

Jozef Kelemen and Jana Horáková

How Adam Goes to Post-Human

This essay deals with the development of machines and the influence it has had on the self-image of human beings. It reflects on machines' gradual improvement from the very times when people started constructing them (because we are inventive beings) to make physical work easier for themselves (because we are weak creatures), all the way through the period during which people gradually passed routine mental work onto machines because there was too much of it (because we are slow) or because it required great precision and attention (because we are inattentive). We end our thoughts with speculations about the possible future, about the times when we will coexist with machines. (Because we cannot deal with our future by ourselves?) We will start with a very old (and surprisingly productive!) human idea, and then, step-by-step, we will approach the present concept – to a certain extent unifying – of the man and the machine.

From Adam to Golem

The starting idea in this essay is the idea of a man created by man. This idea is presented in the first book of the *Old Testament* of the *Bible*: ... *the Lord God formed the man from the dust of the ground and breathed into his nostrils the breath of life, and the man becomes a living being* (*Genesis*, 2.7). Moreover, the *Bible* gave to the Western civilization also the ideological assurance that man is able to perform creative acts as well, by stating: *Then God said: Let us make man in our image...* (*Genesis*, 1.26).

Homer in the *Iliad* (book XVIII) expressed this – or a very similar – dream on artificially constructed human-like beings in the following verses: “*There were golden handmaids also who worked for him, and were like real young women, with sense and reason, voice also and strength, and all the learning of the immortals*”.

The idea of artificially constructed intelligent machines which communicate and cooperate with human beings is present in philosophical contemplations of Aristotle in his fundamental work, *Politics*: *For if every instrument could accomplish its own work, obeying or anticipating the will of others, like the statue of Daedalus, or the tripods of Hephaestus, which, says the poet, “of their own accord entered the assembly of the Gods”; if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workman would not want servants, nor master slaves* (*Politics*, Book 2, Chapter 4, 33-39).

In Prague a middle-age legend on this dream or belief has survived up to now. According to it, the famous rabbi from Prague, Judah Loew ben Bezalel (a real historical person active in the so-called Old-New Synagogue in Prague, buried in Prague Old Town Jewish cemetery), constructed a human-shaped creature sometime at the end of the 16th century – Prague's Golem. He created it in two main phases: First, he and his coworkers constructed an earthen sculpture of a man-like figure. Second, he found an appropriate text; he wrote it down on a slip of paper and put it into Golem's mouth. So long as this seal remained in its mouth the Golem could work, do the bidding of his master and perform all kinds of chores for him, helping him and the Jews of Prague in many ways, etc. The Golem was – at least to a certain extent and in a certain specific meaning of the word – alive.

The development of the old idea of the Golem has two basic sources: well-developed practical skills in pottery, perhaps the highest form of technology of the age in which the contents of the *Old Testament* originated, and the magical exegesis of the *Sefer Yezirah* (the *Book of Creation*) with the idea of the creative power of letters (symbols).

Fig. 1: The Old-New Synagogue in Prague and the usual look of Golems nowadays in Prague gift shops.

Our today's highest technology is based on computers and informatics. Observing the cognitive projects in the field of robotics at the end of the 20th and the beginning of the 21st century, we could realize some rather interesting and maybe surprising coincidences with the work of the rabbi Loew from five centuries before: The power of the appropriate interplay of the silicon-based computer hardware with the software, which includes the form of strings of well-combined symbols, animate our artificial agents in many surprising ways. Constructing hardware and writing software in order to create an autonomous system – especially a robot – are in a certain sense the same activities as those related to the construction of the Golem.

Metamorphose of Robots

Centuries after the “rabbi Loew and his Golem” story entered into the depository of stories about man creating man-like creatures, in 1915 an extremely sensitive writer from Prague, Franz Kafka, published in the journal *Die weißen Blätter* the short story *Die Verwandlung* (*The Metamorphosis*).

In short the story: When one morning Gregor Samsa wakes up as an insect, all his life is dramatically changed as the result of this metamorphose. The new body perceives its unchanged environment in a completely different way as the original human body. Moreover, Samsa is, because of the change of his body, perceived in completely different ways by his relatives, friends and colleagues. Samsa's unchanged mind, full of unchanged knowledge, reasoning and emotions, finds the new ways of existence and social functioning of the new body very hard and dramatical.

This typical Kafka's story can be understood in our context as referring to the high importance of the environment for the (human) body and its (human) mind, or as referring to the high importance of both as a kind of interface with the environment that shapes our experience of it and as a sensory source of information for the mind.

Kafka's artistic sensitivity and his reflection of problems inside the triangle of mind, body, and environment as described in *The Metamorphosis* is in present days converted into fundamental scientific problems concerning the construction of intelligent and embodied humanoid robots. However, there is also another perspective from which we can observe Samsa as embodied in a changed form situated in an unchanged environment. It is the perspective of Samsa's original environment (the environment formed besides others also by his family, his colleagues...). From this perspective we can realize the importance of the collective interactions for the formation of the individuality of the subject. We can realize that the individuality of a subject emerges, at least to a certain extent, from his or her interactions with other individualities. This experience is another important motivation for the present-day efforts in the field of creating and studying collective intelligence and societies of robots.

Five years after the time Kafka wrote the *Metamorphosis*, Karel Čapek, influenced by political, economical, and social conditions of the period (understood in a broad sense), and by eternal human dreams and ideas (which we pointed to in the above text), wrote his famous play *R. U. R.* (*Rossum's Universal Robots*). The text, in which the word *robot* appeared for the first time, has been fascinating and provoking our imagination till nowadays. As Karel mentioned in the Prague newspaper *Lidové noviny* (December 24, 1933), the name used by him for his “artificial workers” was at first *labor* – a neologism derived from Latin *labore* (to work). But he was not satisfied with this word – it sounded too academical to him – and therefore he asked his brother Josef – a recognized Czech painter, a representative of the Czech cubism, author of novels, poetry, and drama – for help. Josef suggested to him the word *robot*. The word is a neologism

etymologically rooted in the archaic Czech word *robota*. *Robota* means in Czech – as well as in the Slovak language of these times – the *obligatory work of servants*, the work of all slaves, servants, and workers in history.

The official opening night of the play *R. U. R.* took place in the Prague National Theatre on January 25, 1921. (But otherwise a company of amateur theatre group Klicpera from Hradec Králové, a town about 100 km eastward from Prague, had presented the play already on January 2, 1921 in the municipal theatre of Hradec Králové, which is even three weeks before its “official” first staging.) The first performance in the National Theatre in Prague was a great success. The play was soon performed in New York (1922), London (1923), Vienna (1923), Paris (1924), and Tokyo (1924) as well as in many other cities around the world.

The play brought great fame to its author in the USA and soon the story of the play, including its robots, became part of American pop-culture, especially of science fiction. In the context of American culture Čapek’s play lost its social satirical edges and the theme was re-interpreted as a metaphor of a hi-tech future that would destroy mankind because of human’s inability to prohibit its abuse. The motive of powerful machines jeopardizing humankind seems to have been already in Čapek’s days so strong that it entered his text even unknowingly, without the intention of the author; for more about this see (Horáková, Kelemen, 2004) and (Horáková, Kelemen, 2008) and also (Horáková, Kelemen, to appear).

The simplifying of Čapek’s metaphor to its dominant interpretations spread further from cultural contexts of the interpreters: The concept of robots in the American tradition (wide-spread all over the Euro-American cultural world!) was transformed from the original socio-critical view of robots as *simplified workers* into the form of robots as *complicated machines*, stuffed by electronics and processors and perhaps programmed in order to be able to replace the human workers in certain situations. With this feature the idea is a continuation of the dreams of ancient European efforts in mechanics like Vaucanson, Jaquet-Drozs, Kempelen, and many others.

Fig. 2: The nowadays view of the building of the Prague National Theatre – the place of the first staging of Karel Čapek’s *R. U. R.*, and the first picture of a robot – the costume design of robots by Josef Čapek for the first performance in Prague. (On the breast of the robot there is the date of the Prague opening night, the head is a caricature of Karel Čapek’s face.)

The Post-Modern Age

Discussing the human-machine ideas and metaphors as sketched above and trying to look into the near future, we are approaching the concept of the machine with which we will live in our (near) future. We have to make up a concept of the machine in relation to the *post-modern man*, a concept of the machine in the *post-modern age*. If the classical machines made physical work easier for us, then we can say that modern machines are devices that store, process, and provide information for us. Somebody had the idea of using another term instead of post-modern for the machines we construct due to our weakness, impatience or inattentiveness, but it is appropriate for the kind of machines that we construct to live with in a way more fitting to the conditions in the new environment and in the framework of new relationships (perhaps in some kind of partnership) formed by our cyber-culture, information age or postmodern era – as you like it.

But there are certain difficulties with the term *post-modern*. Jean-François Lyotard specifies post-modernism as a paradox of *futura exacta*, as what is future (post) and has at the same time

just become past (modo). This paradoxical explanation of a contemporary condition refers to the fact that there is no single artifact, artistic or scientific work, or anything else that we could label as post-modern and so separate it from modern, but there is an age, in which opposites can be found side by side and still be taken as normal and obvious. In this context the post-modern age can be characterized by an absence of criteria through which we could exclude opposites. We may consider this absence of criteria as a transitional phase in cultural development and deduce that in this sense the *post-modern situation* might be not a temporal period of time but a milestone from which we can count our *post-human* future in which we will have to leave our old criteria of understanding the world by means of dualist dichotomies in favor of “liquid complexity” of eternal metamorphoses of the world we live in.

The meaning of the term post-modern can be understood in connection with machines in the above-indicated sense. The post-modern machine will transcend preceding generations of strong and skilful machines. Today precisely the effort to understand this transcendence is bringing about a temporary coexistence of opposing opinions about the form and social impact machines will have in the future. They will undoubtedly still be machines, in relation to which the original perception of the user (subject) relationship towards the used (object) will change into a much more complicated relationship of more or less mediated mutual dependence. In any case, machines are already part of our everyday life and are gradually starting to make decisions for us. If this development continues in the direction and at the pace that we have witnessed and participated in over the last few decades, then it can be expected that we will live with future machines in a more or less *equal relationship*. Formulating stances on this situation is a source of discourse evoking the feeling of the post-modern era.

What is meant by *equality*? Above all that the machine of the future will no longer be treated as instrument. It will not be a cold and high-performance object with which we cannot be equal (in the sense of both human intellectual superiority, and limited abilities of human bodies). It will no longer symbolize man’s estrangement from machines in factories but cooperation with skilful machines everywhere.

The intellectual community, under pressure from this evidence, is gradually dividing into techno-optimists and techno-pessimists. Where do these hopes or concerns stem from? Is the machine a must of our further successful journey through history, or is it an embryo of a future apocalypse? This is a very sensitive topic. Let us cite Umberto Eco from the introduction of his book *Apocalittici e integrati: An apocalyptic ... basically consoles the reader, gives him the opportunity to catch sight, somewhere beyond the catastrophe, of a community of super-humans able to rise above the average banality by simply refusing the given state of affairs. It could be a small community, comprising of only two individuals, the one who is writing and the one who is reading, “ only us two, you and I, only we understood and saved ourselves: the others are the masses.*

Eco stresses that he used the term *superhuman* deliberately, to underline the pseudo-Nietzschean origin of many stances like this. He is probably right because man – even the most insignificant – wants to be an individuality at least in his micro-collective, for instance in the eyes of his family, although paradoxically, he very often has to join the majority to get the feeling that he has become an individuality. He becomes one in his own eyes by attaching himself to a group – usually the weaker the personality the larger the group. But because Man is in principle – whatever he thinks about himself – a member of society, this artificial individualization goes hand in hand with the creation and characteristics of the age in which Man lives.

The Post-Modern Machine

Now we return to a fairly urgent problem of philosophers such as Richard Rorty, Jacques Derrida or Jean-François Lyotard. These philosophers look for answers to the question of where today – in a world of cognitive pluralism, in a world where freedom of the individual really has a high value, in a world of cultural differences and tolerance towards others – where in this world can we plant a certain system of equality and where in this world can we find a

political and legal system, which we – despite these differences – could subordinate ourselves to. How can we reach an unforced agreement in these matters? This is a topic, which is being discussed quite helplessly in philosophy.

What is coming? A society in which there will be more space for individualization. Above all, thanks to new technology imprinting gradually its mark on culture. Let's realize that if somebody actively stands up against the culture industry today, he should see its roots not in the first television broadcast, but in the invention of the printing press and in the ideology of people's equality and sovereignty, writes Eco in the above quoted book, that has come out recently. When today's intellectual finds him or herself face-to-face with modern computer technology, he or she probably does not doubt the usefulness of the creation of conditions for general literacy or the creation and usefulness of a social structure, whose foundations could also be built thanks to the spread of literacy and therefore thanks to Guttenberg's – principally technical – invention, and which determined the nature of a whole exceptionally important era in the development of our civilization. It seems to us more likely that something similar to the feelings of the former British proletariat survives – the fear that he or she would no longer be indispensable. However, every intellectual must him- or herself overcome this barrier of the fear of estrangement.

The machine, as soon as it had been created and we had started to use it, started to integrate into our society. Reflexive interaction began within the framework of which the machine was an improvement, but at the same time, under the influence of the world it created, its creator has undergone changes, too. The conditions of his life have changed, his desires, hopes, and the aim of his effort to create new machines, he has refined his vigilance, etc. It is enough to realize what influence the automobile has had on our behavior, our urban concepts, transport possibilities, legislation, etc.

But to predict anything about the impacts of reflexive interaction is extraordinarily difficult. *If a reflexive interaction can change*, writes George Soros, influenced by Popper's concept of sciences and his own experience in analyzing developments on financial markets, *both participant's thinking and the actual state of affairs, timeless generalization cannot be tested. What happened once does not necessarily recur when you repeat the experiment and the whole beautiful structure collapses. No wonder! Underlying the model is the unspoken assumption of a deterministic universe. If phenomena did not obey timelessly valid universal laws, how could those laws be used to produce predictions and explanations?* (Soros, 1995, p. 216).

What we now witness is an embryo of the post-modern machine. But this is not a vision of the future! We must not only think of supercomputer super-intelligence, skilful robots or of a highly personalized computer of the future from the pages of science fiction. What we now have on our mind is for example the cardio-stimulator. Probably in the future it will be even improved. Maybe it will be linked to a healthcare center, where it will send information about the state of the body, under the skin of which it will have been implanted. It will receive messages from the center on what to do in a given situation. It will be something more than just an electrical stimulator of muscle activity. It will maintain the optimum balance of different hormones, minerals, vitamins, and who knows what else in the host organism and thereby protect it from stress, over-excitement, microorganisms and virus attacks, from the use of addictive chemicals and so on. Maybe we will not even carry cell phones in pockets and handbags because they will maybe also be implanted somewhere under the skin. And they will function completely differently than we use them today. Maybe a mere intention to announce something to somebody will be enough and he or she will know that we want to speak to him or her and if he or she has time and wants to, a connection will be made. There will be a town or a building that will follow your movements, receive your messages, and guide you through its insides to get you where you want to go. It will draw your attention to things that interest you, warn you of dangers... Maybe you will even be able to take a walk without leaving the comfort of your favorite chair, through strange visions and strange worlds of virtual realities. We are just setting out on a road that will soon historically bring us to the already mentioned pragmatic understanding of questions, which Jean Baudrillard submitted in a rhetorical and performative question: *Am I a man or a machine?*

The Cyborg

Jean Baudrillard (1989) asked whether he was *a man or a machine*? He claims that today we, who are searching for an answer, are obviously and subjectively people. But virtually, as he points out, we are approaching machines. It is an ambivalent statement of uncertainty, that joins us, living in modern times of worker's relationship to machines in industrial plants, with the post-modern approach to machine processing and mass disseminating of information. Technology, through which it produces, gradually eliminates basal dichotomies of the man/machine and object/subject type but also of freedom/restraint type and perhaps some others.

Their effects have started to change our environment and even we – being a part of this environment – have changed. We try to convince the reader that these new, autonomous machines interwoven with their surroundings will, using their bodies, gradually acquire something that can exclusively be thought of as their body, and that will very much resemble what we call – when we talk about ourselves – our individual experiences of “being” and is defined by cognition through an interaction with our environment. These beginnings of subjective experience of (postmodern) machines can form the basis of individual, subjective reactions, for which we presume the existence of something we call – again only when talking about ourselves – a subjective consciousness.

People often have the impression that the problem of the relationship between the mind and the body is only philosophical, fairly reliably remote from something that can actually affect us. As if thoughts about a subject in other than an anthropomorphic context were – as a result of their very nature – *ex definitio* pointless and thoughts about cyborgs belonged exclusively in science fiction or in the arsenal of themes for post-modern thoughts. But it is not the case. The mind and body – namely the mind and body of a machine – today stand in the center of the very current tangle of problems in the artificial intelligence theory, technology, and robotics, i.e. disciplines that on one hand evoke the greatest concern and on the other the greatest hope in connection with cyborgs. At the same time it is impossible to exclude certain philosophical implications of current researches. It is actually much more realistic to expect them.

If Čapek's *robot* can be seen as an artificial modern humanoid machine (the body of a worker or a soldier as an ideal prototype of members of a modern society) than the *cyborg* is a symbol of the post-modern human being (as a metaphor of our experiencing of the information society). As long as we are able to free ourselves from the binary mode of articulating reality, there is nothing to stop us from seeing reality as basically a “hybrid”. Then reality seen in terms of binary opposition (e.g. human vs. robot, man vs. machine), is more product of our thought than anything else. What we have in mind can be more closely explained using the metaphor of twilight; cf. (Kelemen, 1999). Twilight is not a hybrid of light and dark, but light and dark (human and machine) are opposite extremes of twilight. Similarly, the cyborg is perhaps not a hybrid of the organic and mechanical, but the “*organic*” and the “*mechanical*” are two extremes of the *cyborgic state*. This is the basic thesis of the ontology of twilight.

Towards the Post-Human

In the beginning of the fifties of the past century, an ambitious way of human professional curiosity was started by Alan M. Turing's famous article on intelligence and computing machinery, and then, in 1956 it was baptized as Artificial Intelligence (AI). “...*finding useful mechanistic interpretations of [...] mentalistic notions that have real value [...] is associated in its most elementary forms with what we call cybernetics, and in its advanced forms with what we call artificial intelligence*” writes Marvin Minsky (1968, p. 2), co-founder of the Artificial Intelligence Laboratory at MIT. But building an android, an autonomous robot with human-like form, abilities, and intelligence, has been both a recurring theme in science fiction stories and

movies, and an ambitious goal for the Artificial Intelligence community: “...here for the first time we see a sufficiently concrete (i.e., technical) foundation for the use of mentalistic language as a constructive and powerful tool for describing machines” (Minsky, 1968, p. 2). During decades of research, two main ways towards this goal have been discovered – the *top down* way trying to analyze the human mind from the position of a computational paradigm and then to (re)construct it step by step on the basis of its precise understanding. Minsky explains this style (through his Shakespeare-like verses of the Prologue of *The Society of Mind*) as follows: “What can we do when things are hard to describe? We start by sketching out the roughest shapes to serve as scaffolds for the rest; it does not matter very much if some of those forms turn out partially wrong. Next, draw details to give these skeletons more lifelike flesh. Last, in the final filling-in, discard whichever first ideas no longer fit” (Minsky, 1986, p. 17).

The alternative way consists in proceeding from *bottom up*, trying to synthesize step-by-step more and more clever machines behaving in intelligent manner in their specific environments. The basic idea behind this approach emphasizes the exploitation of decentralization of systems as intensively as possible. According to it the robot does not need any central representation of its outer environment. A control system is necessary only for managing the conflicts between the decentralized modules. It is supposed that the robot does not need any coherent concept of the outer world. Instead of it the robot must have good enough possibilities to learn what to do directly of its interactions with its environment. A famous project in this field, the project Cog at MIT Artificial Intelligence Laboratory, started with the declared goal to build “...an intelligent physical humanoid robot including active vision, sound input and output, dexterous manipulation, and beginnings of language, all controlled by a continuously operating large scale parallel MIMD computer” (Brooks, Stein, 1994). The result of this project is an upper-torso humanoid robot, which approximates human movements and sensing thanks to 21 degrees of freedom movement actuators and visual, tactile, auditory, and vestibular sensors (Brooks et al., 1999). In order to allow social interactions between robots and humans C. Breazeal developed a platform with capabilities to simulate human emotive facial expressions, the emotional head Kismet (Breazeal, 2000). But some other experiments with robots also evoke a deep impression of their activities at least in a certain meaning and to a certain extent similar to those of human beings; see Fig. 3, for instance.

Fig. 3: A “human-like” gesture of a simple robot Khepera (by the K-Team) during an experiment (at the Robotic Lab of the Institute of Computer Science, Silesian University at Opava, Czech Republic).

For many of those working in the field of artificial intelligence and advanced robotics, writes Hayles (1999, p. 246), “...becoming a posthuman means [...] envisioning humans as information-processing machines with fundamental similarities to other kinds of information-processing machines, especially intelligent computers. Because of how information has been defined, many people holding this view tend to put materiality on one side of a divide and information on the other side, making it possible to think of information as a kind of immaterial fluid that circulates effortlessly around the globe while still retaining the solidity of a reified concept. [...] Other voices insist that the body cannot be left behind, that the specificities of embodiment matter, that mind and body are finally a “unit” [...] rather than two separate entities. Increasingly, the question is not whether we will become posthuman for posthumanity is already here. Rather, the question is what kind of posthumans we will be.”

The Czech philosopher (and also poet and writer) Zbyněk Fišer (alias Egon Bondy), who contemplated the future of mankind sometime during the end of the sixties, sketches two possible answers to the previous question of Hayles: “The emancipation from the biological base, dispose of it or overcome it, surely does not mean, and cannot mean in any case, the achieving of any immaterial form of existence of intelligence. [...] Something goes on which we cannot characterize in some other way than as an artificial form of existence, artificial in the meaning that it is not biological but fabricated” (Bondy, 1993, pp. 52-53). As a story writer he sketched the following view: “The man-machine combination is sci-fi. [...] Man is a biological

unit – a digestive tube plus sexual organs. You can add whatever you want to it but it will remain a hybrid. The biological evolution added the brain. And immediately as the brain becomes productive enough, it begins to collide with the digestive tract and with sexual organs. The result is a jewel! The potential aggression on one side of the coin and the never-ending feeling of vain boredom on the other one. And if you add some machines to all that, the result will be much wrong. So we must develop a new kind of completely artificial beings. Beings which will survive because they will really re-produce and not because they will increase their gratitude to their digestive tracts and their sexual organs” (Bondy, 1997, p. 121).

The man-machine combination has been an attractive idea up to now for authors of science fiction and cyberpunk – let us mention e.g. *The Turing Option* by H. Harrison and M. Minsky published first in 1992. The idea of the robot initiated in the play *R. U. R.* recalls Bondy’s second alternative of human being of the future quoted above. In both cases their authors refer to biochemically based creatures. Already in the *R.U.R.* factory were the first robots made on assembly lines from parts/organs, they didn’t reproduce sexually.

The above-sketched twilight metaphor presupposes a continuum between the two extremes – the dark and the light. In other words, we suppose a monistic view of the physical and mental in our considerations. Whether the human mind and consciousness are based on a material substrate or the surrounding physical world is an illusion of our minds. Most of contemporary philosophers and scientists accept a physicalistic view of the mind; no success has been achieved up to now in answering the question of how consciousness might be physically explained. This is the reason why others offer alternative formulations of physicalism, or others forgo it in favor of more dualistic concepts for solving the traditional mind/body problem; for an overview of the approaches see e.g. (Van Gulick, 2001).

A position not refereed in Van Gulick’s overview just mentioned above is that of Susan Blackmore dealing with machine consciousness and rejecting its possibility because of a simple reason: the human consciousness is from her point of view not more than an illusion of human minds. Humans she consider as the only species “...that readily and easily imitates a wide variety of sounds and actions. This suggests that we alone are supposing a second evolutionary process, cultural or mimetic evolution. If it is so, human evolution must have taken a very different course from that of other species once we became capable to imitate” (Blackmore, 2003, p. 22). Humans are – thanks to their biological and cultural evolution – not general-purpose machines but rather highly selective imitating machines. The level of selectivity of present humanoid robots, according to Blackmore, does not suffice. Mentioning Cog and Kismet she writes: “*Cog barely imitates. Cog can point to objects, recognize joint attention and imitate simple behaviors such as head nods, but it was not designed as a mime machine. The position is additionally complicated by the fact that observers often attribute imitation and other abilities to robots, especially “sociable robots” such as Kismet, even when they do not have any (Breazeal, 2001). [...] Cog, Kismet and other robots will never acquire a human-like consciousness unless they are dramatically redesigned to be capable of true imitation learning”* (Blackmore, 2003, pp. 26-27).

There exists another possible (and plausible) view of the essence of mind and consciousness. It is expressed clearly by defenders of the bottom up approach to artificial intelligence and intelligent robots. R. Brooks writes: “*My feeling is that thought and consciousness are epiphenomena of the process of being in the world. As the complexity of the world increases, and the complexity of processing to deal with that world rises, we will see the same evidence of thought and consciousness in our systems as we see in people other than ourselves now. Thought and consciousness will not need to be programmed in. They will emerge.*” (Brooks, 1999, pp. 184-185). So, mind and consciousness are, according to this view, comprehended as emergent phenomena in the sense of J. H. Holland, who explains emergence as „... all a product of coupled, context-dependent interactions. Technically these interactions and the resulting system are nonlinear: The behavior of the overall system cannot be obtained by summing the behaviors of its constituent parts... However, we can reduce the behavior of the whole to the lawful behavior of its parts, if we take nonlinear interactions into account“ (Holland, 1998, pp. 121-122; emphasized by him). The explanation of mind and consciousness

as emergent phenomena is acceptable also in the case of the top down approach. M. Minsky in his social theory of mind tries to „... *explain how minds work. How can intelligence emerge from non-intelligence? To answer that, we'll show that you can build a mind from many little parts, each mindless by itself.*“ (Minsky, 1986, p. 17).

Conclusions

In the above text we tried to mention several important complement sources of post-humanism connected with the human-machine metaphor, with robots and cyborgs. We have seen at least two different developmental lines of thinking on robots, both rooted in the play *R. U. R.* The first one – intended by K. Čapek – starts from the deep conviction that mankind is able to survive its own (political, scientific, and technical) adventures thanks to its culture and its deep conviction in the moral potential of the human race. The second one starts from a message discovered in the play *R.U.R.* in U. S. A. – a more or less technological message on the dangers of technical creativity of mankind.

Then we sketched two directions of the development of the idea of human-like machines in AI and robotics as a scientific and engineering discipline and we mentioned several possibilities how to react to this ambitions from the position of philosophy. We have involved into this context, rather than commonly known philosophers, an internationally unknown one from the same intellectual climate of Prague, where Čapek developed his ideas and worked. However, in order to provide a picture as complete as possible on the topic, we add the position of S. Blackmore, too: “... *rather than thinking in terms of how “we” should design future machines, it may be better to think in terms of how the mimetic co-evolutionary process will design them. As machines copy more and more information from one to another, the machinery they use will co-evolve with that information, ultimately becoming self-replicating. We are not likely to have control over this process.*” (Blackmore, 2003, p. 29).

Concluding that all: What is clear today, what is out of any doubt at the beginning of the 21st century, is that for our destiny as mankind an unconditional assumption of our survival as species in our civilized world is to live together with the main product of our civilization – with machines. While in May 1961 M. Minsky wrote that “... *we are on the threshold of an era that will be strongly influenced and quite possibly dominated by intelligent problem-solving machines*” (Minsky, 1961, p. 406), in the present time machines are already part of our lives. We have started to construct them (because we are inventive beings) to make hard physical work for us (because we are weak). Then we gradually passed routine mental work onto them of which there was too much for us (because we are slow) or which required great precision and attention (because we are inattentive). Now machines are gradually starting to make decisions instead of us (because we are slaves of our own psycho-physical limitations) and they are starting to behave autonomously in the environment we mutually share. In any case, if the development continues in the direction and at the pace that we have witnessed and participated in over the last few decades, then it can be expected that we will live with future machines in a relationship of greater or lesser equality. Such relationship – reflected at least from the play *R. U. R.* on – may become the crucial ethical base for the coming age of post-humanism.

References

- Aristotle: *The Basic Works of Aristotle* (R. McKeon, Ed.). Random House, New York, 1941
- Baudrillard, J.: Videowelt und fraktalen Subjekt. In: *Philosophie der neuen Technologie*. Merve Verlag, Berlin, 1989, pp. 113-131
- Blackmore, S.: Consciousness in meme machines. In: *Machine Consciousness* (O. Holland, Ed.). Imprint Academic, Exeter, 2003, pp.19-30

- Bondy, E.: *Juliiny otázky*. Dharma Gaia, Prague, 1993 (Julia's Questions, in Czech)
- Bondy, E.: *Cybercomics*. Zvláštní vydání, Prague, 1997 (in Czech)
- Brooks, R. A.: *Cambrian Intelligence*. The MIT Press, Cambridge, Mass., 1999
- Brooks, R. A. et al.: The Cog project – building a humanoid robot. In: *Computation for Metaphors, Analogy, and Agents* (C. Nehaniv, Ed.) Springer, Berlin, 1999, pp. 52-87
- Brooks, R. A., Stein, L. A.: Building brains for bodies. *Autonomous Robots* 1 (1994) 7-25
- Breazeal, C. L.: *Designing Sociable Robots*. The MIT Press, Cambridge, Mass., 2001
- Čapek, K.: *R. U. R.* Simon and Schuster, New York, 1973
- Eco, U.: *Appocalittici e integrati*. Bompiani, Milano 1964
- Harrison, H., Minsky, M.: *The Turing Option*. Simon and Schuster, New York, 1992
- Hayles, N. K.: *How We Became Posthuman*. The University of Chicago Press, Chicago, 1999
- Holland, J. H.: *Emergence*. Addison-Wesley, Reading, Mass., 1998
- Homer: *The Iliad*. Penguin Books, New York, 1998
- Horáková, J., Kelemen, J.: From Rossum's Universal Robots to the post-human. In: *Cybernetics and Systems – Proc. 17th European Meeting on Cybernetics and Systems Research, vol. 2*. (R. Trappl, Ed.). Austrian Society for Cybernetic Studies, Vienna, 2004, pp. 774-779
- Horáková, J., Kelemen, J.: The robot story. In: *The Mechanical Mind in History* (C. Husbands, P. Holland, O. Wheeler, eds.). The MIT Press, Cambridge, Mass., 2008, pp. 283-306
- Horáková, J., Kelemen, J.: Artificial living beings and robots. *Journal of Artificial Life and Robotics* (accepted)
- Kafka, F.: *The Metamorphosis*. Vitalis, Prague, 1999
- Kelemen, J.: On the post-modern machine. In: *Scepticism and Hope* (M. Kollár, Ed.). Kalligram, Bratislava, 1999, pp. 221-243
- Liotard, J.-F.: *La postmoderne expliqué aux enfants*. Editions Galilée, Paris, 1986
- Minsky, M.: Steps toward artificial intelligence. In: *Computers and Thought* (E. A. Feigenbaum and J. Feldman, eds.). McGraw-Hill, New York, 1963, pp. 406-450
- Minsky, M.: Introduction. In: *Semantic Information Processing* (M. Minsky, Ed.). The MIT Press, Cambridge, Mass., 1968, pp. 1-32
- Minsky, M.: *The Society of Mind*. Simon and Schuster, New York, 1986
- Petiška, E.: *Golem*. Martin, Prague, 1991.
- Soros, G.: *Soros on Soros – Staying Ahead of the Curve*. John Wiley & Sons, New York, 1995
- Van Gulick, R.: Reduction, emergence and other recent options on the mind/body problem. In: *The Emergence of Consciousness* (A. Freeman, Ed.). Imprint Academic, Thorverton, 2001, pp. 1-34

Authors Addresses:

Prof. RNDr. Jozef Kelemen, DrSc.

Institute of Computer Science, Silesian University
Bezruc Sq. 13, 746 01 Opava, Czech Republic
kelemen@fpf.slu.cz

Mgr. Jana Horáková, PhD.

Institute of Musicology, Masaryk University
Gorkého 2, 602 00 Brno, Czech Republic
jana-horakova@volny.cz