

## Stéphanie Walsh Matthews\* and Marcel Danesi

# AI: A Semiotic Perspective

**Abstract:** Artificial Intelligence (AI) has become a powerful new form of inquiry into human cognition that has obvious implications for semiotic theories, practices, and modeling of mind, yet, as far as can be determined, it has hardly attracted the attention of semioticians in any meaningful analytical way. AI aims to model and thus penetrate mentality in all its forms (perception, cognition, emotion, etc.) and even to build artificial minds that will surpass human intelligence in the near future. This paper takes a look at AI through the lens of semiotic analysis, in the context of current philosophies such as posthumanism and transhumanism, which are based on the assumption that technology will improve the human condition and chart a path to the future progress of the human species. Semiotics must respond to the AI challenge, focusing on how abductive responses to the world generate meaning in the human sense, not in software or algorithms. The AI approach is instructive, but semiotics is much more relevant to the understanding of human cognition, because it studies signs as paths into the brain, not artificial models of that organ. The semiotic agenda can enrich AI by providing the relevant insight into human semiosis that may defy any attempt to model them.

**Keywords:** abduction; cybernetics; embodiment; modeling systems

---

\*Corresponding author: **Stéphanie Walsh Matthews**, Ryerson University, Toronto, Canada, e-mail: [swalsh@arts.ryerson.ca](mailto:swalsh@arts.ryerson.ca)

**Marcel Danesi**, University of Toronto, e-mail: [marcel.danesi@utoronto.ca](mailto:marcel.danesi@utoronto.ca)

## 1 Introduction

In the 1980s, Artificial Intelligence (AI) became a dominant mode of inquiry into the nature of cognition, shedding light on the nature of intelligence in ways that continue to hold promise for truly revolutionizing research in both psychology and semiotics. AI is an ever-broadening field that has implications for semiotic theories, yet, as far as can be determined, has hardly attracted the broad attention of semioticians in any significant way. However, AI attempts to study the nature of mentality in all its forms (perception, cognition, emotion, etc.) by modeling it in computer software, and, more radically, to build artificial minds

that will purportedly surpass human intelligence in the imminent future. One area has proven to be an exception: the cybersemiotic movement, spearheaded by Søren Brier (2007), which aims to integrate principles of cybernetics with basic principle of semiotics, extends the semiotic paradigm to cover the study of semiosis in humans, animals, and machines, putting the spotlight, however, on the uniqueness of human semiosis, which is autopoietic, that is, self-organizing in a creative, not deterministic fashion.

The purpose of this paper is to look at AI from a semiotic perspective in view of expanding semiotic method to encompass implications of AI for the study of signs and sign systems. The underlying assumption in AI is that human intelligence is not unique – an idea that dovetails with current philosophies such as posthumanism and transhumanism. These are based on the assumption that technology will guide the future progress of the human species, obliterating its historical dichotomies based on human ideas such as gender, race, age, etc. Machines do not make such distinctions. The AI movement may be seen in this philosophical light – as a means to “improve” humanity – a goal that recalls Descartes’ dream of a world that obeys only the laws of logic and mathematics (Damasio 1994).

## 2 The AI approach

Unlike physical objects or natural phenomena, the mind cannot be studied objectively as a separate entity. It cannot be taken out of the body for observation or inspection. The mind is a result of an interaction between the body and brain. It is, in other words, an epiphenomenal by-product of physiological and neural activities working in tandem.

The AI approach of studying the mind without the body can be traced to Descartes’ *De Homine* (1633), in which he gave the mind-body problem a radical formulation. He argued that the two are distinct entities, and that the body worked like a machine that was animated by “animal spirits” flowing through the nervous system. When these spirits reach the pineal gland, which is the “seat of thought,” humans become aware of the animal spirits. The mind could thus double back on the body by instigating the flow of the animal spirits to a particular part, activating it as the case may require. Contemporary neuroscience has largely dismissed Descartes’ “error,” as psychologist Antonio Damasio (1994) termed it. However, Cartesian dualism is still a subtle factor in AI research.

The actual starting point for AI can be traced to the rise of cognitive psychology in the 1960s – a branch of psychology that started adopting insights and terms from AI, seeking parallels between the functions of the human brain and those of the computer, such as the “coding,” “storing,” “retrieving,” and “buffering.” The premise was to model mental phenomena on computers so that they could be observed in their functions. Ulrich Neisser (1967: 6) put it as follows:

The task of the psychologist in trying to understand human cognition is analogous to that of a man trying to discover how a computer has been programmed. In particular, if the program seems to store and reuse information, he would like to know by what “routines” or “procedures” this is done. Given this purpose, he will not care much whether his particular computer stores information in magnetic cores or in thin films; he wants to understand the program, not the “hardware.” By the same token, it would not help the psychologist to know that memory is carried by RNA as opposed to some other medium. He wants to understand its utilization, not its incarnation.

Neisser realized, however, that the computer metaphor, if brought to an extreme, would actually lead psychology astray. So, only a few pages later he issued the following warning (Neisser 1967: 9): “Unlike men, artificially intelligent programs tend to be single-minded, undistractable, and unemotional; in my opinion, none does even remote justice to the complexity of mental processes.” As cognitive psychology progressed throughout the 1970s, eventually it came under the direct influence of AI, leading to the emergence of *cognitive science*, as a new science of the mind. From the outset, two main schools within this science surfaced. One was based directly on the notions and methods of AI researchers, viewing the mind as essentially a computing device, separate from lived reality. As Gardner (1985: 6) put it, the guiding assumption of this “strong” version is that there exists “a level of analysis wholly separate from the biological or neurological, on the one hand, and the sociological or cultural, on the other,” and that “central to any understanding of the human mind is the electronic computer.” The second version, known as the “weak” version, aimed to study the mind-body problem in a new light, that is, in terms of how the two are interactive agents in the generation of cognition. This version is sometimes also called the “embodied cognition” movement. It rejects the idea that the mind can be modeled separately from bodily processes and that cognition is hardly an abstract entity, but rather an epiphenomenal outgrowth of bodily experiences. The metaphor of the mind as a processing container is thus rejected within embodied cognition.

The underlying premise of the strong version is that we can best understand intelligence through algorithmic models. Human beings are thus a particular

kind of “Turing machine,” a concept developed by the great mathematician Alan Turing (1936). Turing described such a machine as an “automatic typewriter” that used symbols instead of letters. It could in theory carry out any recursive function – the repeated application of a rule or procedure to successive results or executions. Recursion became, and still is, a guiding principle underlying the strong version of AI. However, Turing himself was skeptical about comparing his machine to human cognition because of the nature of logical rules themselves. A few years earlier, Kurt Gödel (1931) had showed that there is always some statement in a logical set of rules or propositions that is true, but not provable in it. Turing referred to this as the halting problem, which he articulated in the form of a question: Is there a general procedure for deciding if a self-contained computer program will eventually come to a halt or run forever? Turing concluded that it is impossible to construct an algorithm (set of rules) that always leads to a yes-or-no answer to this question.

The strong version claims that all human activities, including emotions and social behavior, are not only representable in the form of computer models, but also that machines themselves can be built to think, feel, and socialize. The following early citation from Konner (1991: 120) is a case in point:

What religious people think of as the soul or spirit can perhaps be fairly said to consist of just this: the intelligence of an advanced machine in the mortal brain and body of an animal. And what we call culture is a collective way of using that intelligence to express and modify the emotions of that brain, the impulse and pain and exhilaration of that body.

In effect, strong AI aims to take the mind out of the body, so to speak, and study it in the machine. The embodied cognition movement, on the other hand, sees the body as critical in the production of mind. This version actually traces its ideas, consciously or not, to the significant work of the biologist Jacob von Uexküll (1909). The key to understanding the nature of mind lies in the anatomical structure of an organism. Animals with widely divergent anatomies do not live in the same kind of mind world, because each species filters information according to its own particular *Bauplan* – the mental modeling system that allows it to interpret the world in a biologically determined way. A machine also has a *Bauplan* – the human-made computer program. But this is not grounded in the body, but in wires and electrical impulses. The computer and the human being have, in effect, widely divergent “anatomies” and, as von Uexküll would have it today, do not “live” in the same kind of mind world. AI and natural intelligence are not equivalent, only analogous in the best of cases – and possibly so because of the inherent metaphors we use to describe them.

The human *Bauplan* activates emotional areas of the brain that may be beyond algorithmic modeling. The limbic system – which includes portions of the temporal lobes, parts of the hypothalamus and thalamus, and other structures – has been found to play a larger role than previously thought in the processing of cognition (Damasio 1994). It is not clear how AI can produce a model of the limbic organ.

For semiotics the mind-body problem is resolved in terms of how these two agents produce a system of interpretation through semiosis. A computer processes input information to produce a required output; a human mind *interprets* information, even if it is not clear what kind of output is involved. In effect, the human brain is an “interpretive machine,” not an algorithmic one. And its activities are governed in large part by actual lived experiences. This is, actually, the underlying premise of so-called phenomenology, originally developed by the Edmund Husserl (1891), who wanted to understand how awareness of sensations and emotions unfold in tandem with rational processes. Modern-day phenomenology characterizes the forms of consciousness as *phenomena*, and the processes involved in consciousness-formation, such as perception and desiring, as *acts*. These are related to objects of consciousness and thus are also considered to be phenomena. The link between phenomena and acts is intentionality. Phenomenologists also claim that past experiences will limit people’s ability to understand phenomena and thus to act accordingly. The French psychologist Maurice Merleau-Ponty (1942, 1945) conducted a series of significant experiments that showed, in a phrase, how the body and the mind interact to produce meaningful forms, such as words, which then double back on reality to provide interpretations for it. Phenomenology seeks to understand why “meaning” is an intrinsic part of the interconnection between body and mind. It is little surprise, therefore, that phenomenology has achieved favor among semioticians.

### 3 Current research paradigms

Research in AI has become very sophisticated since the 1990s, and AI theorists anticipate that in the next few decades it will be possible to generate a “super intelligence.” This scenario was predicted by Ray Kurzweil in his 2005 book, *The Singularity is near*, in which he maintains that there will AI will autonomously outperform human intelligence – an event known as the *Technological Singularity*, which will occur when an upgradable software becomes self-sufficient without human intervention, thus becoming capable of self-improvements. Each

new self-improvement will bring about an intelligence explosion that will, in turn, lead to a powerful super-intelligence that will surpass human intelligence. Kurzweil predicts that the Singularity should occur around 2045, when AI technologies cannot be stopped by human intervention (see also Kurzweil 2012). By that year, networks of computer chips known as silicon neurons, will be able to mimic with a high degree of fidelity the information-processing functions of brain cells and thus operate at the speed of neurons.

This is truly a remarkable claim, which semiotics clearly needs to address. So, it is worthwhile examining a few of the main tenets and methods of current AI. Computer models known as *parallel distributed processing* (PDP) models have become common, since these are designed to show how, potentially, algorithms can be devised to simulate brain modules that process information holistically. The PDP models perform the same kinds of tasks and operations that, for example, linguistic syntax does (MacWhinney 2000). This type of modeling has produced interesting ideas, the paramount one being that for true AI to emerge, it must produce systems that function through interconnectivity, not as simple sequential algorithms.

But, as Max Black (1962) pointed out at the start of AI, the idea of trying to discover how a computer has been programmed in order to extrapolate how the mind works is fraught with Cartesian dualism – a critique reinforced by physicist Roger Penrose (1989), who emphasized that computers can never truly be intelligent because the laws of nature will not allow it. Aware that this is indeed an effective counter-argument to strong AI, Allen Newell (1991) responded early on by pointing out that the use of mechanical metaphors for mind has indeed allowed us to think conveniently about the mind, but that true AI is not based on metaphor. He summarized his case as follows (Newell 1991: 194):

The computer as metaphor enriches a little our total view of ourselves, allowing us to see facets that we might not otherwise have glimpsed. But we have been enriched by metaphors before, and on the whole, they provide just a few more threads in the fabric of life, nothing more. The computer as generator of a theory of mind is another thing entirely. It is an event. Not because of the computer but because finally we have obtained a theory of mind. For a theory of mind, in the same sense as a theory of genetics or plate tectonics, will entrain an indefinite sequence of shocks through all our dealings with mind – which is to say, through all our dealings with ourselves.

There is no reason to assume that AI is anything more than a set of rules devised by humans to help carry out their knowledge tasks more efficiently. Natural intelligence is based on semiosis, which is a product of an interaction between the body, the mind, and the environment. In strong AI, the assumption

is made that the mind can be extricated from the body and the environment, and work independently of them. A basic principle of semiotic analysis posits, in fact, that expressions, symbols, representations, and traditions are interconnected to each other through historicity and bodily experiences.

Overall, the work in AI that aims to understand how humans think, perceive, and remember is yielding some truly incredible computational feats, from voice to facial recognition technologies. An early approach, which has remained key to all subsequent facial recognition technologies is that of David Marr (1983). Marr attempted to reproduce in computer software the essential features of vision (perception, recognition, etc.), which he did successfully. He then used his algorithms as the basis for developing a theory of human vision; in other words, he sought to explain visual perception, not by working directly with the visual nervous system, but by designing programs to be consistent with the processes known, observed, or suspected to underlie visual perception. This has been valuable in having forced psychologists to reconsider many of their assumptions about perception and to seek out much more clearly formulated explanations of visual processes. But the more critical question for strong AI is the following: Does the computer, following Marr's instructions, really "recognize" objects, people, and events in the same ways that people do? A machine might be capable of "perceiving" an object, so to speak, but it does so according to its own *Bauplan* – the human-made computer program.

## 4 Intelligence amplification

An area of AI that falls outside both the weak and strong versions, comes under the rubric of Intelligence Amplification (IA). This is the use of AI to enhance natural intelligence by amplifying it through prosthetic technologies. Research on IA goes back, in fact, to the emergence of cybernetics, which was introduced in a 1948 book *Cybernetics, or control and communication in the animal and machine* by Norbert Wiener and developed a little later by his *The human use of human beings: Cybernetics and society* (1950). William Ross Ashby then made the science known more broadly in his 1956 book, *Introduction to cybernetics*. These two books were followed up by J. C. R. Licklider (1960) and Douglas Engelbart (1962), who came to be designated as the founders of the IA movement.

IA dovetails with cyborg theory, or the view that physical and mental abilities can be extended beyond normal human limitations by mechanical elements built into the body. It espouses the view that the amalgamation of

humans with machinery and artificial systems is bringing about a veritable paradigm shift in human evolution. A *cyborg* is a human whose functions are taken over in part by various electronic or electromechanical devices, or else whose anatomical or psychological capacities are bolstered by prosthetic technology. Cyborg theory is often inserted into a larger philosophical discourse called posthumanism, or the view that humans should no longer dominate the world but instead merge with animals and machines to create a new world order. The theorist most associated with this view is Donna Haraway (1989, 1991). Since the cyborg is not bound by notions of race and gender, it will rise, claims Haraway, to efface the “isms” of traditional human-centered worlds. She also claims that the cyborg will efface the belief in a Self contained inside the human body as well as the traditional notions of the uniqueness of human consciousness. She calls the cyborg a “posthuman subject” whose identity will undergo “continuous construction and reconstruction.” In posthumanism, humans are just small organic particles in the overall scheme of things and thus there is a need to move beyond archaic concepts of human nature and to establish a society that is without the traditional prejudices and biases. Posthumanism will be taken up in the next chapter.

It should be mentioned that the idea of augmenting human faculties through technology was a central one in the work of communications theorist Marshall McLuhan (1964), who suggested that all technologies were amplifications of human abilities. McLuhan framed this notion in the context of his Four Laws of Media (McLuhan and McLuhan 1988) – amplification, obsolescence, reversal, and retrieval. A new technology or invention will at first amplify some sensory, intellectual, or other human psycho-biological faculty. While one area is amplified, another is lessened or rendered obsolete, until it is used to maximum capacity whence it reverses its characteristics and is retrieved in another medium. A well-known, and now classic, example given by McLuhan is that of print technology. Initially, it amplified the concept of individualism. It did so because the spread of print materials encouraged private reading, and this led to the view that the subjective interpretations of texts was a basic right of all people. In turn, this rendered group-based understanding obsolete until it changed from a single printed text to mass-produced texts, leading to mutual readings, albeit typically displaced in time and space. This allowed for the retrieval of a quasi or secondary communal form of identity – that is, readers of the same text were connected in an imaginary way.

McLuhan always foresaw danger in the enthusiasm over new inventions, such as IA. He warned against the “amputations” that these might bring about. Modern technologies, such as the Internet, may in fact make us mere “spectators,” inclined to abrogate our responsibility to think and act



independently, thus debilitating true democracy and meaningful discourse. McLuhan actually saw the evolution of humanity optimistically, maintaining that people can always become active and express themselves freely and, in fact, challenge the leadership. By understanding what is going on in terms of IA technologies, we are better able to make practical decisions vis-à-vis those amputations that these would otherwise bring about and this deal with them concretely.

## 5 The Singularity

Engineer and futurist Ray Kurzweil has claimed that around 2029 AI will pass the Turing Test, achieving human levels of intelligence. Then, around 2045, AI will have surpassed these levels on its own, thus multiplying human intelligence enormously in a cyborg fashion. That moment is called the (*technological*) *Singularity*, as mentioned above. The term was introduced broadly by science fiction writer Vernor Vinge in his 1980 story, *True Names*. He followed this up with a 1993 article, “The coming technological Singularity,” in which he maintained that his fictional vision would become a reality in the first part of the twenty-first century. It is from that article that the term technological Singularity took hold and spread among AI scientists.

The notion can actually be traced back to a comment made by mathematician John van Neumann, cited by Stanislas Ulam (1958: 5): “[The] ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue.” The research in this domain of AI is now called *Seed AI*; it seeks to create machines that are capable of improving their own software and hardware, by recursively rewriting their own source code without human intervention. The scenario for this occurrence was put forth hypothetically by I. J. Good in 1965, who claimed that, as computers increase in computational power, AI researchers can build a machine that surpasses their own intelligence. This machine then will design even more intelligent machines ad infinitum. This generative process was called a “law of accelerating returns” by Kurzweil, whereby AI would increase exponentially.

The Singularity implies that at a certain point AI will become conscious of what it is doing, or awareness by the artificial mind of itself (Bor 2012). But the use of the word *consciousness* is itself ambiguous. So, to avoid this ambiguity, Kurzweil used the designation “mind-beyond-machine,” which alludes to the

original Cartesian view that the mind is a machine that is activated by some animal spirit. The difference in Singularity Theory is that the animal spirit is replaced by algorithms. There are two obvious problems with this idea – one is that we still do not know what the mind really is in human terms and second there is no way to program the imagination in the form of algorithms. Moreover, at the emotional level of mind the brain produces “affective” models of reality, which are connected to the functions of the body, not to any algorithmic rules.

There have been many critiques of Singularity Theory, which need not concern us here. However, there are psychological tests that can be used to support or refute it theoretically. Kurzweil’s test for the advent of an artificial super-intelligence, before the Singularity, is the Turing Test. As is well known, this is an argument devised by mathematician Alan Turing to show that one could program a computer in such a way that it would be virtually impossible to discriminate between its answers and those contrived by a human being. Suppose someone is in a room that hides on one side a programmed computer and, on the other, a human being. The computer and the human being can only respond to the person’s questions by writing on pieces of paper which both pass on to the observer through slits in the wall. If the observer cannot identify, on the basis of the written responses, who is the computer and who the human being, then the observer must logically conclude that the machine is “intelligent.” It has passed the Turing Test.

An early counter-argument to the Turing Test was formulated by American philosopher, John Searle (1984). Searle argued that a computer does not know what it is doing when it processes symbols, because it lacks intentionality. Just like an English-speaking human being who translates Chinese symbols in the form of little pieces of paper by using a set of rules for matching them with other symbols, or little pieces of paper, knows nothing about the “story” contained in the Chinese pieces of paper, so too a computer does not have access to the “story.” It does not have the inhering ability of human symbols.

Actually, in addition to the Turing Test, AI must pass another test, which can be called “Gödel’s Test,” after Kurt Gödel’s (1931) famous proof that within any formal set of rules there are results that can be neither proved nor disproved. Turing (1936) himself gave a version of this test, which he called the halting problem. Given a computer program and an input, the problem is to determine whether the program will finish running or will go into a loop and run forever. Turing proved that no algorithm for solving this problem can exist. He reasoned as follows: If a solution to a new problem were to be found, then it could be used to decide an undecidable problem by changing instances of the undecidable problem into instances of the new problem. Since we know that no

method can decide the old problem, no method can decide the new problem either.

Another test that AI must pass is known, simply, as “Affective Test,” whereby AI shows the ability to grasp emotions. Actually, there is now a sub-field aiming to study and develop computer systems and devices that can recognize and simulate human emotions. So-called Affective Computing (AC) software has been developed with the capacity to detect emotional states in people on the basis of the algorithmic analysis of facial expressions, muscle tension, postures, gestures, speech tones, pupil dilation, etc. The relevant technology includes sensors, cameras, big data, deep learning software, etc. The aim is to construct machines that can decode emotional states or influence them. This line of research has led to the building of so-called Empathy Machines, which are companion robots that display the ability to respond to human emotional states or language. To actually achieve empathy, however, a robot would have to be able to experience emotion and that means being able to recognize and comprehend it. It is also difficult to predict how an Empathy Machine will process, reproduce, or simulate emotions. Incidents have been documented with Alexa (a commercially-available machine) that bring this out. Apparently, the machine started producing laughter for no reason. The technical diagnosis was that Alexa somehow processed the command, “Alexa laugh,” when the user had not, in actual fact, uttered the statement.

The last comment leads to a subtest of the Affective Test, which may be called the “Humor Test,” or the ability of a computer to grasp and use humor for intelligence purposes. As Wells (1988: 7) aptly points out, humor is dependent “on a skeleton of double meanings, surprise, and the familiar in strange dress.” A subtest of this test can be called the “Irony Test,” or the ability of the machine to understand and produce ironic modes of language. Irony has various social and cognitive functions, falling into three main categories – verbal, dramatic, and situational. The first one involves strengthening the intended meaning of an utterance by inducing the interlocutor to seek it indirectly in a subjective way. The second implies relaying some meaning to which some interlocutor has no access. In the Greek tragedy *Oedipus Rex* by Sophocles, for instance, Oedipus kills a man. He does not know that the man is Laius, his father. Oedipus puts a curse on the slayer of Laius. The irony is that Oedipus has unknowingly cursed himself, and only the audience has access to the irony. Situational irony involves emphasizing events that work out contrary to expectations. Suppose that a town is preparing a celebration for a returning soldier. But the soldier is killed in an accident on his way home. The irony comes from the contrast between the expectations of the people and the actual situation.

AI might be able to pass the Turing Test, but it is unlikely ever to pass the Gödel and Affective Tests (and subtests). The former implies that logical systems are undecidable; and thus it is not clear how AI can eliminate undecidability from its algorithmic systems. The latter means that AI would have to a limbic system that allows it to understand the affective components of information. AI might also not be able to pass a version of the Turing Test as well, which maintains that the essence of human understanding is creative inference – a process called *abduction* by a Peirce (1931–1958, volume 5: 180).

As far as can be told, it is unlikely that AI will ever pass this Abduction Test, since inferences are unpredictable and are based on bodily experiences. Despite such huge obstacles, AI has continued to pursue the goal of replicating and optimistically passing human intelligence. This specific and focused approach now comes under the rubric of Artificial General Intelligence (AGI), which is the study of intelligence irrespective of its carrier – human, animal, or machine. Sirius and Cornell (2015: 14) define this field as follows:

AGI describes research that aims to create machines capable of general intelligent action. The term was introduced in 2003 in order to avoid the perception that the field was about creating human-level or human-like intelligences, which is covered by the term “Strong AI.” AGI allows for the inclusion of nonhuman as well as human models of general intelligence.

An important experiment by Weisberg, Keil, Goodstein, Rawson, and Gray (2008), however, seems to show that humans do not process information in the same way as the algorithms of computer scientists do. The researchers tested people’s abilities to critically consider the underlying logic of a computational explanation, giving naïve adults (those with no knowledge of neuroscience), students in a neuroscience course, and neuroscience experts brief descriptions of psychological phenomena followed by one of four types of explanation. The actual information was irrelevant to the logic of the explanation, as confirmed by the expert subjects. The subjects evaluated good explanations as more satisfying than bad ones. But those in the two non-expert groups additionally judged explanations with logically irrelevant information as more satisfying than those without. The neuroscience information, in other words, had a particularly striking effect on judgments of bad explanations, masking otherwise salient problems in these explanations. Although the experts were not fooled by the explanation, the experiment did issue a warning about the nature of explanations and their purported realism, referring to them as having, because of their neuroscientific explanation, a seductive and alluring quality.

## 6 Enter Baudrillard, Postman, and semiotics

The late French semiotician-philosopher Jean Baudrillard (1983) traced the belief that AI would become equivalent, and even surpass, natural intelligence to a breakdown between people's perception of reality and fantasy. As is well known, both within semiotics and in cognate fields, Baudrillard called people's engagement with fantasy, *hyperreality*, which produces a sense that the world of artificial simulations is more real than real. He called this the *simulacrum*, which is an artificial form of consciousness that emerges spontaneously. The term *simulacrum* comes from Latin, meaning "likeness" or "similarity," and was used in the nineteenth century by painters to describe drawings that were seen merely to be copies of other paintings rather than emulations of them. Aware of this designation of the term, Baudrillard insisted that a computational simulacrum is not the result of a simple copying or imitation, but a form of perception, which he called hyperreal. The subtext in Baudrillard's notion is that AI is merely a simulacrum, and that we are inclined to accept all kinds of simulacra as real, rather than hyperreal, which is only possible in a world characterizable as a *technopoly*.

The term *technopoly* was coined by Neil Postman in 1992, in *Technopoly: The Surrender of Culture to Technology*, defined as a society that has become so totally reliant on technology that it seeks authorization in it, as well as deriving recreation from it, and even taking its orders from it. This is a coping strategy that results when technology saturates the world. Technopoly, Postman suggests, is a "totalitarian technocracy," evolving on its own. It reduces humans to seeking meaning in machines and in computation. Postman saw negative consequences for the human condition in a technopoly, since the promises of advanced technologies would turn society into an amorphous mass of non-thinkers. He altered McLuhan's phrase of "the medium is the message" to "the medium is the metaphor," insisting that new media are mind-numbing tools. Postman was particularly concerned with children's upbringing in a technopoly. While children were once seen as little adults, the Enlightenment brought broader knowledge of childhood, leading gradually to the perception of childhood as an important period of development. Since children now have easy access to information, the result is a diminishment of their potential. He thus warned that those who do not see the downside of technology, constantly demanding more innovation are silent witnesses to a new cognitive form of brain pollution.

In addition to this, Merlin Donald (2014) considers the eroding effect of externalized memory on the cognitive faculties of the species. He warns against

the exploitation of memory suggesting that what lead the human mind to produce such extensive cultural systems is what is inherently put at risk by the very products it created.

A basic operating principle of AI is that of recursion (applications of rules over and over) is the underlying structure of knowledge tasks. In AI, recursion refers more technically to the process of repeating items in a self-similar way and, more precisely, to a method of defining functions in which the function being defined is applied within its own definition, but in such a way that no loop or infinite chain can occur. The so-called “recursion theorem” says that machines can be programmed to guarantee that recursively defined functions exist. Essentially, it asserts that machines can encode enough information to be able to reproduce their own programs or descriptions (Berlinski 2000).

Now, the question of characterizing the brain as a recursive organ is at the core of the semiotic critique of AI. The human mind produces “habits” not “recursions.” The concept of habit enlisted here is not the behavioristic one – it is the one prefigured in the philosophical pragmatism of Alexander Bain (1868) and shaped by Charles Peirce’s well-known antipathy to psychologism (the over-reliance on psychology as an exclusive approach to human mentation). To Bain’s notion of habit, Peirce added that “the essence of belief is the establishment of a habit; and different beliefs are distinguished by different modes of action to which they give rise” (Peirce 1931-1958, volume 5: 398). Peirce also saw a connection between evolutionary change and “habituescence,” as he called the conscious awareness of habit-formation (Peirce 1931-1958, volume 6: 302–303):

As habits become imprinted in the brain’s neural pathways, they become virtually impossible to break, increasing the automatic responses that we have to risky behaviors. However, because they constitute a semiotic system – habituescence – they can be changed and revised at will. Moreover, these responses are not solely the product of the brain, but rather, the occurrence of a dialectical relationship of mind and body.

## 7 Concluding remarks

Perhaps the most staggering “thought” of all is that we have the capacity to “think” sentiently in the first place. It is thus little wonder that AI research has recently started to focus on consciousness and what it means. Consciousness, or life aware of itself, is indeed a difficult phenomenon to explain. From a semiotic perspective, only one thing would be given as certain: namely that

consciousness would have been impossible without the imagination and the semiotic blends that it engenders. In embodied cognitive science, *blending* is defined as the ability of the brain to take concepts in one domain and blend them with those in another to produce new ones or to simply understand existing ones. Changing the blends leads to changes in understanding and cognition. Blending theory thus makes it possible to connect, say, language and mathematics, in a way that goes beyond simple analogies (Lakoff and Núñez 2000, Fauconnier and Turner 2002, Walsh Matthews 2018).

Blending is unconscious and that is why we hardly ever are aware of what we are doing when we think. It is analogous to Peirce's abductive hunches, which are attempts to understand what something unknown means initially. These eventually lead to inferences through a matrix of associative (blending) devices to previous knowledge such as analogy and metaphor. So, the Pythagorean triangle, which came initially from the hunches of builders, led to an inference that all similar triangles may contain the same pattern, and this led to the insight that we call the Pythagorean theorem, which was given a logical form through proof. Once the form exists, however, it becomes the source for more inferences and abductions, such as the previously-hidden concept of Pythagorean number triples. It is clear that cognition cannot be separated from blends that result from visualization, intuition, imagination, reasoning, and all the other aspects of human mentation that can hardly be algorithmicized.

Clearly, semiotics must respond to the AI challenge. It must focus on how abductive responses to the world generate meaning in the human sense, not recursive processes. The AI approach to the human mind is, nonetheless, a very instructive one. But it is just that – an artificial approach that semioticians can help enlarge with their own approaches to modeling (for example, Sebeok and Danesi 2000). Semiotics does not attempt to answer the all-encompassing question of mentality, a does AI or Singularity Theory, because it knows that an answer is unlikely. Rather, it limits itself to a less grandiose scheme – describing the representational activities that semiosis animates. It is in these activities that consciousness can be considered in an indirect fashion. The semiotic agenda can both enrich AI and provide it with caveats about its more radical claims, since it sees mentality as being shaped by a search for the biological, psychic, and social roots of the human need for *meaning*.

## References

- Ashby, William R. 1956. *An introduction to cybernetics*. London: Chapman and Hall.
- Bain Alexander. 1868. *The senses and the intellect*. London: Longmans.
- Baudrillard, Jean. 1983. *Simulations*. New York: Semiotexte.
- Berlinski, David. 2000. *The advent of the algorithm*. New York: Harcourt.
- Black, Max. 1962. *Models and metaphors*. Ithaca: Cornell University Press.
- Bor, Daniel. 2012. *The ravenous Brain: How the new science of consciousness explains our insatiable search for meaning*. New York: Basic Books.
- Brier, Søren. 2007. *Cybersemiotics: Why information is not enough*. Toronto: University of Toronto Press.
- Damasio, Antonio R. 1994. *Descartes' error: Emotion, reason, and the human brain*. New York: G. P. Putnam's.
- Descartes, René. 1633. *De homine*. Amsterdam: Elsevier.
- Donald, Merlin. 2014. The digital era: Challenges for the modern mind. *Cadmus* 2(2). 68–79.
- Engelbart, Douglas C. 1962. *Augmenting human intellect: A conceptual framework*. SRI Project No. 3578, Stanford Research Institute.
- Fauconnier, Gilles & Mark Turner. 2002. *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic.
- Gardner, Howard. 1985. *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.
- Gödel, Kurt. 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, Teil I. *Monatshefte für Mathematik und Physik* 38. 173–189.
- Good, Irving J. 1965. Speculations concerning the first ultraintelligent machine. *Advances in Computers* 6. 31–88.
- Haraway, Donna. 1989. *Primate visions: Gender, race, and nature in the world of modern science*. London: Routledge.
- Haraway, Donna. 1991. *Simians, cyborgs, and women: The reinvention of nature*. London: Free Association Books.
- Hawkins, Jeff & Sandra S. Blakeslee. 2004. *On intelligence*. New York: Times Books.
- Husserl, Edmund. 1890. *Philosophie der Arithmetik*. The Hague: Nijhoff.
- Konner, Melvin. 1991. Human nature and culture: Biology and the residue of uniqueness. In James J. Sheehan & Morton Sosna (eds.), *The boundaries of humanity*, 103–124. Berkeley: University of California Press.
- Kurzweil, Ray. 2005. *The singularity is near*. Harmondsworth: Penguin.
- Kurzweil, Ray. 2012. *How to create a mind: The secret of human thought revealed*. New York: Viking.
- Lakoff, George & Rafael Núñez. 2000. *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
- Licklider, Joseph C. R. 1960. Man-computer symbiosis. *IREE: Transactions on Human Factors in Electronics, HFE-1*. 4–11
- MacWhinney, BRIAN 2000. Connectionism and language learning. In Michael Barlow & Suzanne Kemmer (eds.), *Usage models of language*, 121–150. Stanford: Center for the Study of Language and Information.
- Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.



- McLuhan, Marshall. 1964. *Understanding media: The extensions of man*. Cambridge: MIT Press.
- McLuhan, Marshall & Eric McLuhan. 1988. *Laws of media: The new science*. Toronto: University of Toronto Press.
- Merleau-Ponty, Maurice. 1942. *La structure du comportement*. Paris: Presses Universitaires de France.
- Merleau-Ponty, Maurice. 1945. *Phénoménologie de la perception*. Paris: Gallimard.
- Neisser, Ulrich. 1967. *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Newell, Allen. 1991. Metaphors for mind, theories of mind: Should the humanities mind? In James J. Sheehan & Morton Sosna (eds.), *The boundaries of humanity*, 158–197. Berkeley: University of California Press.
- Peirce, Charles S. 1931–1958. *Collected papers*. Cambridge, Mass.: Harvard University Press.
- Penrose, Roger. 1989. *The emperor's new mind*. Cambridge: Cambridge University Press.
- Postman, Neil. 1992. *Technopoly: The surrender of culture to technology*. New York: Alfred A. Knopf.
- Searle, John. 1984. *Minds, brain, and science*. Cambridge, Mass.: Harvard University Press.
- Sebeok, Thomas A. & Marcel Danesi. 2000. *The forms of meaning: Modeling systems theory and semiotics*. Berlin: Mouton de Gruyter.
- Sirius, R. U. & Jay Cornell. 2015. *Transcendence: The disinformation encyclopedia of transhumanism and the singularity*. San Francisco: Disinformation Books.
- Turing, Alan. 1936. On computable numbers with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society* 41. 230–265.
- Uexküll, Jakob von. 1909. *Umwelt und Innenwelt der Tiere*. Berlin: Springer.
- Ulam, Stanislaw. 1958. Tribute to John von Neumann. *Bulletin of the American Mathematical Society* 64. 5.
- Vinge, Vernor. 1993. The coming technological Singularity: How to survive in the post-human era. *Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace*, NASA Publication CP-10129, 11–22.
- Walsh Matthews, Stéphanie. 2016. How fit is the semiotic animal? *The American Journal of Semiotics* 32(1). 205–217.
- Weisberg, Deena S., Frank C. Keil, Joshua Goodstein, Elizabeth Rawson & Jeremy R. Gray. 2008. The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience* 20. 470–477.
- Wells, David. 1988. *Hidden connections, double meanings: A mathematical exploration*. Cambridge: Cambridge University Press.
- Wiener, Norbert. 1948. *Cybernetics, or control and communication in the animal and the machine*. Cambridge, Mass.: MIT Press.
- Wiener, Norbert. 1950. *The human use of human beings: Cybernetics and society*. Boston: Houghton Mifflin.

## Bionotes

### Marcel Danesi

Marcel Danesi (b. 1946) is Full Professor of Linguistic Anthropology and Semiotics at the University of Toronto. His research interests span areas from semiotic theory and pop culture analysis to metaphorical analysis and mathematical representation. Recent publications include: *Marshall McLuhan: The unwitting semiotician* (2018), *Ahmes' legacy: Puzzles and the mathematical mind* (2018), *An anthropology of puzzles: The role of puzzles in the origins and evolution of mind and culture* (2018), and *Memes and the future of pop culture* (2019).

### Stéphanie Walsh Matthews

Stéphanie Walsh Matthews (b. 1977) is an Associate Professor of Languages, Literatures, and Cultures at Ryerson University, in Toronto. Her research interests include cognitive semiotics, Autism Spectrum Disorder, language and cognition, and postcolonial theories. Recent publications include: *A.J. Greimas: Life and semiotics* (2017), *Semiotics post-Greimas* (2017), "Semiotics and literary criticism," *Oxford Encyclopedia* (2017), "How fit is the semiotic animal?" (2016).