Commentary 21 on
Karl Jaspers Forum, Target Article 22, 2 November 1999

MENTAL ACTIVITY AND CONSCIOUSNES...
By Timo Jarvilehto

Commentary 20
THE ORGANISM-ENVIRONMENT AND ROBOT-ENVIRONMENT THEORIES
by Chris Malcolm

DIGITAL vs. ANALOGOUS CONSCIOUSNESS

by Paul Jones

10 May 2000

Chris Malcolm's article is a remarkable example of the convergence of different disciplines at approaching the necessity of not only declaring (or denying) the existence of the mind, but also attempting to develop a sound strategy of its practical implementation, at least to a limited extent. As expected, people dealing with a particular material would find their specific way to organize it into a something possessing a glimpse of consciousness – still, there are objective laws of subjectivity underlying any kind of reason, and one has to satisfy them in order to produce a thinking creature, either of flesh and blood or of silicon chips.

This is yet another illustration of the general principle stating that praxis is the only criterion of truth – or, rather, the clearest manifestation of it. When people have to do something, they forget about any subjectivism, spiritualism, positivism...and other possible –isms – and start to act in accordance with how it is all ordered in the world; this naturally induces thoughts adequately servicing the required mode of action. Of course, one may stick to a wrong theory and spend one's life in vain occupations – this might be regarded as a primitive "bad example" mechanism of inducing the others to switch to a different track.

The resemblance of the history of human-made tools to the development of organic life is no wonder to those who read vol. I of "The Capital" by K. Marx, where this similarity was attributed to the same "objective logic" behind the both processes. More discussion of the relevant problems can be found in Marx' manuscripts of the early 1860's, which, however, do not seem to be commonly known to the learned audience.

The binary-tree illustration of the traditional approach to the description of the mind, as provided by Malcolm, beside it certain ability to serve as a basis for systematic comparison of different schools, is useful in yet another respect: it shows how lack of hierarchical reasoning can lead to the questions that cannot be answered without arbitrary decisions about where the "true" distinction should be made. However, no distinction can be complete without specifying the way of linking the distinguished things to each other – as soon as we specify the level of hierarchy we deal with, any distinction at all is thus made true, and there is no need to argue with other viewpoints, which are, in their turn, valid under specific conditions. The discussions between different schools based on a binary discrimination mode of reasoning might be assimilated to argument between two people, one saying that 3 = 2 + 1, and the other opposing this statement, saying that 3 = 1 + 1 + 1, both disputants blaming a heretic, who claims 3 = 1 + 2 to be the ultimate truth; there is also a group of "agnostics", who say that 3 = 3, and nothing can be said beyond that absolute totality that would not be an entirely artificial construction; those who admit that 3 = 4 – 1, or even 3 = –2 + 5, may be commonly believed to be mentally insane, nothing to say about "mystics" deriving 3 from such an irrational entity as sqrt(3), multiplying it by itself! In fact, 3 can equally be defined through any of these procedures, and a million others, which all, in their unity, constitute the hierarchy of the number 3, possessing that many properties, forms or representations (the different unfoldings of the hierarchy).

Most people would agree that a conscious being is different from a consciousness-devoid animal, or an inanimate thing. However, few people can say anything coherent about where is the difference. Comparing the two "intelligent" balls of Malcolm – one with a mechanical device inside, and one with sensors and reversible motors – that can exhibit superficially the same behavior, I cannot find any essential difference: surface deformations and elastic forces can be considered as a kind of sensor on exactly the same grounds as metal contacts, piezoelectric elements, chemical triggers or photodiodes – it is their ability to reflect certain aspects of the object situation that matters. It does not matter whether the information about the state of the sensors is transmitted to effectors through some electric circuitry, or via strain waves in a solid body (which virtually have electric nature too). That is, the presence of specialized sensors cannot be said to be a necessary prerequisite of any intelligence – though the very ability to reflect the world, regardless of the particular ways of doing that, is the most fundamental property of any material thing that underlies the possibility of eventually developing consciousness in certain kinds of creatures, not necessarily looking like humans on the Earth. [V. Ilyin, *Materialism and Empiriocriticism*, 1909]

The assertion that digital computation is more favorable for the development of consciousness that analogue computation does not seem indisputable. I would not say that analogue computers are more limited to the kinds of computation they can perform than digital processors. For any particular problem, one could design an analogue system that would efficiently solve it, with reasonable accuracy. Moreover, there may be adaptive systems that can change the modes of their behavior depending on the situation, thus mimicking algorithmic computation. On the other hand, I would not overestimate the potential of digital (discrete algorithmic) computation: Turing's theorem dealt with a very limited class of algorithms, and a "general-purpose" Turing machine is bound to get stuck in a just a little bit more serious problem, as the Goedelian line of reasoning suggests. It is a combination of analogue and digital computation (or, rather, continuity and discreteness in general) that can provide flexibility enough to support conscious behavior.

It may be amazing to observe how the ideas once clearly expressed and well established get flattened with time, to suddenly become re-discovered in a quite different social environment many years after. While Malcolm is engaged in demonstration of how memory can originate from the marks left by a living creature in its environment, I recollect the works of 1920-30s by L.Vygotsky, where the theoretical derivation and experimental investigation of the formation of consciousness from certain kinds of productive behavior starts with that very explanation of memory as interiorized activity of making traces in the world. Dozens of former Soviet psychologists worked in that direction, providing the same arguments as Malcolm's, along with many other considerations – however, all that work had to be forgotten due to misfortunate political circumstances, so that modern thinkers have to make their own way through the wilderness to find the same answers.

Malcolm:
"The fact that the brain is involved tempts us to think that the brain is more involved than it is"

Similarly, a computer cannot work without the processor, and this may make one think that the processor is the principal part of the computer. However, no processor can operate as such without a proper environment, including other processor chips (controllers etc.) as well as "lower-level" devices like power suppliers, connectors etc. One could observe that replacing one processor with a quite different one, or even with a many-processor system, with appropriate adjustments made to the operating system, would retain basically the same functionality, as long as the peripherals are the same, which heavily undermines the idea of the processor's priority in computing. In the same manner, in the early age of radio, the vacuum tube might be thought to be indispensable for signal amplification – now we know about transistors, digital logic chips etc.

Moreover, no computer can operate as such without being *intended* to do so, that is, without being fed an appropriate sequence of instructions and prompts from its exterior. For instance, I turn on my laptop, and it starts Windows'98 and initializes a number of applications I typically use, as if it knew what to do by itself; however, it is I who configured it that way, and it would stay idle after startup until I tell it what is to be done next. Well, one could suggest less trivial examples like a server communicating with the other computers through the Internet in the 24x7 regime and performing a lot

of tasks of all sorts without any human interference; this superficial autonomy can fool nobody who ever dealt with server administration. However, this latter example brings us closer to artificial reason than any trick with artificial intelligence. If computers are ever to develop a kind of consciousness, this can only occur through their integration in a kind of society, in which the sense of any single computer's existence comes from its place in the whole computer network. No computer will become conscious on itself, without being *meant* to become conscious. It does not really matter whether that will be the "society" of computers of the same or different kinds, or computer-human communication.

The problem of the relations between thought and language has always been considered very important for understanding consciousness. However, many thinkers narrowed it to the problem of the expressibility of the inner psychological processes or subjective states in words, which lead to conceptual difficulties and false conjectures. Language can in no way be reduced to words, and even less to the spoken word, speech. The same things can be communicated by different means, like behavioral hints, gestures, facial expressions – or even refraining from any perceptible action. Words and speech are related to all those modes of communication like digital to analogue processing, or credit card payment to direct goods exchange. Here, the old theory by K. Marx comes up in a probably unexpected aspect: the development of language as a universal tool of human communication follows exactly the same route as the development of money exchange from the primitive forms of exchange to complex indirect exchange systems, and the adherence of some layers of society to the idea of the priority of the word (and thought) rather than physical action can be assimilated to the banker's belief that the capital is the motive force of industry. Yes, it is, in a society of definite type – and it is not, in some other kinds of society. There are different levels of thinking, and words may dominate on some of them, while non-verbal communication plays the role of language in other domains.

Philosophy is replete with various mental constructions designed to illustrate one or another ideological standpoint. Two hundred years ago, that was the Robinson abstraction; today, there is much ado about zombies. Philosophers imagine dull robots that mimic human behavior and say that there is no way to distinguish such creatures from conscious people. In such mental games, it is implicitly assumed that, first, consciousness is a "local" property confined to a single organism and, second, consciousness can only be detected by some specifically "conscious" operations that cannot be it reproduced by non-conscious beings. Both assumptions are wrong, and following this line of thought is no better than trying to find a molecule of life that would drastically differ from any other molecule in its chemistry. In reality, there are no operations that can be performed by humans and never by an artificial device – any single operation, and any sequence of operations, can be reproduced by a robot if it has once been performed by a human; this, however, won't make the robot conscious. This is one of the distinctive features of a conscious being, the ability to make non-conscious things do something for humans, thus releasing their time for more creativity.

The keyword to understanding consciousness is *universality*. One way or another, people can do anything, reach anything, perceive anything. Using Spinoza's words, a conscious body can build its motion following the pattern of any other body. Consequently, deprivation of productivity leads to lack of reason. That is why Asimov's "laws of robotics" can only apply to machines, and never to conscious beings. You cannot just impose the rule that no robot can make harm to humans, without explaining why and to what extent this rule should be applied. Otherwise, one will deal with a slave, a machine, albeit walking like men. In the sphere of social relations, universality means freedom.

The important aspect of universality is the ability to exercise self-control. This means that any "absolute" freedom knowing no restrictions is identical to the complete absence of freedom, ultimate slavery. One cannot be said to be conscious without directing one's activity to the objectively open routes, rather than trying to force the world to something impossible; in a conscious person, this does not imply lack of creativity – inversely, without being objectively justified, creativity degenerates into primitive field behavior more appropriate for animals.

To be different from the rationality of the robot that can work with humans and communicate with them in human language, behavior must be properly motivated. This is what T. Jarvilehto stressed in his definition of consciousness as human cooperation directed to a common result beneficial for

everybody. However, Malcolm is right indicating that consciousness does not originate from mere cooperation and requires something else that could become a kind of *shared* experience, allowing one person to be mentally put in the place of another, hence forming the very idea of personality. This is what can never be found in animals, who live entirely for themselves, never trying to act and feel like the others. The product of conscious activity is *intended* to be used by the others – and its producers in particular, just for instance. Social memory described by Malcolm is a part of this shared (objectified) experience. The origin of consciousness is not in mere self-awareness, the ability of feeling one's own body and its motion – rather, it is the ability to share the activity of the others that makes consciousness. A person acquires self-consciousness through the outer world, abstracting from one's organism through the shared parts of one's "inorganic body" (Marx) common to many people simultaneously. This is how we could eventually build a conscious machine – and it does not matter which senses that machine would have and which language it would use to communicate with us.

In the end, I would like to touch Malcolm's "programming" dilemma: is it more appropriate to split human (or robot's) actions in separate operations, or it is the sequence of goals that makes more sense for a conscious being? The problems like that arise from the implicit assumption that operations and their results are quite different from each other, and operational description treats the same thing in a manner different from that of structural approach. Once again, the number 3 can be defined either as an equivalence class on some set-theoretical universe, or as a finite sequence of operations of an abstract automaton (e.g. Turing machine). The universality of conscious behavior enables us to identify the entities obtained in different ways, and the very idea of consciousness implies that both operational and structural approaches will be used, their ratio being adjusted to the requirements of the situation. Recalling A. N. Leontiev's theory of activity, one could readily observe that the operational approach will be appropriate for describing how actions get composed of operations, while the decomposition of activities into individual actions would result in a sequence of goals. So far, no robot incorporates all the three levels (operation, action, activity), and hence the opposition of the aspects co-existing in every conscious act.

---

http://unism.pjwb.org/kjf/index.htm
http://unism.pjwb.net/kjf/index.htm
http://unism.narod.ru/kjf/index.htm
unism@ya.ru