Artificial intelligence in science fiction as a model of the posthuman situation of mankind *

MIROSLAV KOTÁSEK

ARTIFICIAL INTELLIGENCE AS A SCIENTIFIC FIELD AND FICTION

The article[1] maps a range of topics and contexts which, in my view, necessarily come into play when one of the most frequent science fiction topics – artificial intelligence – is to be thoroughly considered. The main paradox the reader faces (however indirectly or unconsciously) while reading a story making use of artificial intelligence is that he or she is confronted with a concept that views the features of human mind and intelligence as identifiable, fabricatable or programmable – the paradox stemming from the fact that it is exactly these features that enable him or her to read and interpret the given story. Implicitly, the suggested answers indicate the difference of such a self-referential structure from that of self-conscious fiction, meta-novels, and anti-illusive narrative in general.

In today's cultural situation it is almost impossible to have clearly defined borders between scientific "concepts and terms" and their "vulgar" usage in everyday discourse. Considering the proliferation of technology in everyday life, transforming the very cognitive capacities of the human mind (especially through changing the ways people experience the physical world), this situation can be seen as especially relevant if we want to discuss "intelligence", "mind", and "cognition" and its relevance as far as its role in constructing human identity is concerned. At the same time, language used in "everyday discourse" is not in any way naïve, innocent or uninformed, but forms a complex cultural construct that incorporates the beliefs and values of a given society (say: Western society of the last seventy-five years). Simultaneously, it also influences the use, functioning, and definition of these beliefs and values – in other words: cooperates in codifying them. My question is: what is the function of science fiction in shaping people's ideas about the plausibility of constructing an artificial intelligence and hence the knowability of the processes that define the human mind, further evolution of the human race, and so forth?

There were several questions I asked myself before writing this article. Firstly, what allows me to talk about "artificial intelligence" as a scientific field (in itself a rather unstable domain where psychology, computer sciences, neurology, genetics, study of perception, semiotics, information theory, and linguistics intersect and influence each other) and science fiction at the same time? Can anything relevant be concluded

for cognitive literary theory from this possibility and/or impossibility? Secondly, how do people process the different and seemingly incompatible discourses offered by the field of "artificial intelligence" (describing and modelling cognitive processes; testing and producing machines, robots, and software), on the one hand, and "popular literature" (whose main and only purpose is usually regarded – even by academics – mere entertainment for the public), on the other hand? By claiming these discourses do intersect at a certain point, most probably around the same topic (though arriving at it by different means) am I in fact suggesting that science fiction can still be viewed as *the* genre offering a colloquial notion of "science", thus fulfilling one of the main roles it acquired in the second half of the nineteenth century when it came into existence? Can it be then argued that we still feel a profound alliance between science fiction and scientific discourse, with SF explaining, domesticating the scientific discourse, the findings of natural sciences, and the resulting new technologies, at the same time instigating "cognitive estrangement" (cf. Suvin 1979) and "defamiliarisation" (Jameson 2005, 286) of what the inhabitants of a given era have stipulated to call "reality"?

If we answer in the affirmative, it will be only part of the whole story, as the language of science is thus delegated a primary position – subsequently, science fiction can only react passively and use such a language in a specific (i.e. literary) way. I want to argue that science fiction might compete with other discourses (political, pedagogical, artistic, social) in the attempt to orient people in the contemporary world, offering a model for understanding modern human life (and more to the point, human mind) in a more blatant and instructive form than any of the above mentioned discourses, not to mention other "literatures". In the same vein, Fredric Jameson, commenting on the work of William Gibson, claims that "the representational apparatus of Science Fiction […] sends back more reliable information about the contemporary world than an exhausted realism" (Jameson 2005, 384). What is even more relevant for my argument is the way he describes the situation of high modernist literature: "social 'contract' between writer and reader [which] has had as one significant structural consequence, the transformation of the cultural text into an *auto-referential* discourse, whose content is a perpetual interrogation of its own conditions of possibility" (292; italics author), specifying the situation of science fiction oriented towards the future: "the centre of gravity of such narratives shifts towards an auto-referentiality of a specific, but far more concrete type: such texts then explicitly or implicitly, and as it were against their own will, find their deepest 'subjects' in the possibility of their own production, in the interrogation of the dilemmas involved in their own emergence as utopian texts" (293). Enhancing Jameson's ideas further, it might be claimed that in science fiction texts employing the topic of artificial intelligence the reader is confronted not only with a model of the possibility of producing such texts (the sociocultural contract), but also with a model of its reception (the narrative contract).

Having mentioned reception it should be obvious that the reader here is no sovereign subject capable of negotiating all the different meanings offered to him or her by the text, but rather a subject always already situated, positioned, fragmented, defamiliarized, and estranged. In a way he or she plays the role of an observer, central to the thinking of Maturana and Varela, who stress that

all cognitive experience involves the knower in a personal way, rooted in his biological structure. There, his experience of certainty is an individual phenomenon blind to the cognitive acts of others, in solitude which, as we shall see, is transcended only in worlds created with those others (1992, 18).

Consequently, reading can be viewed as a model of our everyday conduct in the given world: "we are experiencing a world. But when we examine more closely how we get to know this world, we invariably find that we cannot separate our history of actions – biological and social – from how this world appears to us. It is so obvious and close that it is very hard to see" (23). The act of reading can provide the experience of this getting into contact with the world most effectively, especially because of the self-referential quality that can be ascribed to it. I have so far mentioned self-referentiality on the socio-cultural and intentional (reading) levels, the latter having, according to Maturana and Varela, more or less a direct link to self-referentiality on the biological level.

Some would argue that this link to the biological level – though not exactly to the "lowest level" where only, as Hofstadter claims, formal descriptive rules apply (Hofstadter 2000, 686) – can be quite explicit in reading. Such a concept of reading sees the body involved in the hermeneutic process – thus cognition cannot be located inside the "brain". Christopher Keep is very open in this respect, commenting on and getting inspiration from Roger Chartier, Elizabeth Grosz and Jacques Lacan:

> To read is not to withdraw from the body so much as to occupy it differently. Conversely, the body, too, is configured and altered by our perceptions of it, perceptions which are themselves constituted in the ensemble of gestures and actions which we perform. […] Reading is one such social and political inscription, a writing of the body which reveals its culturally and historically contingent character. Bodies, or more precisely, our different senses of inhabiting a physical self, are as variable as the texts that they read (1999, 165).

Although Keep associates this reading exclusively with "hypertext" as opposed to "Codex book" I think his ideas can be applied to reading in general, especially of those science fiction texts that propose (on the level of plot) decentred views of individual identity and the identity of the human species.

However, the relationship between the scientific discourse of the field of artificial intelligence and the fictional discourse of science fiction is even more complicated. Since the 19th century the technology produced has changed profoundly – the attempts at constructing artificial intelligence are an important part of this sea-change. Wired communication, trains, planes, steam and internal combustion engines, the phonograph, and photography all seem laughable in comparison with technologies that change the very way people experience the world and live their lives. In his cyberpunk anthology *Mirrorshades,* Bruce Sterling pointed out that it was precisely this "technological shift" that initiated the cyberpunk movement in the first half of the 1980s:

> for the cyberpunks […] technology is visceral. […] [I]t is pervasive, utterly intimate. Not outside us, but next to us. Under our skin; often, inside our minds. Technology itself has changed. […] Eighties tech sticks to the skin, responds to the touch: the personal computer, the Sony Walkman, the portable telephone, the soft contact lens (1991, 346).

MIROSLAV KOTÁSEK

Thirteen years later, Roy Ascott describes the peculiarity of the situation and its difference from nineteenth-century technologies in this way:

> the paradox is that, after the gross materialism of the old industrial culture, relentlessly fueled by technology and science, the issues of consciousness, mind and spirit which are rising to the top of the artist's agenda are themselves instituted by technology and science. It is technology which is both informing and challenging our models of mind. It is hard science which is confronting the soft questions of philosophy (1999, 176).

Usually, the discourses of the natural sciences and fiction are regarded "incompatible" (hence requiring "adaptation" through science fiction, textbooks, and popular-scientific texts). The situation is much more general than we would probably suspect. Marie-Laure Ryan in her essay "Cyberspace, Virtuality, and the Text" instructively recapitulates the history of the term "cyberspace", comparing its usage in literature, other media, and in scientific discourse, commenting on the first Conference on Cyberspace and the resulting book of contributions *Cyberspace: First Steps*. In her analysis of the language used in (what is to be taken for) scientific discourse she finds strong parallels between the two linguistic regimes of fiction and science. She concludes: "for most contributors, however, cyberspace […] is foremost a catalyzer of dreams. True to Gibson's definition, the cyberspace of *Cyberspace: First Steps* is a collective hallucination of unlimited dimensions, since the number is determined by the imagination" (Ryan 1999, 83). What has to be stressed is that Ryan is in no way illustrating the general postmodern thesis of reality as language or discourse, which would be in this case even further highlighted by the fact that language is forced to deal with a phenomenon whose existence is "virtual", has no clearly defined borders and the sovereignty executed in imagining and constructing such a domain is rightfully equal on both the scientific and fictional side. I do think Ryan is pointing to something more substantial: the scientific discourse of the natural sciences, as a discourse supposed to deliver unbiased, rational descriptions and data that can be mathematically verified, obviously borrows the language, hopes and fears of an unknown territory from the realm of fiction – where that phenomenon actually "came into being". The interconnectedness between science fiction literature and the field of artificial intelligence appears to offer some structural similarities. And it does not stop at the relationship between fictional and scientific discourse. It can be rightfully argued that the usage of a term (the vocabulary) influences the way an individual can think. At the same time, the usage of a shared discourse or vocabulary by a community (and I am not referring here to readers of science fiction) forms grounds of its "culture", lays out the foundation of sociability, and provides evidence thereof. Consequently, Haraway's remark: "the boundary between science fiction and social reality is an optical illusion" (1991, 149) can be taken literally.

In science fiction tendencies can be observed similar to those described by Ryan in connection with the term "cyberspace". It is sometimes very difficult to decide whether the ideas presented in a story originate in the science fiction "repertoire" or represent scientific speculations. As an example of the blurring of the division between the two discourses Hans Moravec and his book *Mind Children* (1988) can be mentioned. Katherine Hayles was deeply disturbed by his ideas proposing the

possibility of constructing and copying the human mind and placing it inside an artificial body or machine. She suggested that she began writing her book *How We Became Posthuman* as a direct reaction to the "nightmare" she encountered in Moravec's book:

> I was reading […] *Mind Children: The Future of Robot and Human Intelligence*, […] when I happened upon the passage where he argues it will soon be possible to download human consciousness into a computer. To illustrate, he invents a fantasy scenario in which a robot surgeon purees the human brain in a kind of cranial liposuction, reading the information in each molecular layer as it is stripped away and transferring the information into a computer. At the end of the operation, the cranial cavity is empty, and the patient, now inhabiting the metallic body of the computer, wakens to find his consciousness exactly the same as it was before (Hayles 1999, 1).

The uneasiness with which Hayles orients herself in these and other visions of the future of mankind appearing in the fifty years of the history of artificial intelligence is obvious throughout the whole book – it is in fact its main point. Similar attitudes and theories about the future of mankind have been offered by Ray Kurzweil. He envisions the possibility to "free" the human mind and store it in an information network:

> The longevity of one's mind file will not be dependent, therefore, on the continued viability of any particular hardware medium. Ultimately software-based humans, albeit vastly extended beyond the severe limitations of humans as we know them today, will live out on the web, projecting bodies whenever they need or want them, including virtual bodies in diverse realms of virtual reality, holographically projected bodies, physical bodies comprised of nanobot swarms, and other forms of nanotechnology. A software-based human will be free, therefore, from the constraints of any particular thinking medium (Kurzweil).

The ideas and images offered by Kurzweil are not dissimilar to those we find in Greg Egan's novel *Permutation City*, where they are not presented as straightforward, seemingly rational "scientific facts", but as extrapolations presenting grave ethic and existential dilemmas.

There are other consequences resulting from the blending of a matter-of-fact tone and science fiction imagery by Kurzweil, Moravec, and others. First, their theories indicate the construction of artificial intelligence is in itself not a problem at all. Second, this matter-of-factness projected into the future implies that not only is there no problem with artificial intelligence, but the same applies to futurity – the future can be reasonably predicted, the exponential growth of technological progress, once it has been stated, cannot be stopped. It can, of course, be questioned whether these future visions really are factual. Through presenting the future in seemingly realistic descriptions almost all differential characteristics of fiction and scientific discourse become ineffective.

Although I have been paying attention almost exclusively to science fiction, it has to be emphasized that the way science fiction contributes to the construction of the social world is, of course, not exclusively limited to a science fiction audience. The social world is constructed by borrowing and utilizing components identifiable especially on the level of plot and story, used in science fiction to model and suggest new types of sociability up until now not fully realized. What I am especially interested

in are the possible new social systems brought about by changes in the definition of humanity. The concept of "folk mythology" as defined by Jerome Bruner could prove as relevant in that it problematizes the locus of meaning, which is no longer identifiable with a Cartesian subject: "culture is also constitutive of mind. By virtue of this actualization in culture, meaning achieves a form that is public and communal rather than private and autistic" (1990, 33). Taking this idea seriously would mean that the construction of artificial intelligence inevitably involves construction of a corresponding culture. Hayles comments on the interconnectedness of different discourses, culture, and science:

> The larger global pattern that emerges when literary culture and artificial life are integrated together suggests that technology and culture are bound together in complex feedback loops that themselves have self-organizing properties. […] Science and literature can no longer proceed as separate discourses. They speak, if not in the same voice, from the same sites (1999a, 219–220).

Last but not least, our understanding of "intelligence" is also co-defined or at least influenced by the "vulgar" uses of a great variety of adjectives semantically connected with "intelligence" as employed by PR, advertisement, journalistic clichés, and mass media in general. In designations like "smart city", "smart home", "smart house", "smartphone"; "intelligent building", "intelligent kitchen", "intelligent control system", and many others, intelligence is equated with mechanical memory operations, distant checking of temperature, security, heating, ventilation, solar panel electricity usage, operations of household machinery, and the like. The resulting semantics can be understood as a devaluation of both the terms "thinking" and "intelligence" – these are then out of definition relatively easily programmable and operatable, machine-like; their actions affect almost exclusively automated behaviour. And because all of these usages seem to be evidently far-fetched from the "real" meaning of "intelligence", they are enormously effective as people are not taking the semantics literally (the designations would then appear as denigrations) but rather are influenced by them subconsciously. The resulting semantic map is further enriched by other nuances originating from different usage of words belonging to the same semantic cluster such as "philosophy" (philosophy of petrol stations, philosophy of retailing), "logic" (logic of buying), and probably also "literacy" (whether financial or digital). If we take into account the complete spectrum of values ascribed by different societies and social classes to "intelligentsia" and "intellectuals", we get the whole picture. Despite the tendencies to downplay the effects these usages can have on our understanding of "cognitive operations" and on the construction of the "social world", I would stress that they probably have greater impact when they leave the sphere of semantics alone and enter the sphere of (social, cultural, and political) action – people start to act as if the adjectives were describing "real" properties. And that is being done more often than not, whether by scientists, politicians or artists. Thus taking the aspects outlined seriously does not imply a "naive" viewpoint, but can serve as an indication of the processes by which a given culture constructs itself – the idea of human identity definitely lies at the core of all socio-cultural structures. Speaking of "artificial consciousness", Ascott describes part of it as having to do, among other aspects,

with "the increasingly cognitive quality of the tools, products and buildings that are now invested with intelligence: we are surrounded by things that think, or that at least are smarter than they used to be (and learning fast!)" (1999, 177). The semantic processes described accompany the mutually dependent, undistinguishable discourses of fiction and science: "when someone maintains that machines can think, that computers have minds, or that digital simulations are alive [...] the terms themselves undergo transformation and modification as new players vault onto the field" (Hayles 2005, 215–216). The two processes can be said to be analogous.

"THE SITUATION"

The situation briefly outlined above is one of the symptoms of what we are accustomed to call "the posthuman" or "posthumanism". I find the term especially relevant as the definitions of "intelligence", "mind", and "human identity" shift and have undergone substantial revisions. By no means do I see the role of science fiction in the destabilization of the concepts as central, yet I do consider its role in the whole process indispensable. It might seem ironic that in this regard there opens up the possibility to read science fiction (usually considered a typical example of "the fantastic") in "realistic" mode – either bringing forward or renewing the historical links with realistic narrative. It is exactly in this way that we could read Larry McCaffery's statement from the *Columbia Literary History of the United States* in which he states that science fiction is "arguably the most significant body of work in contemporary fiction" (1988, 1167) as well as Jameson's comment on the "representational apparatus of Science Fiction" quoted above. Thus I do not understand McCaffery's claim as referring to science fiction's ascent towards mainstream "aesthetic values". I would say that one of the main preconditions that allows me to view the position of science fiction in the way suggested is the above described "mésalliance" (as it is often viewed) of the discourses of science and fiction, as well as the postmodern idea of the world being constructed primarily by "language games", which enables language to play an essential role in the construction of socio-cultural reality, underlining social reality as an autopoietic system as well. Social systems can be then seen as autonomous and self-sustaining.

Mentioning above the construction of the "social world" as well as the self-referential bind between the producers and receivers of texts suggested by Jameson, we can legitimately ask what kind of world it is that Moravec, Kurzweil and others are counting on; which world it is that guarantees them that they will be heard. In other words: what conditions their discourse? We will probably find that there is no easy explanation. Of course, from what has been said it can be concluded that these texts are changing the cultural and social world in which they originate and put forward their ideas, acting as a kind of virus. The circulation of the texts creates favourable conditions for the enhanced acceptability of similar texts. The methodological hindrance we are facing when viewing social systems as autopoietic (self-referential) is how change can be introduced into a closed system (see Luhmann 1990, 15). But it is not our main concern. Hayles asks the same question, commenting very openly on Moravec:

I consider this a pipe dream. Human mind has evolved over millennia in such complex symbios with specific biological structures that it is inconceivable consciousness could exist in anything like its present form apart from the physical processes that constitute it. The interesting question, then, is not whether such a transformation can take place, but why it is a compelling imaginary at this cultural moment. What does it mean to lose your body, or more precisely to fantasize that you can lose your body and inhabit a computer? What cultural formations speak to this question, and through what dynamics are their responses generated and structured? (1999a, 208)

From the variety of responses these texts have provoked it can be said that they are being read in two opposing ways – either as scientific predictions about the future of human race (or at least the future of technology), or as deliric hallucinations ("pipe dreams").

By now it should be evident that this possible "double reading" is not so much a result of a particular text's characteristics but rather is coded into the culture producing such texts. Taking science fiction into consideration and concentrating on the role it has played in the last thirty-five years (in other words: focusing on the development of science fiction from cyberpunk onwards), I am becoming more and more sceptical about the possibility to subjugate science fiction to a "formal" reading of any kind. Hence I do not view narratives as transgressing the world in which they are created, giving birth to derivative "possible worlds"; the "worlds" narratives offer, the forms they accommodate and develop, and the contents they thrive upon are coextensive with the cultural formations and meanings they use and within which they operate, as well as the mental (semantic) contents assigned to them in individual processes of reading. No mimesis, then, no modality. The current cultural situation of the West can be seen at the same time as the product and producer of this jointure of model and original, simulacrum and reality, pointing at the breakdown of classical dichotomies underlying the concept of representation. A parallel observation can be made regarding the changing nature of humanity – the dynamics offered by current society is taken over by literature allowing us to see clearly the connections and mutual dependences of man and machine:

the continuing development of intelligent machines and the shifting meanings of key terms – work together to create a complex field of interactions in which humans and intelligent machines mutually constitute each other. Neither kind of entity is static and fixed; both change through time, evolution, technology, and culture (Hayles 2005, 216).

Stating no or minimal difference between fictional and factual texts, the understanding of science fiction as "realistic" discourse might seem provocative, probably even naïve, at the same time pointing to the obsoleteness of the terms.

I have implicitly suggested that cyberpunk initiated a change in the traditional science fiction treatment of exactly those oppositions that both postmodernism and posthumanism are so fond of dismantling. Veronica Hollinger sees the difference between pre-cyberpunk and cyberpunk science fiction as a radical one:

While SF frequently problematises the oppositions between the natural and the artificial, the human and the machine, it generally sustains them in such a way that the human

remains securely ensconced in its privileged place at the centre of things. Cyberpunk, however, is about the breakdown of these oppositions (1991, 204–205).

In other words, there is no sovereign, enclosed position from which the human subject could hope to observe and describe the ongoing changes. For cyberpunk authors the classical aesthetics of science fiction was not interesting in any way, as William Gibson points out in an interview conducted by Larry McCaffery: "My SF is realistic in that I write about what I see around me. That's why SF's role isn't central to my work. My fiction amplifies and distorts my impressions of the world, however strange that world may be" (1991, 276). Timo Siivonen, commenting on Gibson's novels, voices a similar idea: "in the world represented by Gibson, human experience is defined in terms of humanity's relationship to technology. This, however, is not an instrumental relationship, in which humans, according to their own intentions, exploit technology to subjugate the outside world. Modern technology is no longer an entity discrete from the user, but rather an environment in mutual interaction with human beings" (1996, 228).

The first volume of Bernard Stiegler's *Technics and Time* can be regarded as the epitome of this viewpoint in philosophy. In it Stiegler offers his interpretation of Plato's dialogue *Protagoras*, more specifically, the passage in which Protagoras narrates the myth of Prometheus and his brother Epimetheus. According to Protagoras, Epimetheus wanted to help his brother Prometheus with the task of distributing qualities among all the creatures living upon the Earth. While fulfilling his voluntary task, Epimetheus forgets about the humans. When Prometheus comes to check how Epimetheus is doing, he discovers that Man is standing there naked, hungry and forlorn. To compensate for Epimetheus's forgetfulness Prometheus steals fire from Hephaestus and practical wisdom from Athena, giving them to Man. From that time on, Stiegler argues, Man can be defined as a "technological animal". Man is thus defined by a profound "lack" that has to be "corrected" by technology – by definition Man is a "prosthetic" or "technological" animal (Stiegler 1998, 185–203). Using the terminology of Donna Haraway, people have always been (literal or metaphorical) "cyborgs", although she stresses the situation as a current one: "By the late twentieth century, our time, a mythic time, we are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs. The cyborg is our *ontology*; it gives us our politics" (1991, 150; italics added by M. K.). The concept consequently breaks down the oppositions between "mind" and "body", "nature" and "culture", "subject" and "object".

In the last section I would like to indicate how the "hybridity" of opposing conceptual frameworks changes our understanding of the role literature plays not only in the socio-cultural system (I have tried to outline this function in the first part of my article), but also in our understanding of the reading process.

HYBRID DISCOURSE, HYBRID MINDS, HYBRID READING

I used a computer to write this article. That unquestionably qualifies me to be viewed as a "cyborg". Even worse: to fully grasp the consequences of what I am doing I have to view myself as one. I can still vividly remember what it was like to write semi-

　　　　　　　　　　　　　　　　　　　　　　　　　　　MIROSLAV KOTÁSEK

nar papers on a typewriter, with little or no possibility to edit the text. I can empirically state that in this particular case computer technology is definitely influencing not only the way the text is being created but also the way one can compose arguments, i.e. the way of thinking. Ascott claims:

> My computer is an intrinsic part of me at two levels. At the general level, it is integral to all those technological tools and instruments of seeing, prostheses of the senses, that augment our view of the world. On the particular level, it is the interface to my communication with others, apart from those with whom I live or work in close physical proximity. It is where most of my memory is located. It prompts essential actions. It enables me to plan and configure. It checks for accuracy, it assists my imagination, it gives me access to databanks in every realm of thought in every part of the world. Both scientists and artists are interested in what it is to be like this extended organism (1999, 178).

The fact that this relationship is in most cases "unreflected" and "natural" points best to its importance, indicating that the hybridity of man and machine has become the precondition of "standard" functioning of people (as in, for example, car driving).

Is it possible to incorporate all these aspects into a "theory of reading"? Although I am not sure that I am prepared to assert all the qualities Timo Siivonen ascribes to "cyborg discourse": "the manner […] in which various technological, natural, biological, social, linguistic and cultural changes are inscribed into the text's rhetorical structure" (1996, 229), I am tentatively asking what the features of "cyborg reading" would be. Siivonen builds his analysis of Gibson's novels around the term "oxymoron". I have already pointed to the mutual hybridity of the texts of science fiction and science and pointed to some possible consequences regarding the traditional dichotomies of real vs. fiction/simulacrum/virtuality. In trying to identify the possible way hybridity can be applied to reading, my starting point differs from both Haraway and Siivonen.

I have chosen Søren Brier as my starting point, who has devised an intricate theory of human cognition, partly built on the theory of the autopoiesis of living systems (as defined by Maturana and Varela), adapted further by Niklas Luhmann for use in the analysis of social systems and production of meaning. Brier distinguishes three stages (systems) that I find relevant for the reading process as well. The three separate systems are the socio-cultural sphere of meaning, whose intricate nature and functioning regarding the circulation of different texts has been outlined in the first part of my article. The second one would be the classical sphere of "intentional ego", usually regarded as the only sphere that can be rightfully called "mind":

> One way to understand our inner mental world is to see it as a way of representing our bodily interactions with the environment through the constructions of a felt signification sphere. In this way, an individual 'point of view' as a center of cognition, interest, and interpretation is created (Brier 2002, 115).

The second system is closely related to the body and the processes through which it influences cognition and houses "mind". Body represents the third (biological, physiological) autopoietic self-sustaining living system.

What is especially important is that Brier answers the methodological question about the possibility of introducing change into a closed system. He underlines the

different types of interpenetration, which I will demonstrate on reading, that allow the three different systems ("closed boxes") to "communicate": "The three closed systems produce different kinds of semiosis and signification through different types of interpenetration, plus a level of structural couplings and cybernetic 'languaging', as Maturana and Varela [in their book *Autopoiesis and Cognition*] call it" (Brier 2002, 121). As regards my approach to the reading process, the way "meaning" arises in these three systems is certainly crucial:

> Meaning is then seen as generated by the interpenetration of the systems. For example, language is a part of the socio-communicative system, but it does not really get a meaning before it interpenetrates with the psychic system and gets to indicate differences of emotions, volitions and perceptions 'putting words' on our silent inner being. But our cognitive, emotional and volitional qualities would only have a weak connection to reality if they were not connected to the survival of the living systems' organisation as a body in its interacting with the environment's differences in the development of a signification sphere in the evolution of the species (116).

I have tried to show beforehand how science fiction can be understood as a part of this definition of human mind (human semiosis). When Hayles points out the loss of "materiality" in the "posthuman situation", she has literature (among other phenomena) in mind, suggesting that literary texts are also "losing [their] bodies", and proposes to redescribe literature and literary works as a "space of encoding and decoding", which "links it with an operator outside the text who performs the coding operations" (1999a, 208). Hayles is specifically referring to the changes hypertexts initiate, claiming, in my eyes rather incoherently with her argument just quoted that

> narrative sequencing […] is one way in which narrative changes with literary hypertexts. The transformation of narrative extends beyond this concern, however, for increasingly it is bound up, in both print and electronic hypertexts, with a desire to explore the specificities of the medium, as the materiality of the book or electronic text enters into the space of representation in powerful new ways. This emphasis on materiality reverses several centuries of legal and literary constructions which asserted that the literary 'work' consists of language and concepts alone, not the book as physical object (2001, 22).

Let me mention a few books that stress, implicitly or explicitly, their own materiality and process of their creation (usually aided by computer software). Katherine Hayles's *Writing Machines* (Cambridge: MIT Press 2002) and Avital Ronell's *The Telephone Book: Technology, Schizophrenia, Electric Speech* (Lincoln: University of Nebraska Press 1989) are illustrative examples of books in which the typographic layout is stressed and subsequently challenge reader's abilities to orient himself or herself in the written text, forcing him or her to take into account the materiality of the book, its design and composition of information, words, and sentences on the page and in the whole "open work" (as Hayles decided to call this type of texts borrowing Eco's term). Hofstadter sends his "Words of Thanks" in GEB, among others, to "the text-editing program to which this book owes its existence […] I was able to do something which very few authors have ever done: typeset my own book" (Hofstadter 2000, xliv), and in the dedication part of his *Le Ton beau de Marot*, he mentions: "in writing it, I have depended from the start to finish on

the power, flexibility, and intelligent design of FullWrite 2" (Hofstadter 1997, xxii).

Beforehand I have quoted Christopher Keep who describes the role of body in reading, underscoring the materiality not of the text, but of the reader's body that becomes the object on which meanings are inscribed, pointing at the way these meanings are historically, culturally, and politically conditioned. I want to combine this idea of reading that is always already conditioned from the outside with the conceptual framework of Brier. I have pointed to the diverse ways in which science fiction (literary texts in general) cooperate in designing and transforming socio-cultural systems. Talking about materiality of texts and reading, I suggest viewing books as machines which are storing information, in very much the same manner as computers partake in the process of their creation. Then not only in writing a text on a computer but also in reading it could we talk about a machine-human hybridity (cyborgization). Take away the human and the book becomes a numb, material object stored in special buildings, mechanically and digitally reproducible. The importance given to embodiment (both thematically stressed in cyberpunk and post-cyberpunk science fiction and by cognitive scientists like Brier) can be analogically transferred to the relationship between the book and its (human) reader. Meaning is also embodied in the material form of a book, while the transformation of the body form – e.g. replacing print by digital text and hypertexts – leads to a transformation in meaning (content, if you will).

Hence the dichotomy of inside and outside would be another one that is being questioned by this understanding of the reading experience. When I open a book printed on paper, it offers me its material inside (pages covered by print), but only when I begin to read it do I come into contact with the meaning. Yet the book does not code meaning in sheer materiality – the reader and the book are using language which can never be privatized (despite empirical evidence it cannot be enclosed inside a book). I, as a reader, cannot escape the radical situatedness of reading – by definition, the fact I am reading "foreign" stories never lets me forget that the language which is coming at me I will never be able to make my own – language by the sheer fact of being used inside a book points at itself as the mediator of an interpersonal sphere of communication. In reading I can never escape the cultural and social dimension of understanding. The book does nothing more and nothing less than that it models this function of language. Although I can sometimes regard reading an intimate, sovereign act of will, such an understanding of language protects me from the dangers of solipsism. The book models exactly those interpenetrations used by Brier in his description of the theory of links existing between different levels of human cognition. Obviously, the book achieves to function as such a model only through and after making the reader part of the modelling process. From the point of view of the reader it equals making the book and its meanings part of his reading experience by turning the material signs into cognitive content in his or her mind, giving them back the internality stolen from them after the reader forcefully opened the book and began to read/stare at the print.

But it does not stop here. In such an outline of reading it is always necessary to ask similar questions Hayles did when wondering about the world cultural situation that

allowed the texts of Moravec and Kurzweil to function and be accepted. I have pointed to the role science fiction (fiction in general) plays in constructing such cultural and social systems also to implicate the genre as a cultural, social, and political institution. There is also a temporal aspect involved in the reading experience (apart from the temporal structures provided by plot and the fact that "it takes some time to read a book") – during reading the reader is confronted with texts that define (have defined, had defined) the cultural situation he or she is currently inhabiting. Taking all these aspects into consideration, the theory of autopoietic systems can be applied on all levels of human life and sociability from the very basic level of living organism as self-sustaining system, through the ways it forms "minds" (which then project themselves into cultural forms), as well as on the procedures used by individual "embodied minds" to interact in society, which is articulated precisely through this interaction.

How does literature use the resulting cultural meanings; how does she respond to them? Regarding artificial intelligence (which should have formed the red line around which my article centred) and its usage in science fiction (despite the claim that science fiction can and probably should be read as "realism"), the topics that define posthumanism and can be found in science fiction lead to the recognition of the situation of the individual reader as problematic. As one is confronted with a situation in which human identity is questioned, never secure, escaping the control of "mind", eluding individual projects, giving birth to assemblages of man and machine – whether it is a computer as I write this article, or a book as a machine storing meaning – the resulting hybridity cannot be understood as an external expression of a literary genre providing the reader with "fantastic" topics, but as a model of the cultural situation in which the literary text is created. I have also suggested that it can also influence and question the nature of the reading process, claiming this questioning might be inevitable. Literature is an autopoietic system connected to other autopoietic systems: the social system, cognitive embodied system, and the systems of living organisms securing their material survival.

LITERATURE

Ascott, Roy. 1999. "Art, Consciousness and Artificial Life." *Artificial Life and Robotics,* 3 (3): 176–180.

Brier, Søren. 2002. "Intrasemiotics and Cybersemiotics." *Sign Systems Studies,* 30 (1): 113–128.

Bruner, Jerome. 1990. *Acts of Meaning.* Cambridge (Mass.) and London: Harvard University Press.

Gibson, William. 1991. "An Interview." In *Storming the Reality Studio*, ed. by Larry McCaffery, 263–285. Durham: Duke University Press.

Haraway, Donna J. 1991. *Simians, Cyborgs, and* Women. New York: Routledge.

Hayles, Katherine, N. 1999. *How We Became Posthuman.* Chicago: Chicago University Press.

Hayles, Katherine, N. 1999a. "Artificial Life and Literary Culture." In *Cyberspace Textuality*, ed. by Marie-Laure Ryan, 205–223. Bloomington: Indiana University Press.

Hayles, Katherine, N. 2001. "The Transformation of Narrative and the Materiality of Hypertext." *Narrative* 9 (1): 21–39.

Hayles, Katherine, N. 2005. *My Mother Was a Computer: Digital Subject and Literary Texts.* Chicago: Chicago University Press.

Hofstadter, Douglas R. 1997. *Le Ton beau de Marot: In Praise of the Music of Language.* New York: Basic Books.

MIROSLAV KOTÁSEK

Hofstadter, Douglas R. 2000. *Gödel, Escher, Bach: An Eternal Golden Braid*. London: Penguin.

Hollinger, Veronica. 1991. "Cybernetic Deconstructions: Cyberpunk and Postmodernism." In *Storming the Reality Studio*, ed. by Larry McCaffery, 203–218. Durham: Duke University Press.

Jameson, Fredric. 2005. *Archaeologies of the Future*. London and New York: Verso.

Kurzweil, Ray. "The Law of Accelerating Returns", http://www.kurzweilai.net/the-law-of-accelerating -returns [acc. 8 Sept., 2015]

Keep, Christopher J. 1999. "The Disturbing Liveliness of Machines." In *Cyberspace Textuality*, ed. by Marie-Laure Ryan, 164–181. Bloomington: Indiana University Press.

Luhmann, Niklas. 1990. *Essays on Self-Reference*. New York: Columbia University Press.

Maturana, Humberto R. and Francisco J. Varela. 1992. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: Shambala.

McCaffery, Larry. 1988. "The Fictions of the Present." In *Columbia Literary History of the United States*, ed. by Emory Elliott, 1161–1177. New York: Columbia University Press.

Ryan, Marie-Laure. 1999. "Cyberspace, Virtuality, and the Text." In *Cyberspace Textuality*, ed. by Marie-Laure Ryan, 78–107. Bloomington: Indiana University Press.

Siivonen, Timo. 1996. "Cyborgs and Generic Oxymoron: The Body and Technology in William Gibson's Cyberspace Trilogy." *Science Fiction Studies,* 23 (2): 227–244.

Sterling, Bruce. 1991. "Preface from *Mirrorshades*." In *Storming the Reality Studio*, ed. by Larry McCaffery, 343–348. Durham: Duke University Press.

Stiegler, Bernard. 1998. *Technics and Time, 1*. Stanford: Stanford UP.

Suvin, Darko. 1979. *Metamorphoses of Science Fiction: On the Poetics and History of a Literary* Genre. New Haven: Yale University Press.

Artificial intelligence in science fiction as a model of the posthuman situation of mankind

Artificial intelligence. Cognition. Cyberpunk. Posthumanism. Science fiction.

The article focuses on the way in which science fiction genre and scientific texts alike model the ideas of Western culture concerning the functioning of the human mind (brain or intelligence), its knowability, and the probability of its successful simulation. The most problematic issue arising is that this kind of text rests both on the idea of absolute knowability of the human mind (thus stepping outside the "strange loop" defined by Douglas Hofstadter) and on the belief in the possibility of creating such an artificial live system that would reconstruct exactly this "loop". The article sees the distinction between scientific and literary texts concerning artificial intelligence as problematic – the resulting "hybridity" is further employed and enhanced in connection with theories regarding "the posthuman". Building upon the theories of autopoietic (self-referential) systems, the article concludes with an outline of a "theory of hybrid reading".

Mgr. Miroslav Kotásek, PhD.
Ústav české literatury a knihovnictví
Filozofická fakulta Masarykovy univerzity
Arna Nováka 1
Brno
Česká republika
18163@mail.muni.cz