The Singularity of AI in the Eyes of a Jester

Paul Martin
2000

A Comparative Analysis of the Inevitable Dawn of Artificial Intelligence as Presented in William Gibson's *Idoru* and Masamune Shirow's/Mamoru Oshii's *Ghost in the Shell*.

"If poets are the unacknowledged legislators of the world, science-fiction writers are its court jesters."
-- Bruce Sterling in the preface to William Gibson's *Burning Chrome*

Introduction

If people knew what scientists are up to, they would not be sleeping as calmly as they do today. If only they knew, they would read more carefully what the cyberpunk authors have to say.

The purpose of this work is not only to compare the pictures of Artificial Intelligence (hereafter referred to simply as AI) included in two major works of cyberpunk genre, but also to show the connection between those images and the reality we all live in or its nearest future. So what is the future of AI depicted in cyberpunk works like ? And if, as Sterling suggests, science-fiction writers pass on a cautionary message to the contemporary society in a satirical disguise, are these pictures really parallel to our reality ? And if they are, should we be afraid ?

The subject matter is Artificial Intelligence. The environments that house the AIs described hereafter are the environments of cyberpunk literature. To understand the findings and conclusions of this work, it is necessary to know the meaning of both these notions.

Defining Basic Notions : Introduction to Cyberpunk

Cyberpunk has been present on the literary scene for almost twenty years now, being the most trendy and mainstream sub-genre of science-fiction and, although announced dead by some critics, has not been replaced by any other science-fiction movement so far. But how should we define what cyberpunk is and distinguish it from what is merely cyberpunkish ? Appignanesi points out that the major feature of the cyberpunk world is a "total intrusion of technology into human lives" (129) and this may be used as a sufficient definition : cyberpunk talks about very common people, often punks, who are lost in a society of high technology. A useful addition to it is provided by the author who is considered one of the founders of the genre, William Gibson. In one of his interviews he stated that it "is about the present. It's not really about an imagined future. It's a way of trying to come to terms with the awe and terror inspired... by the world in which we live". This opinion is in accordance with the opening quotation from Sterling. Both opinions, too, answer the question of why the

examples from this particular genre were chosen for the sake of comparison shown in this very work. Indeed, there is a reason for doing this : the resemblance to reality cyberpunk works bear.

Obviously, there are other genre-specific features that characterize cyberpunk : like the presence of the Internet (that is often combined into one entity with virtual reality and takes up different names : Cyberspace, Matrix, Multiverse, partly because it was invented in cyberpunk literature before it was invented in reality and named the Internet) or the influence of almighty organizations in the form of international corporations or mafia-like criminal orders of different kinds. These elements, too, appear in the works described herein, but they are just other elements of a picture already defined.

Defining Basic Notions : Introduction to Artificial Intelligence

Let us now consider the term Artificial Intelligence. What is AI then ? And do we live in the age of AI just now ? Both these questions can be answered together.

Yes, we do live in the age of AI if we assume that Artificial Intelligence is a capability of solving by a machine (i.e., hardware) or by a program run on this machine (i.e., software) specific problems that are precisely defined (which means that the solver, either a program or a machine, is provided with algorithms for solving them or have such algorithms built into itself). A machine (either hardware or software) that possesses such capability would also be called AI. We already have machines that translate documents in real time (like babel fish on the Internet) or play chess and win with grand masters (like the computer Deeper Blue).

And no, we can only hope (or fear) to be living in this age if we attribute such a machine to possess the independence of intelligence, this very special human-like intelligence, the ability of thinking.

In order to clarify this picture we need to define one more notion : thinking. The author of this work proposes to use here one of the simplest definitions of thinking, one provided by Buller, who adopts Turing's test (called, at its very beginning, the imitation game), in which a judge is having a conversation (by means of a computer screen or any similar device, in order not to see the interlocutors) with a human and a machine (9-11). He tries to find out which one of the beings "on the other side" of a computer screen is AI by asking questions (Buller 9-11). The AI's task is to pretend to be a human being : if a machine is undistinguishable from a human, it thinks (Buller 9-11).

Obviously, that definition is only satisfactory if we interpret thinking in terms of answering questions, or, more generally, problem-solving. An AI that possesses this capability will differ from the one controlled by algorithms because of its ability to solve (or, at least, to try to solve) any kind of problem, instead of solving just the problems it is programmed for. But what about consciousness ? Indeed, when we speak of AI, we usually think of a conscious being, similar to a human one. But consciousness can be treated as a secondary function of an artificial brain. This question will be discussed in detail in the second part of this work, entitled The Psychology of AI.

No wonder cyberpunk that, as mentioned before, presents the caricaturized image of the contemporary society, is so interested in AI. For there are numerous premises that scientists may be on the verge of constructing AI. The likeness of this act of creation is further described in the first part, called The Birth of AI : They are Inevitable.

Many cyberpunk works depict AI in various forms. To give just a few examples, consider the cyberpunk trilogy (*Neuromancer*, *Count Zero* and *Mona Lisa Overdrive*) by William Gibson, where the deeds of an almighty AI,

called Wintermute, are the perpetrating power behind all the events that happen to the main characters. Or films like James Cameron's *Terminator* or the Wachowski brothers' *The Matrix*, describing the world in which AIs have taken over (perspectives of AI are more closely discussed in the third part, The Future of a Newborn AI : They Have Something in Common). And these are just the most significant examples.

Nevertheless, for the purpose of my work I have chosen to discuss and compare the AI described in William Gibson's *Idoru* to the one depicted in Masamune Shirow's graphic novel (or manga) *Ghost in the Shell* and its filmed version (anime) by Mamoru Oshii. Why those two ? They seem to be the AIs closest to our reality, not only because of the specific features of cyberpunk itself, discussed above, but also because in both literary works the AIs are described as the very first in the world and, as such, similar to the hypothetical AIs that could possibly be created in the world we live in (see the next part).

In *Idoru* the AI is the title character, Rei Toei. Idoru originally means idol in Japanese and that is exactly what she is : an extremely popular pop singer. However, she is not a human singer, but a virtual idol, because that is what idoru means nowadays : a simulation of a person that neither really exists nor has ever existed in a real world. It is a "personality-construct, a congeries of software agents, the creation of information-designers" (Gibson 92; ch. 13), in other words, a computer program in a form of a girl. Such programs are popular in Japan even now (in much more primitive form obviously) and people do treat them like real TV stars. The difference is this idoru, although she has no material form and can only appear as a hologram, is real : she evolved to become an AI with a human-like consciousness. Therefore, she is the first real AI in the world in which AIs used to be only a theory, just as in our world, so we may assume that Idoru portrays our hypothetical future.

The AI described in *Ghost in the Shell* is very similar. Here, it evolves from a spy program, code-named Project 2501, to the form of an independent, intelligent and dangerous terrorist, nicknamed Puppet Master.

The Birth of AI : They are Inevitable

Are they really ? In the discussion initiated by Vernon Vinge, most of the participating science-fiction writers admitted that he is right, predicting that mankind will have created AIs within some thirty years and they will most probably become our successors (Cyran 74-76).

Singularity. That is what they call it. Singularity, according to the writers mentioned above, is the point in time (we may also call it a horizon of prediction) after which humans will invent both real nanotechnology and real AI, the point in time after which all that happens is unpredictable and incomprehensible for us, as all the rules that govern our world of today become absolutely insignificant (Cyran 74-76).

Most of it sounds like a pure fantasy. But science is a funny thing. It reaches beyond human comprehension. What was laughed at yesterday often becomes a standard of today and today's ideas are sometimes incomprehensible even for their inventors : science is creating itself, one may say. And so, serious scientists that work upon creation of thinking machines agree with authors of science-fiction literature, saying that completion of their project is just a matter of time as they already know how to do it and they just lack the appropriate devices. Andrzej Buller states that the participants in the international CAM-Brain project aimed at creating AI have already invented the architecture of a future artificial brain the only problem is that they do not have the hardware capable of holding enough software neural structures and of evolving to meet the needs of evolving software (50-62, 69-79).

From what was already said about the origins of the AIs described herein one may easily and truthfully conclude that both of them were not created as such by humans, but they evolved from human designs of

different purposes : the idoru was just a simulation of a singer and Project 2501 was a spy program.

This act of pseudo-auto-creation even makes Puppet Master from the anime *Ghost in the Shell* reject his being called AI and call himself "a life-form that was born in the sea of information". In fact AI is the proper term to describe him. He was artificial, as according to Buller the real AI is not going to be built from a scratch to its very final form (17-49). Instead, human scientists will devise basic hardware and software structures and let them evolve to achieve AI perfectly adjusted to living in certain environment (Buller 17-49). It has already been proved that programs show a tendency to evolve very much like living organisms do according to Darwin's theory (Buller 17-22). One may argue that comparing software to living creatures is exaggeration, but all the living things are also coded, just as all computer programs are. What else is DNA if not a sophisticated code that contains information about the structure of organisms, their brains included ? Obviously, the question of consciousness, even soul maybe, remains and this is further discussed in the next part.

One can see that both cyberpunk works compared promote the same idea that absorbs the minds of contemporary researchers : AI cannot be created but it can evolve just as it happens in natural world with gene-coded organisms. Gibson states that "the genuine AI (is) most likely to evolve in ways that (have) least to do with pretending to be human" (247-248; ch. 37) and goes as far as to suggest "that AI (may) be created accidentally and that people (may) not initially recognize it for what it (is)" (248; ch. 37). Similarly, Project 2501's evolution is the result of his interacting with different sources of data and (what is not directly stated neither in the novel nor in the film, but is a logical conclusion of what is stated there) of his residing in different pieces of hardware, probably very sophisticated ones as he was designed as a military program and military technology is usually (and definitely in the world depicted by Masamune) the most advanced one. Therefore his designing as an application meant for penetrating different hardware environments and extracting data resulted in his turning against his designers, as these environments made him develop independent intelligence and free will.

The Psychology of AI : Ghost in the Machine

Will AI be similar to human beings in terms of psychology ? Nick Bostrom, one of the writers participating in the discussion mentioned herein, claims that a crucial factor in shaping the psyche of AI is its creators and that the moral standards of AI will reflect the norms obeyed by its designers (Cyran 74-75).

This approach, however, is reflected in neither of the cyberpunk works discussed here. Rei Toei may take the sensibility to human problems after their designers (which is especially shown in her attitude towards Chia), but she surpasses her creators in terms of wisdom and insight and, therefore, seems to judge the events around her with a cold certainty of a superior being that recognizes human emotions, but, as a being practically immortal, is beyond them. Project 2501 behaves likewise, combining in itself features of a rebel child and those of a being sensitive as a human, but more important than any of our kind, for whom humans are only pawns in his plan, used according to his will... Are we dealing here with psyche of children gods ? And who is right then, Bostrom or the creators of those cyberpunk visions of AI ?

There are some premises that Bostrom is indeed wrong. Witold Marciszewski observes that machines are free from problems (both practical and theoretical) and doubts that biological organisms have to deal with and, therefore, are not driven by the same stimuli and cannot think like human beings (78-81).

Actually, there are views even more radical than that. According to them, it is possible that consciousness is only a by-product of evolution and that more advanced intelligent creatures (such as AIs) may get rid of it, as it would not aid them in solving problems and, therefore, in being intelligent (Marciszewski 18-19).

Buller is more optimistic, but also denies the necessity of AI being only a reflection of its creators' moral norms. He agrees with Alan Turing, saying that engineers can only build an artificial mind of such parameters that it will be a proper housing for soul (Buller 14-15). In his opinion, it is up to God to create or deliver a soul to a mind provided by humans and, therefore, engineers cannot replace God in creating souls and are nothing more for AI than parents for children (Buller 14-15). We may conclude that, just as opinions of children do not always reflect their parents' views, a psyche of any AI may not be the exact counterpart of the psyches of its creators.

This idea of creating housing for souls may, of course, seem blasphemous. But if people assume presence of souls in creatures coded by genes, it should be possible as well to seek it in entities coded in a different, yet similar, way. The idea of people being only providers of housing for something (be it a psyche or a soul) created independently (be it a spontaneous creation or an act of God) somehow goes together well with the fact that neural networks that are considered the basis for constructing AI are mystery even to their builders. Scientists can design neural networks and influence their evolution by promoting learning certain models of behavior, but they do not fully understand how neural networks work (Buller 38).

Bearing these facts in mind, one cannot avoid observing that the title *Ghost in the Shell* is very significant. It suggests the existence of incorporeal energy that, together with a body (be it an organic body or a mechanic shell) forms a thinking entity. The presence of this energy is required and without it even an organic person cannot be considered conscious. The same force would then be responsible for emotions that spontaneously evolved in Rei (like her love for Rez).

What differs both AI from each other is the stimuli that have aided in shaping them. It is similar to the situation of two children growing up in different circumstances. In case of Project 2501, the environment is the above mentioned "sea of information". Basing his evolution on pure (and often military) data, free from insignificant emotional background, he develops and perfects male personality of cold logic, sometimes even cynical or merciless, as he uses and even abuses people to do his bidding, regardless of harm he does unto them. Unlike him, Rei was originally designed as a female personality. Moreover, she was meant to be the final result of a process of creating "desiring machines, aggregates of subjective desire, an architecture of articulated longing" (Gibson 237; ch. 35). A combination of these factors produced an AI that is more devoted to feeling than to knowing.

The Future of a Newborn AI : They Have Something in Common

As the above presented stages of comparison have already proved, both Puppet Master and Rei Toei share different features. It is likewise with their goals. Both artificial entities know exactly what they want to do and adjust their ways of life accordingly to make it possible. Both AIs are expansive and their want to further evolve and spread. They want to reproduce. In order to do this, they choose partners that are living creatures, human beings but very different from the rest of humans and that enables them to understand the purposes of the AIs. In case of Puppet Master it is major Motoko Kusanagi, who is a member of special forces, a person of anti-totalitarian (a bit anarchistic perhaps) views. She is, therefore, a perfect defender of Puppet Master, who is pursued by corrupted government agents. But, what is more important, she is almost one hundred percent cybernetized (only her mind is supposed to be human, and even this she is not absolutely sure of), therefore alienated and unsure of her own personality. These qualities let her understand Puppet Master, who is an artificial entity, the only one of his kind and, because of that, alienated in a way similar to hers. Rei chooses Rez, a charismatic leader of a famous rock band, whose fame and eccentricity have also isolated him from the

society, making him one of the kind, just what she is, and whose wealth and influences make him a perfect person to help her achieve her goals.

The notion of partners in both cases means something different. Puppet Master becomes a part of Motoko's mind, reproducing psychically and creating a new entity out of this union, an entity capable of reproducing physically in order to give beginning to a new race. Rei marries Rez virtually but, as one may suppose, probably intends to have sex with him physically, as the nanotech unit Rez's people steal from Russian mafia, described as "the medium of (their) marriage" (Gibson 257; ch. 38), can help them make Rei a real person by building a body for her mind. This tendency of AIs to reproduce is, again, very truthful in its parallel to the reality of AI that already exists. Although the real AI does not exist yet, the evolution of experimental applications that are basis in researching AI is driven by evolutionary strategies of survival that are based on reproducing of the most healthy specimens. Once again, the cyberpunk authors only pretend to write about the future.

Conclusion

The picture emerging from the above comparison is not as pessimistic as one may expect. It is true that the cyberpunk works compared, in their significant similarities and astounding resemblance to the real achievements of contemporary science and technology, give us warning about what we may expect from our world in not so distant a future. But these stories are not tales of future horrors rather of future changes. The message is clear : something is going to change. But if there is some truth in these tales, the direction of change depends on the decisions thinking entities (either humans or AIs) will take. And it does not really change anything in our picture of the world as it has always depended on such decisions the only difference is that these were taken by humans only up to now...

Bibliography

1. Gibson, William.*Idoru*. London : Penguin Books. 1997.
2. *Ghost in the Shell*. Dir. Mamoru Oshii. Masamune Shirow/Kodansha/Bandai Visual/Manga Entertainment. 1995.
3. Appignanesi, Richard. *Postmodernism for Beginners*. Cambridge : Icon Books Ltd. 1995.
4. Buller, Andrzej. Sztuczny mózg. To ju nie fantazje (Artificial Brain. Not a Fantasy Any More). Translation mine. Warszawa : Prószy ski i S-ka. 1998.
5. Marciszewski Witold. Sztuczna inteligencja (Artificial Intelligence). Translation mine. Kraków : SIW Znak. 1998.
6. Cyran Janusz, ed. "Requiem dla ludzko ci ?" ("Requiem for Humanity ?"). Translation mine. Nowa Fantastyka Oct. 1999 : 74 76.
7. Anonymous collective work. The Cyberpunk Project. Authors' Definition. 1999.