What is It Like to be Wintermute?

Virtuality and Consciousness

Blanka Earhart
Center for Advanced Enquiry in the Interactive Arts
Planetary Collegium
Plymouth University
Plymouth, United Kingdom
4blanka@gmail.com

Abstract—**The world around us is becoming more complex and better connected every day. Humans and machines relate among themselves in a network of information and action. Through these relationships human mind and consciousness is changing to accommodate the environmental demands. While humans are becoming technologically augmented, new forms of existing in the world become possible thanks to the electronically connected system. Despite these new developments, human consciousness is yet to be fully understood. Presently, there is no conclusive definition of this phenomenon. In the mean time, the evolution of human consciousness carries on. Thanks to the shifting technology, human existence is increasingly reminiscent of the reality presented in cyberpunk novels. Interestingly, the approach of this literary genre reframes the questions circulating around the subject of consciousness. It is perhaps only through fiction that we are able to contextualize our present position vis-à-vis our technology. This view has far-reaching repercussions in philosophy of mind and technology.**

Consciousness; cyberpunk; technology; Internet; evolution; agency; intelligence; teleonomy

I. INTRODUCTION

The complexity of the technological inventions created at the beginning of the XXI century is exponentially increasing. Sophisticated modes of communication are becoming ubiquitous among humans and machines alike. The complex systems used by humans in every aspect of our lives from medicine to entertainment are increasingly connected. At the same time, they enable gathering, storage and exchange of information. This electronically linked system is a fertile ground for an onset of new forms of being in the world. It is up for a debate if the newly created yet still nebulous entities can have some autonomy and approach anything close to human consciousness. There are however certain qualities common to both carbon-based and electronic intelligence.

While the electronic world is becoming more sophisticated, human experience in the world is becoming augmented and fuses with technology. Different modes of mind and body enhancements are available to ever-increasing numbers of people. In fact often, we find ourselves living through our technology. Be it virtually or in reality our consciousness is filtered through electronic and mechanical devices that make us see farther, move faster and know more. By using technology we alter our consciousness and create our sense of self in an unprecedented ways. In this process, human consciousness is becoming closely related to the machines that it produces. This proximity creates a feedback loop which brings us ever more closely to our product and confuses the boundaries between human and artificial.

The two-fold development of connected artificial systems and technologically enhanced humans initiates a new kind of consciousness and raises questions about its boundaries. At which point can we talk about a system as conscious? Additionally, what is the nature of such consciousness and how does it relate to our familiar, even though little understood, human consciousness? When dealing with complex virtual systems with built-in logic, such as social networks, computer viruses or any sophisticated software, there are moments when the experience becomes uncanny for the human participant. It is a result of the system behaving in a manner reminiscent of our consciousness. One may recognize traits that under different circumstances would signal interaction with a human being. This sense of uncertainty about the participants' identity signals that our every-day experiences are approaching the scenarios invented in literary genre such as cyberpunk. The fictional stories throw a different light on the traditional questions regarding consciousness.

II. APPARENT CONSCIOUSNESS

Despite the fact that there is no simple definition of consciousness, certain characteristics of a system can signal its possibility. An agglomeration of such traits creates a sensation of encountering a sentient entity. In every day life, this uneasy feeling is experienced when inanimate

object appear to be invested with life. For example, plastic bags blown by the wind behave as though they were acting out of their own volition. This perceived life-like behavior I will designate as apparent consciousness. I also propose that agency, intelligence and teleonomy contribute to creation of this appearance.

Agency is the capacity of a system to act in a coherent manner. One characteristic of complex software is its ability to execute actions and movements of its own accord. Large systems, such as a flock of geese or a message board, behave in what appears to be spontaneous manner that is united through action. Further, understanding what the system is doing, gives one the ability to name the collective and perceive it as a unified field. In the virtual world, the most basic activities performed by a social network, for example, are propagation, accumulation of posts according to a program and formation of new affiliations. As a result, this unity in action endows the collectivity with a sense of agency.

Moreover, complex software often exhibits intelligence specific to systems that are auto-corrective in a biological and cybernetic sense. A piece of software, operating within certain logic, has the capability of organizing content in what appears to be an intelligent manner. The single mind, acting through the multiplicity, emerges as a function of the system's directionality. Theoretical biologist Ernst Mayr explains that the direction is inspired by a goal. "This end point might be a structure, a physiological function, the attainment of a new geographical position, or a consummator act in behavior. Each particular program is the result of natural selection, constantly adjusted by the selective value of the achieved end point [1]."

Mayr suggests that biological systems exhibiting apparently purposeful behavior can be viewed as teleonomic, which is a term derived from the Greek telos meaning goal or end point. "All teleonomic behavior is characterized by two components. It is guided by a 'program' and it depends on the existence of some end point, goal, or terminus that is foreseen in the program that regulates the behavior [2]." Considering this definition, software in its interaction with human agents is a teleonomic phenomenon. First, it is running a 'program,' which enables it to incorporate electronic and human elements as its components and second, there is usually an end goal in a form of a document, feed, profile or any other outcome of the operation.

Further, one engages the software through a portal such as a website, widget or other type of user interface. Establishing a one-on-one relationship further perpetuates the sensation that a person is engaging with a discrete and responsive entity. The intimacy of a relationship coupled with the ability to perform actions, intelligence executed through code, and directionality of the system gives the accumulative impression of sentience. Thus manifested apparent consciousness floats on the edge of the system escaping our reason and creating an uncanny sensation.

III. BEHAVIORAL AND SUBJECTIVE CONSIDERATIONS

For the time being, the most coherent conversation about consciousness revolves around its characteristics and therefore appearance. The full definition of consciousness escapes layman as well as scientists. It is perhaps because we are unable to gain perspective while existing in the middle of the experience of being conscious. As much as questions revolving around the subject baffle scientists, they are often addressed directly in science fiction writing. The father of cyberpunk literary genre, William Gibson, is often credited with prediction of the Internet and the behavior that comes as a result of human-machine interaction. In his seminal work, *Neuromancer*, he explores different types of consciousness, both silicon and carbon based.

In the novel, technologically altered humans struggle against and work in tandem with artificial intelligence and lesser electronic entities. The story revolves around a plutocratic family preserving its power through cryogenics and cloning, which stands in for extreme form of inbreeding. In a world where new generations of cloned humans mange and are being managed by artificial intelligence, it is difficult to decide if the silicon or carbon-based life is more real. In fact, the question about one possessing greater reality than the other is rendered null and is replaced by notions of fluid identity. A division line however can be established between individual beings and hive entities. The individuals living on Earth, no matter how augmented, are of a different order of existence than the compound super intelligence and its human family counterpart living on the edge of the world in Villa Straylight.

One of the characters, Wintermute, is a semi-autonomous artificial intelligence tasked with working for the Tessier-Ashpool family. This virtual entity is of particular interest as it relates to some structures populating the Internet today. While Wintermute is housed on a physical piece of hardware in a location on Earth, it moves through the cyberspace and outside world alike hacking and manipulating software as well as human psychology. The only obstacle to its free reign is the Turing lock imposed as a safety precaution to curb the AI. The removal of this lock is what stands between Wintermute and its lost half, another artificial intelligence named Neuromancer. Wintermute's desire to free itself drives the plot of the novel and highlights the conflict between the human family hive and the virtual agglomerate.

The built-in will that drives Wintermute's agency makes it stand out among other AIs. While some contemporary software possesses directionality, none comes close to exhibiting the focused will driving Gibson's IA. The need to merge with Neuromancer becomes a single direction providing the logic to Wintermute's behavior. The AI organizes a complex plot traversing human bodies, minds, hardware and software in order to free itself. While being generally disembodied, Wintermute is a shape shifter taking

on personalities of other characters in order to communicate with humans at their comfort level. However, the AI's nature is not human and it is impossible to tell what it is like to be Wintermute. While the author points at the possibility of the AI having a subjective experience, the fact is never clear. It is impossible to judge if the desire for Neuromancer creates an internal feeling approximating human emotion.

The mind philosopher Thomas Nagel holds that while being an intrinsic part of consciousness, the subjective character of experience is impossible to detect. In his essay, *What is it Like to be a Bat?* Nagel states the following: "But no matter how the form may vary, the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism. There may be further implications about the form of the experience; there may even (though I doubt it) be implications about the behavior of the organism. But fundamentally an organism has conscious mental states if and only if there is something that it is to *be* that organism – something it is like *for* the organism [3]." Without the subjective experience, the only judgment that can be made about consciousness of a system is based on its behavior and therefore not conclusive. Wintermute therefore may be only a mirage of actions emulating apparent consciousness.

As discussed previously, complex software today exhibits the traits of apparent consciousness. Additionally, highly connected virtual world enables such software to propagate and recruit humans and hardware as part of its process. Twitter, for example, propagates itself through the Internet according to its program making pacts with humans that wish to participate in its functionality. Both human and electronic counterparts contribute to the overall phenomena that we have learnt to view as one entity. Emulating apparent consciousness Twitter exhibits agency since it acts in a coherent manner as well as intelligence facilitated by the code. It further possesses teleonomy as a function of directionality towards the participating people and displayed topics. In its behavioral essence the software already existing on the Internet performs similarly to Wintermute. The fictional AI's will is an exaggerated and complex directionality that falls under the category of behavior. Will is a functionality that can be programmed in.

The human search for the precise definition of consciousness is rendered irrelevant in Gibson's novel. Presently, philosophy and science are attempting to establish precise criteria for designating a conscious system that will take into consideration the behavioral as well as subjective aspects. Gibson bypasses this problem and shows us a world where humans as well as machines are evolving into entities escaping traditional criteria. The Tessier-Ashpool family is human in form yet their relationships to each other and the outside world are altered by three generations of cloning. Their isolated existence allowed them to move beyond social and biological norms. In effect, their consciousness is closer to a hive, where an individual stands for the totality and the totality is a conscious entity. Lady 3Jane is herself, her own mother, grandmother, and her entire family at the same time. Simultaneously, the AIs are evolving to resemble humans in the complexity of their behavior yet are retaining their hive-like structure. In fact, they desire a polymorphous existence and merging to create larger and better-adapted entities.

IV. CONCLUSIONS

The question of consciousness in thus constructed world becomes secondary to the possibilities for proliferation and manifestation. For Gibson the behavioral and apparent reading of consciousness is sufficient. The entities exhibiting agency, intelligence and directionality may or may not be conscious yet they create tangible consequences in the world of the living. These consequences force us to treat them seriously despite their unresolved conscious status. What we learn form Gibson's theoretical exercise is that consciousness is not a stable phenomenon defined in a number of statements. Instead, consciousness transforms quicker than our ability to define it. As soon as one aspect of it is fixed, new permutations of the phenomena make the original definition insufficient. Additionally, as entities submerged in the experience of consciousness, we attempt to use objective tools to describe a subjective experience. Since our state of being is incompatible with our modes of enquiry and communication, we fail to capture the essence of the experience.

Ultimately, there may be more than just one kind of consciousness. While the fixed definitions are failing us, we still need to produce a meaningful way of contextualizing our experience in the world. Perhaps the essence of our own experience depends on the interactions with other entities since through the action we can define ourselves. The important question is not how consciousness can be defined in the outside objects. Instead, we should ask how does it evolve throughout the territory of one's own phenomenology. The introspection may reveal that we are already widely polymorphic creatures, distributed and merged with other entities both physical and virtual. Today, it is a question of choosing what kind of consciousness we would like to become.

REFERENCES

[1] E. Mayr, "Evolution and the Diversity of Life," Cambridge: Harvard University Press, p. 390, 1997.

[2] E. Mayr, "Evolution and the Diversity of Life," Cambridge: Harvard University Press, p. 390, 1997.

[3] T. Nagel, "What is it Like to be a Bat?," The Philosophical Review, Vol. 83, No. 4 p. 435, October 1974.