ *virtual border btw. “Brain” and “Mind”*
- **Noise:** $Z(t)\xi(t)$: $Z(t)$ = amplitude, $\xi(t)$ = random function (Monte-Carlo)
 - *presenting in RH only*
- Λ : *inter-subsystem connections* \leftrightarrow “*corpus collosum*”: serve to provide “dialog” **RH** and **LH**
 - $\Lambda = + \Lambda_0 (R \rightarrow L) = \Lambda$; $\Lambda = - \Lambda_0 = \Lambda (L \rightarrow R)$: refers to all eqs. ¹²

Comments to the sys. of Equation : 2

- **Bottom block = variables $Z(t)$, $\mu(t)$ = refer to *Emotions***
 - $\mu(t) \leftrightarrow$ “*deep B*” = **effective composition of neurotransmitters (stimulant – inhibitors)**
 - $Z(t)$ = the “*tool*” for self-appraisal = “***emotional temperature***”
 - Z_0 = “**normal temperature**” \leftrightarrow value necessary for normal system’s functioning (*homeostasis*)
- **Final eq. : $\Lambda(t)$ = activity of RH\LH subsystem is controlled by emotional tool $-dZ/dt!$**
 - **unexpectedness (incorrect prognosis) \leftrightarrow negative E \leftrightarrow RH activation is necessary! (= mobilization)**
 - **finding a solution \leftrightarrow positive E \leftrightarrow LH only (relax)**
 - **NB: derivative could be either (+) or (-) !**

noise amplitude $Z(t)$: typical patterns in solving various problems

- **Recognition** (iteration process): Fig. 3a

RH puts forward hypotheses, LH tests them, etc.:

dumping oscillation around normal value Z_0

- **Prognosis** (prediction):

- the same in normal mode

- in the case of **joke** (sudden unexpected but still familiar inf that switches to another prediction):

sharp rise at t^* that is immediate changing by **fall**

down (new solution found) \leftrightarrow **laugh** (Fig. 3b)

- **Aesthetic Emotions** (contemplation of Art objs) :

goosebumps \leftrightarrow $Z(t)$ "vibration" around Z_0

(Fig. 3c)

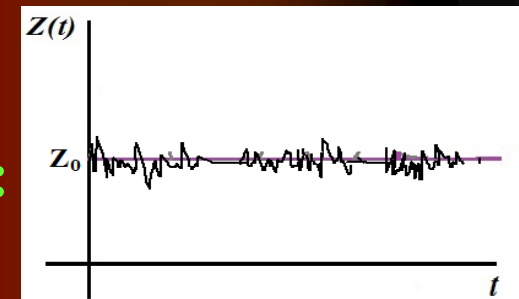
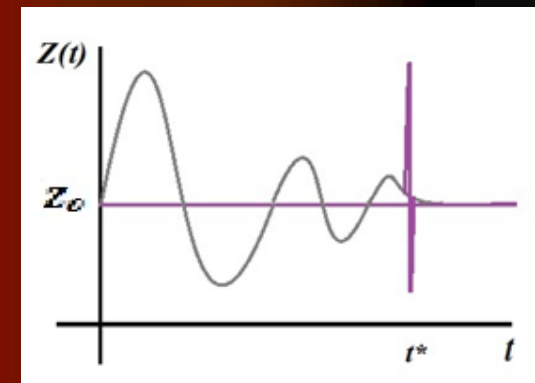
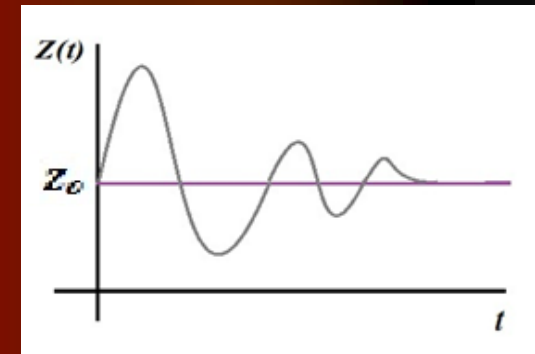


Fig. 3

Intellectual panic mode: *throes of creativity*

- Panic is characterized by unpredictable (*chaotic*) behavior and sudden *jumps* in mood [4]
- Intellectual panic could be caused by necessity to solve some *creative* problem *urgently*
 - e.g., before certain *deadline*
 - *could arise in any creative work: Art as well as Science*

Simulation of intellectual panic mode

- Creative work requires *noise amplitude* $Z(t)$ *increase* for *extracting new* solution from the halo-experience H_0 (sub-consciousness)
- When Z exceeds some critical value $Z^{**} \gg Z_0$, system falls into *chaos*: chaotic jumps around abnormally high value results in noise dominating in RH and *mixing all known images*
- It could result in either:
 - sudden solution (*insight*) at t^{**} that is accompanied by *emotional burst* similar to laugh (Fig. 4a)
 - deep long depression if $Z=0$: the system *can't neither perceive nor generate new inf* (Fig. 4b)
- At $Z > Z^{**}$ the probability of *waking up "sleeping" neurons* can provide new *implicit associations* that could lead to new solution unexpected for the system itself even at the Brain (*halo*) level! (see Fig. 2) \leftrightarrow *enriching the Brain (sub-consciousness) experience*

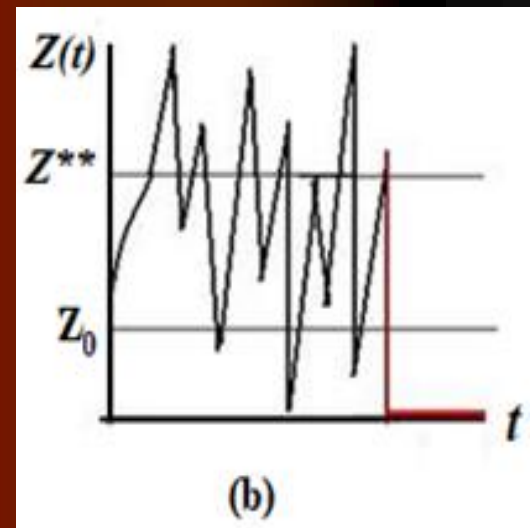
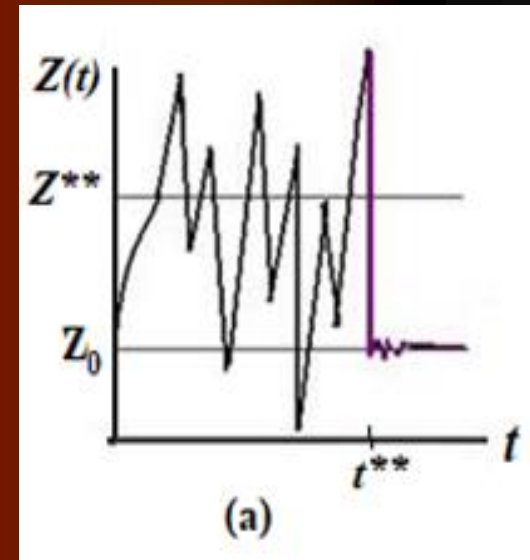


Fig. 4 To 16

Conclusions

- *Negative emotions* (increasing self-excitation amplitude $Z(t)$) represent the mobilization of the system's resources since lead to increasing possibility to find (occasionally, due to *noise*) *desired hidden solution*
- Extreme case = intellectual *panic* \leftrightarrow *chaotic jumps in noise amplitude $Z(t)$ and mood*
- Panic requires significant amount of energy for the neurotransmitter production and thus could not last long
- It could result either in
 - *insight* solution
 - deep long *depression*
- Panic is not good for human health (it is better to escape it), but if succeed, one gets an

ineffable emotional reward (Eureka!!)

Summary

- the intellectual panic (*throes of creativity*) could play *even positive* role in creativity, if results in generation of new information enriching the individual “Brain” (sub-consciousness) experience
- remember Nietzsche: “*You have to have chaos inside you to give birth to a dancing star*”



**Dmitrii
Chernavskii**

Feb 24 1926 –
June 19 2016

*To the memory of my
father, coauthor, and
close friend*

Thanks for attention

References

1. Goldberg E.: The new executive brain. Oxford University Press, 2009
2. Goldberg E.: Creativity: The Human Brain in the Age of Innovation. Carlisle, NY (2018).
3. Levin, J.: Materialism and Qualia: The Explanatory Gap. *Pacific Phil. Quarterly*, 64(4), 354–361 (1983).
4. Quastler H.: The emergence of biological organization. Yale University Press, NH (1964).
5. Solso, R.: Cognitive psychology. 5th edn., Needham Heights, MA: Allyn & Bacon (1998).
6. Chernavskaya O.D. et al. : An architecture of thinking system within the Dynamical Theory of Information. *BICA* 6, 147-158 (2013).
7. Chernavskaya, O. D. et al.: An architecture of the cognitive system with account for emotional component. *BICA* 12, 144—154 (2015).
8. Chernavskaya, O., Rozhylo, Ya.: The Natural-Constructive Approach to Representation of Emotions and a Sense of Humor in an Artificial Cognitive System. *IARIA Journal of Life Sciences*, 8 (3&4), 184—202 (2016).
9. Chernavskaya, O., Rozhylo, Ya.: On the Possibility to Interpret Aesthetic Emotions and the Concept of “Chef-D’oeuvre”. In: 9th Int. Conf. COGNITIVE 2017, pp. 24—31 (2017).
10. Chernavskaya, O., Rozhylo, Ya.: On the Possibility to Interpret Aesthetic Emotions and the Concept of “Chef-D’oeuvre” in an Artificial Cognitive System. In: 9th Int. Conf. COGNITIVE 2017, pp. 24—31 (2017).
11. Chernavskii, D. S.: The origin of life and thinking from the viewpoint of modern physics. *Physics-Uspekhi* 43, 151-176, (2000).
12. Hebb, D. O.: The Organization of Behavior: Neuropsychological Theory. Wiley & Sons (1949).
13. Hopfield, J. J.: Neural networks and physical systems with emergent collective computational abilities. *PNAS* 79, 2554—2558 (1982).
10. Grossberg, S.: Studies of Mind and Brain. Riedel (ed), Boston (1982).