E-LOGOS/2000

ISSN 1211-0442

Karel Pstruzina: **Do Robots need Consciousness?**

Abstract:

The answer on the question in the topic of my contribution depends on other question. What is purpose why we construct robots?

Possible answers are:

- We construct robots as a bondsmen;
- We construct robots as the lords.

I claim that if the robots are bondsmen then they need the consciousness but if the robots are only the lords then the consciousness is something unnecessary.

Though my claim sound rather perplexed, I think it is deeply rooted in Hegel's passage Lordship and Bondage from the Phenomenology of Spirit. I think it could be very inspired for elucidation both evolution of consciousness and future investigations in the domain of consciousness including the robotics.

Hegel wrote that:

Self-consciousness exists in and for itself when, and by the fact that, it exists for another (Hegel p. 111).

It means man strives for recognition by the other man and therefore he undergoes *a life-and-death struggle*, because only through such struggle man can approve himself as independent and *pure being-for self*.

The robots as a lord can function without consciousness. They do need not approve themselves by the means of suppression other being.

Robots can be constructed with the map of their environment and therefore they can reflect concrete situation in which they are but it is long way to conscious being.

To be conscious it means that something like map of environment is previous to percepts that we scrutinize the world through our inner thought's model of world, through our opinion how the world is. We continually expect what will be in the next moment and we are prepared for it. But robots can function very productively out of such mechanism.

If man wants to be lord then he needs the conscious robots. Man long for recognition as *independent consciousness whose essential nature is to be for itself (Hegel p.115).*

Unconscious robot is only machine and it is not enough for our self-consciousness and our self-identity.

Karel Pstruzina: **Do Robots need Consciousness?**

There are many questions about computers, robots, or units of artificial intelligence.

Turing's question: Does can a machine think (?) is most known.*) The question in topic of my contribution is more complicated because it is question on the consciousness, not thinking. A consciousness is more sophisticated phenomenon then thinking is and the answers therefore can very differ, as well.

I characterize thinking as the dis-course with intentional contents of mind. The thinking can deal with concepts, tones, symbols, a numbers, an imaginations etc., but every time with something. The thinking is the blank flux; it carries and mixes concepts, tones, symbols, numbers, etc., but it is under them as something autonomous and independent. Thinking is not the demonstration of being but thinking is spontaneous. Thinking is autonomous in the sense of self-sufficient reason and it is irreducible. It is because the thinking is simple movement (similarly to computation). Thinking is articulated by that that thinking is thinking "about something" at every moment. It could seem that thinking is tied on being and that thinking is only modus of being and that thinking pursuits the differentiation inside of being but my characterization is antipodal. Thinking constitutes the essences because the thinking reflexes the forms of being. An identity and negativity are fundamental operations that are carried out by thinking. It means thinking differs by saying: it is that (it is like that); or it is not that (it is not like that). And by such operations are realized the generalization and abstraction and these operations are a base for constitution of essences, too.

Thinking differs from consciousness. It is only one of the part of consciousness, for my opinion. The consciousness disposes also by the emotions, by imaginations, by pieces of knowledge, by

experiences, by memory contents, etc. Consciousness is also deeply rooted with historical personal development of the individual and the society and culture.*)

I found out about 150 definitions of consciousness. They start with identification consciousness with short term memory (Minsky*)), or short term memory with delay (Pribram*)) to definition consciousness as superordinate term that cover all mental states, events and processes that are experienced (Flanagan and Dryden*)) or identification consciousness with the internal integrative system (Sommerhoff*)). But my question is not oriented to definition of consciousness but to robots and whether robots need consciousness. The answer on this question depends on other one, for my opinion. Especially on the question what is purpose why we construct robots?

Do we want to construct robots as our partners that will be independent and autarkic creatures or we want to construct robots as servants that will make easier our life? Or by the words of Hegel do we construct robots as a bondsmen or as lords?

I claim that if the robots are bondsmen then they need the consciousness but if the robots are only the lords then the consciousness is not needed. Though my claim sound rather perplexed, I think it is deeply rooted in Hegel's passage Lordship and Bondage from the Phenomenology of Spirit.*)

Hegel's conception is founded by his approach to self-identity. The self-identity is mostly interpreted as something what is closely tied with the memory, especially with this part of long-term memory that includes episodic memory. E. Tulving*), M. Warnock*) and others wrote about. But Hegel's attitude differs.

Hegel thinks self-identity leans on the recognition by other people. Other people must respect me as an independent and autarkic creature. The man is held only from himself. The man by every doing is as if speaking this is me. I am such that I can transform my surroundings in way that I want to do. I recognize myself by my deeds. I can see what I am only if I can see the result of my activity. What I am thinking about me it is entirely feeling but not recognition and I need recognition of myself. Therefore I must sign my surrounding by myself. But it is not sufficient. I must recognize myself not only by the transformation of my surrounding but by other people as well.

Other people do the same. They long for recognition also and by this way the activity of one man encounters with activity other man. And it is the reason why they undergo *a life-and-death struggle*, because only through such struggle man can approve himself as independent and *pure being-for self*.

Hegel writes:

The presentation of itself, however, as the pure abstraction of self-consciousness consists in showing itself as the pure negation of its objective mode, or in showing that it is not attached to any specific existence, not to individuality common to existence as such, that it is not attached to life. …

And Hegel continues:

Thus relation of two self-conscious individuals is such that they prove themselves and each other through a life-and-death struggle. They must engage in this struggle, for they must raise their certainty of being for themselves to truth, both in the case of the other and in their own case.)*

What does mean? It means for my opinion that the activity of man must be founded only by his or her own will. If we act on the base of dictation other man then it is of course our activity but it is not our intention. We are acting only what other man wants. And therefore we do not recognize ourselves. Only if our activity is performance of ourselves or of our will then we can speak: it is what I am, what I want.

How it is when we are considering about robots.

The robots as a lord can function without consciousness. They do not need approve themselves by the means of suppression other creatures. They can be very functional out of conscious reflection of their surrounding. They can possess the specific receptors for very different stimuli from the outside that enable to diagnostic concrete things and to orientated themselves in the surrounding. They can possess the feed back receptors by which they can control their activity, as well. They can possess also some map of surrounding for that they are constructed but their form of perception of outside world will be very differ from the way how man is perceiving the world.

If a man perceives the world then his inner thought's model of the world precedes the percepts. There are differences above all in the process of perception during natural aging. The crucial points is, in all probability, puberty when the thalamus as a filter of sensations gets under the control of the cerebral cortex. After puberty perception is not prior to the mind, but on contrary our thinking chooses among the outer stimuli. Vygotsky's thesis is:

A child thinks as it perceives; on the contrary an adult perceives as he thinks;

This, is a summary of sophisticated processes that happen in reality in our brain during aging.

The brain processes differ if a child or adults keeps for example a book and perceives it. A normal adult has an abstract model of a book before his perception. This abstract model of book arises from habituation. It is by such process that we carry out abstraction from very often a repeated perception. By this abstraction we can distinguish both; what is among our percepts invariable and essential and what general properties or features belong to the set of that things. Similarly we have an abstract model of the various things and I label them endocepts.

This is Arieti's term and I prefer this one. K.H. Pribram uses the term *neuron's model of world*; W. Penfield prefers the term *pattern*; D.H. Hubel speaks about *coding's processes*; B. Russell calls it *schema*; and J. Fodor uses the term *prototype*, etc.

Though endocepts are closely related to awareness of percepts there are differences between awareness of our percepts and all sensations that our brain has been recorded. We are consciously aware only a limited amount of these percepts at any moment. Endoception is the opposite of

perception. It is an inner recall of our life through the world. The endocepts is also a storage of all contents of thinking that we carry out and by which we scrutinize the world. The endocepts includes also expectations, values, long-term memory traces, intentions, etc. But an endocept could not be identified with subconscious structure as it is displayed by a Freudian-type analysis. Rather it comprises large systems of past experience, images which do not currently release actions, are not easy to express in the words but are felt as dispositions to thinking.

We can distinguish a thing only when we compare it with endocepts, with the memory traces of previous percepts. We conjecture or we make hypothesis "what it is" at every moment of our perception. And it is thinking which confirms or refuted our conjecture. Thinking makes a zig-zag course between endocepts and the actual sensations and confronts these actual sensations from the criterion of novelty. It means our thinking traces if something new is in our percepts. If it is not, if percepts comport with our endocept for example of a book we do not really perceive a book for a long period when we read it. Therefore our brain chooses the strategy of confirmation of our endocept of a book.

It is very similar in our quotidian life. We are living in a relatively stable environment and it would be very difficult for us to reflect the whole of our environment at every moment. But in the case that the world is some way that we do not expect it to be, of course, we perceive the world too. We instantly give it attention if something new is in our actual perception. For example, a book is damaged, or there are misprints in a book and so on. In this case our thinking reflects differences between endocept and actual perception and we either complete, if it is possible, our endocept with novelty, or we reconstruct it, or we must form a new one.

These processes depend on the quantity and the consequence of the inclusive novelty in the actual sensations. Sometimes new information has only a virtual character, when new stimuli are small or they are not frequent. We shift aside such information into periphery of our consciousness and they make a latent agent. Similarly turbulent information are not able to create a new endocept because they bring too many new stimuli, they are confused and call up chaos. Such information we push aside to the periphery as well. These are two bounds and between them is a possibility for creation of new endocept from percepts.

Thinking is here still in the level which, in tradition of German philosophy, is labeled as "das Verstand". Thinking at this point works with the material object and therefore it is observable thinking.

But thinking can work otherwise as well as with concepts. In such processes thinking does not work with sensations, percepts and images, but with ideas. The latter are most frequently externalized as words and represent a content of thinking. Therefore man can work with ideas as with things in a practical life. Man can combine or compose ideas and so he can create new ideas. Thinking is here on its own field and can evolve concepts.

I claim the creation of new ideas can be described not only by the means of Hegel's negation or logical inference but also by the means of generators and inhibitors of thinking which can act both way; discursively and divergently.

The endocept is opposition of percept. The endocept expresses the active evocation of world in the mind of man. I think that endocept differs from experience because the endocept include not only our imagination but also the concepts.

K.H. Pribram claims that the brain is depended on the information from the outer world when it create the neuronal model of the environment but the brain is independent when it elaborates them. Pribram utilizes the result from the Sokolov's research. The neuronal model represents the memory mechanism. Pribram and Sokolov suppose the neuronal model is tuned on the input information. We have the neuronal model of environment before the outer stimuli begin take effect on us. They are differences among scientist what part of brain generates such process. Magoun supposes that it is mesencefalon; Penfield's hypothesis is that the generator is in diencefalon; and Pribram thinks that it is brought about by limbic system, especially by amygdala and by prefrontal lobe of neo-cortex. Goleman at his work "Emotional intelligence" wrote about the amygdala and hippocampus which has responsibility for generating of emotional and epical memory.

The endocepts precede to percepts. The endocepts are set up before we register the percepts. Such pre-set has its base in the habituation, it means on the base of repetition in our perception. This mechanism of perception guarantees that our brain will be not overworked. The stimuli that we anticipate are not recorded by brain, or more precisely they are not fully aware. When we perceive whatsoever object this object is only confronted with neuronal model of environment, with the endocept, which is set up before our perception. The percept usually affirms endocept. If the percept is not in concordance with endocept then we are aware it and gradually incorporate such novelty into endoceptive structures. When we records information we endeavor to bind them to the endocept. If the information is in coincidence with endocept then it completes endoceptive structure or affirms inner thought's model of world. Some percepts can be new and strange, they do not need agree to nothing what we know, what we absorb. At this case we record them at once. We are aware that something among the percepts is new and the stream of our attention takes them into center of our awareness. But some of such new percepts we can shift off to periphery of mind and they cause only as a latent agent there.

W.J. Freeman investigated how the neuronal model of world or endoceptive structure is generated. He used the PET method. The results of his research show that the impulses are more chaotic when the men is relaxed than he perceive.

This conception is from the point of view philosophy very important. The perception is not understood as adverse transfer of picture (as photography) but it is process by which the outer stimuli are complete by neuronal model of world by endocepts. At the process of perception we does not only record the stimuli but we construct the objects on the base of previous perception.

Notes:

Turing, A.M.: Computing machinery and intelligence. In: Mind LIX/1950

Rose, S.: Brains, Minds and the World. In: From Brains to Consciousness. (Edited by Rose, S. Priceton University Press 1998, ISBN 0-691-00460-2, p. 5 - 6

Minsky, M.: Minds are simply what Brains do. In: http://www.leaderu.com/truth/2truth03.html

Pribram, K.H.: The Deep and Surface Structure of Memory and Conscious Learning. In: Mind and Brain Sciences in the 21st Century, edited by R.L. Solso. A Bradford book 1999, ISBN 0-262-69223-3. (p. 130)

Flanagan, O., Dryden, D.: Consciousness and the Mind: Contributions from Philosophy, Neuroscience, and Psychology. In: Methods, Models, and Conceptual Issues, edited by Don Scarborough and Saul Sternberg. A Bradford book 1995, ISBN 0-262-65045-2. (p. 167)

Sommerhoff, G.: Consciousness Explained as an Integral Interating System. In: Http://

Hegel, G.W.F.: Phenomenology of Spirit, translated by A.V. Miller. Oxford University Press 1977, ISBN 0-19-824530-0 (p. 111 – 119)

Tulving, E.: How many Memory Systems are there? In: American Psychologist, 40/1985 (p. 385 – 398)

Warnock, M.: Memory, London, Faber&Faber Ltd. 1987 (p. 4 – 8)

Hegel, G.W.F.: Phenomenology of Spirit, translated by A.V. Miller. Oxford University Press 1977, ISBN 0-19-824530-0 (p. 113 – 114)

Vygotskij, L.S.: Thought and Language. Moscow 1934, (p. 124)

Arieti, S.: Creativity, the Magic Synthesis. Basic Book, New York 1976, ISBN 0-465-01443-7 (p. 53 – 65)