AI & Soc
DOI 10.1007/s00146-012-0404-4

25TH ANNIVERSARY VOLUME
A FAUSTIAN EXCHANGE: WHAT IS TO BE HUMAN IN THE ERA OF UBIQUITOUS TECHNOLOGY?

Art and robotics: sixty years of situated machines

Simon Penny

Received: 17 September 2011 / Accepted: 6 January 2012

Abstract This paper pursues the intertwined tracks of robotics and art since the mid 20th century, taking a loose chronological approach that considers both the devices themselves and their discursive contexts. Relevant research has occurred in a variety of cultural locations, often outside of or prior to formalized robotics contexts. Research was even conducted under the aegis of art or cultural practices where robotics has been pursued for other than instrumental purposes. In hindsight, some of that work seems remarkably prescient of contemporary trends. The context of *cultural robotics* is a highly charged interdisciplinary test environment in which the theory and pragmatics of technical research confronts the phenomenological realities of physical and social being in the world, and the performative and processual practices of the arts. In this context, issues of embodiment, material instantiation, structural coupling, and machine sensing have provoked the reconsideration of notions of (machine) intelligence and cognitivist paradigms. The paradoxical condition of robotics vis-à-vis artificial intelligence is reflected upon. This paper discusses the possibility of a new embodied ontology of robotics that draws upon both cybernetics and post-cognitive approaches.

Keywords Robotics · Robotic art · Artificial intelligence · Cybernetics · Cognitivism · Situated cognition · Media art

S. Penny (✉)
Claire Trevor School of the Arts, University of California, Irvine, CA, USA
e-mail: penny@uci.edu
URL: http://simonpenny.net

1 Overview and historical context

The purpose of this paper is threefold. First, to document and elucidate robotic art—a minor but stubbornly persistent strain in the history of computer automated cultural artifacts (CACA)—a strain that belatedly must be recognized for its prescient relevance. The second is to discuss some of the history of ideas related to robotics. And third, to explore some of the deeper tensions in the cognitivist worldview as they apply to robotics, to art, and to their combination.

Rather than synthesizing the skills of the programmer and the painter, robotic art integrates the sensibilities of the sculptor, installation artist, and performer with the computational systems equipped with both sensors and mechanical effectors. As such, these systems mirror the ongoing sensorimotor engagement of humans (and animals) in the world. This fundamentally interactive conception traces its roots to the pre-digital and pre-cognitivist cybernetic sensibilities of the mid 20th century, a movement that had substantial influence in the arts as well as in engineering and other sciences. Indeed, one might reasonably assert that the seemingly wildly disparate fields of control theory and performance art both have their roots in the ur-discipline of cybernetics. The argument for control theory is straightforward and uncontestable. The influence of cybernetics in the arts is less known and less celebrated but just as clear.

This confluence of arts sensibilities and technological development produced artifacts that exhibit and explore ongoing computational engagement with the real/material/social world, in a manner analogous to humans and animals. This quasi-biological condition was a key aspect of 1990s Artificial Life research. I argue here (as I have argued in the past) that this field of practice prophetically researched and prototyped systems, relationships, and devices, which more recently became institutional and corporate research topics, garnering descriptors such as *autonomous agents*, *affective computing*, *tangible interfaces*, *haptics*, *ubiquitous computing,* and more generally *human computer interaction* (Penny 2000, 2008, 2010). *Helpless Robot* (1987) by Norman White is a paradigmatic example. It has no motive power. It depends on its verbal persuasiveness to entice humans to do its work. Wittily, as a person is drawn into helping the helpless robot, it becomes increasingly impolite and finally abusive, creating a situation from which the humiliated human helper must sheepishly escape.

From this perspective, such materially instantiated and physically situated practice stands in stark contrast to the better-known media arts practices, which focus on software-based quasi-cinematic practices (Penny 2011b). Quite simply, such practices are preoccupied with representation in the naïve sense of "picturing of something." Contrarily, the kinds of practices I will glibly group as *robotic art* are general or predominantly "the thing itself" (though such manifestations can surely be, and often are, both the thing itself and a representation). It cannot be stressed enough that this is a fundamental ontological split between kinds of practices.

This general taking-for-granted of representationalism is a symptom of the powerful confluence of two cultural streams, which themselves have a deeper common origin. On the one hand, the tradition of western pictorial representation inheres key aspects of a Cartesian Humanism—particularly the axiomatic separation of subject and object implied and dictated by the rules of perspective. But, artisanal and cultural practices have always occupied a marginal condition with respect to the (untenable) Cartesian construction because they maintain a tight and ongoing intercourse with materiality where reasoning takes a back seat. One rides a horse, rather literally, by the seat of one's pants. The act of chiseling cannot exist without the chisel. And likewise, a chisel is not fully a chisel unless it is wielded with appropriate chiseling skills, situated in a chiseling context, with a whole host of chiseling support items—benches, mallets, vises, and wood to chisel.

On the other hand, the ideology that underlies the technology of digital computation trades explicitly in representation, at the level of the code, and in a general commitment to cognitivism—that is, the idea that intelligence or thinking or reasoning can be adequately and more or less completely described as operations on symbolic tokens in some abstract quasi-mathematical reasoning space, and is thus quintessentially representationalist. As will be evident then from this brief exegesis, the common roots of these two cultural vectors are in a modernist Cartesian humanism. Thus, artistic sensibilities bring to robotics a deep and subtle understanding of embodiment, gesture, materiality, physical and social space, while AI pulled the other way into a platonic realm of immaterial ideals.

This is no idle philosophical folly. It is explanatory of the trajectory of history of computational cultural practices over the last six decades. In very broad strokes, the history I pursue below is as follows: Machine Art emerged largely as a post-war phenomenon, with its most famous practitioners being Jean Tinguely, Alexander Calder, Vassilakis Takis, and Nicolas Schöffer. Between the wars, Marcel Duchamp made some gestures toward a mechanical art practice, but abandoned actual working machines (rotoreliefs) for representations of such (*The Large Glass,* 1915–1923).

In the meantime, Cybernetics was in ascendency as a philosophy of the *intelligent* machine (among other things), though what the cyberneticians meant by *intelligence* was a rather different thing from what the artificial intelligence community understood by the term. In this respect, cybernetics was a philosophy of the situated machine and was preoccupied with ongoing sensorimotor engagement with the world, epitomized by the key cybernetic concept of feedback. The notion of autopoiesis proposes a definition of biological life in terms generally compatible with cybernetics.[1] In an historically elegant way, it was Humberto Maturana's main protégé Francisco Varela who applied autopoiesis to cognitive science and, thereby, played a major role in the development of a post-cognitivist position in cognitive science (Maturana and Varela 1980). But I am getting ahead of myself.

During the 1960s, the technology of digital computing—the automation of mathematico-logical reasoning—developed rapidly, due in large part to the invention and production of the transistor, which reduced physical size and power consumption markedly while increasing reliability.[2] Each *logic gate* in a digital computer is composed of transistors, and these gates are deployed to implement a particular flavor of binary logical reasoning as developed by George Boole.[3] This development segued neatly with the rise of cognitivism, catalyzed by the work of Alan Newell and Herbert Simon, particularly their Physical Symbol System Hypothesis, about which they famously said, "A physical symbol system has the necessary and sufficient means for general intelligent action" (1976). It is this general assertion that intelligence consists in the

[1] In Lettvin et al. (1959), a key but almost forgotten early work in neuroethology and the biology of cognition, they offer clear evidence of the fallaciousness of cognitivism long before it became dogma!.

[2] While records of transistor research go back as far as 1921 (Lilienfeld), Bardeen, Brattain, and Shockley are credited with the invention of the bipolar transistor in 1947, but production (by Texas Instruments) did not occur until the 1960s. See http://en.wikipedia.org/wiki/Transistor.

[3] http://en.wikipedia.org/wiki/Boolean_algebra.

manipulation of symbols that set AI on its quintessentially representationalist path, a path that began to be rather unclear by the mid 1980s. It was at this point that some cognitive scientists, roboticists, and others began to cast about for alternative possibilities. These alternatives emerged largely in fields that became known as Artificial Life research and in Embodied and Situated Cognition.

The term *embodiment* has been deployed in various contexts with various meanings. I use the word here consistent with these various applications, so it is necessary at the outset provide a clarification. With respect to humans and cognition, the use of the term first connotes the simple fact of bodily materiality, but further usually implies a recognition of physical instantiation as being a necessary aspect of and for human cognition, a position that refutes "brain in a vat" arguments (Clark 2008) and accepts the reality of morphological computation (MacIver 2001). When applied to robotics, the term has an analogous meaning of "physical instantiation." But it also implies that a machine situated in a physical environment must both manage the qualities of that environment, and also can exploit qualities of its own embodiment and the materiality, as asserted by Brooks (1999): "the world is its own best model". A related underlying theme here is a recognition of an unreflexive representationalism in both computer science and the arts. This recognition leads to a speculative research agenda concerned with the potential of post-representationalist, post-cognitivist techno-arts practices (Penny 2011a).

2 Techno-mythology

In the popular imagination, the conception of *robotics* and *artificial intelligence* is informed largely by sci-fi fantasies disseminated via the dream machines of popular media such as Hollywood film. As a starting point, it is necessary to maintain a double sense of these terms. On the one hand as potent images of cultural dread and longing, and on the other hand as pragmatic exercises of engineering with capabilities drastically short of the sci-fi fantasy version. In this sense, it is useful to regard *automated reasoning* as a less grandiose synonym for AI, and *augmented machine tool* as a similarly more realistic synonym for *robot*.

It is important to draw a rough line around the kinds of artifacts and system that in the current context will be included in the larger *robotics* arena. Firstly, the sci-fi specter of the anthropoid robot will play a relatively minor part here. But, the more general area of biomimetic robotics will be important. The capability of mobility is certainly usual, but not required. Some capacity for electromechanical motion is normal, but again not compulsory. For the purposes of this discussion then, a robot is a technological system, which by the integration of sensors and effectors, evidences contingent autonomous behavior in the physical world. This definition excludes prosthetic and teleoperated technologies.

It is also necessary to disabuse readers of the notion that *robotics* is a research field ensconced within digital computation. This is true neither technically nor historically. *Computation*, even in the Boolean sense, requires neither digital logic, nor transistors, nor even electricity—robotics research certainly predates the *digital*. Neither has robotics research occurred exclusively in robotics labs. As David Mindell has demonstrated, the science of the self-governing machine not only predates AI but also predates Cybernetics and extends back to the 19th century, instantiated in forms neither digital nor electronic, nor even electromechanical (Mindell 2002).

3 Post-war art and cybernetics

The work of the British cyberneticians W. Ross Ashby, Stafford Beer, Gordon Pask, and W. Grey Walter has been the subject of renewed interest in recent years (Pickering 2010). It is yet another odd twist in this story; two of the early cases of what we could now call robotics art were built by theses folk. Grey Walter, a neurologist and cyberneticist, built his *Turtles*, Elmer and Elsie, in the late 1940 (Fig. 1).

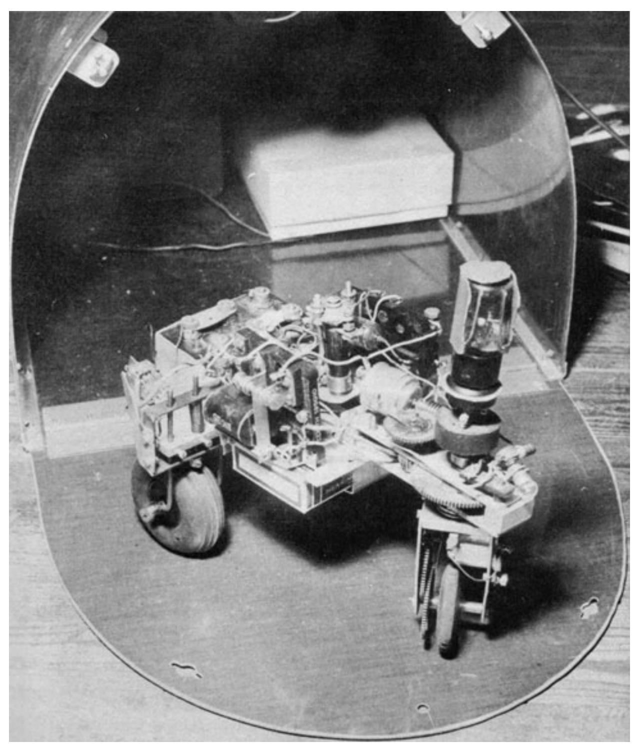

Fig. 1 Elsie, with cover removed

In the current context, the following aspects of Walter's work are important. Firstly, as a neurologist, he was concerned with neurological modeling, and in this case a simple brain of two neurons. Importantly, evidence of the success of the enterprise was not to demonstrate problem-solving ability in vacuo, or in some closed electrico-symbolic realm, but to demonstrate their ability to "survive" in their world. Here, consistent with a cybernetic world-view, "survival" means to materially instantiate a model of effective integration with its world. With respect to robotics generally, the term "survive" is deployed metaphorically. Robots can be, and in this case were, understood as models for human or biological systems, as opposed to being senor-augmented machine tools. The *minds* of Elmer and Elsie were not constituted around reasoning, but around *drives*. Those familiar with the agendas of (reactive) robotics and autonomous agent research will recognize that Walter's work predates similarly motivated work by 40 years. Indeed, 30 years later, that icon of robotics theory, *Vehicles* (1986) by Valentino Braitenberg, includes textual descriptions of possible machines that, seemingly innocently, duplicate the behavior of Walter's *Turtles* (1986).

A more explicitly, artistic project of British Cybernetics was Gordon Pask's *Musicolour*, a performative analog improvisatory multimedia agent that was designed to collaborate with a performing musician.[1] While less strictly *robotic* Pask's *Musicolour* evidences a conception of an actively creative agent, which as with Walter's *Turtles*, preempts such conceptions in academic and institutional contexts by a generation. The question of why Pask's and Walter's work, and other work like it, was ignored by robotics researchers of later decades is, in my opinion, a rather sobering example of the dangers of discursive hegemony—a topic that I will return to below.

The Cybernetics movement increasingly took the form of an ur-discipline and was hugely influential across diverse disciplines, not just engineering, but in meteorology, neurophysiology, sociology, business management, and the arts. Among others, Nicholas Schoffer implemented Cybernetically inspired robotic artworks during the 1960's, and work in interactive installations and architectural environments was pursued by characters as diverse as the Chilean artist Juan Downey and Greek composer Mikis Xenakis.

Between 1965 and 1970, several large-scale, cybernetically inspired interdisciplinary exhibitions or projects occurred. The first of these, *9 Evenings: Theatre and Engineering* (1966), was part of a series of interactions between artists and engineers headed by Bell Labs engineer Billy Kluver. Eventually named Experiments in Art and Technology (EAT), the organization also involved Robert Rauschenberg, Jasper Johns, David Tudor, John Cage, to name just a few prominent members. Many of their projects pushed engineering capabilities to their limits and capture, as proofs in principle, and as modalities of technological interaction and communication, which seemed outlandish at the time. A later project of EAT, called *Some More Beginnings* (1968–1969), included pioneer Canadian robotic artist Norman White.

Significantly, Kluver was personally close to members of the US Cybernetics community. This group of interdisciplinary luminaries who gathered around Norbert Wiener included Claude Shannon, Warren McCulloch, Walter Pitts, Heinz von Foerster, Gregory Bateson, and Margaret Mead. Shannon, clearly a man of some wit as well as artisanal skill, had a penchant for building machines, one of which he called "The Ultimate Machine": a box whose only function when switched on was to project a machine finger out of a hatch and flick its own main switch off.

In addition, three other curatorial projects of the period should be mentioned. *Cybernetic Serendipity*, curated by Jasia Reichardt at the ICA (London 1968), involved several leading theorists and included works by Gordon Pask and Edward Ihnatowicz. Art theorist Jack Burnham was a major vector in the popularization of Cybernetic ideas in the arts (1968). His exhibition, *Software* at the Jewish Museum in New York (1970), included such diverse characters as Hans Haacke and Nicholas Negroponte. *Software* is significant for the way it evidences a transition from Cybernetic to Symbolic AI thinking, and in the way, it evidences how hazily non-technological the notion of "software" was at the time. The LACMA Art and Technology program of Maurice Tuchman (also 1966) paired artists with (the R&D operations of) major corporations. This project became mired in Vietnam War protests that polarized the West Coast art community.

One of the most extraordinary expressions of post-war cybernetic sculpture was *The Senster* (1970) by the aforementioned Edward Ihnatowicz (Fig. 2).

The Senster was a 3-m high, biologically inspired, articulated robotic arm that tracked and leant toward human interactors. *The Senster* deployed an extraordinary array of handcrafted technologies in the process, including two axis acoustic localization, doppler radar, and custom hydraulic articulation. In its structure and behavior, *The Senster* captured agendas that would not be recognized in institutional circles for another 25 years.

Most of the works discussed above languished in an historical twilight for a generation. A simplistic explanation might suggest that they were simply ahead of their time and occurred without an institutional or discursive context. This is only partly true as there was clearly an active discursive context. A second approach would be to propose that the disciplinary location in which these works arose had no historical connection to the disciplinary

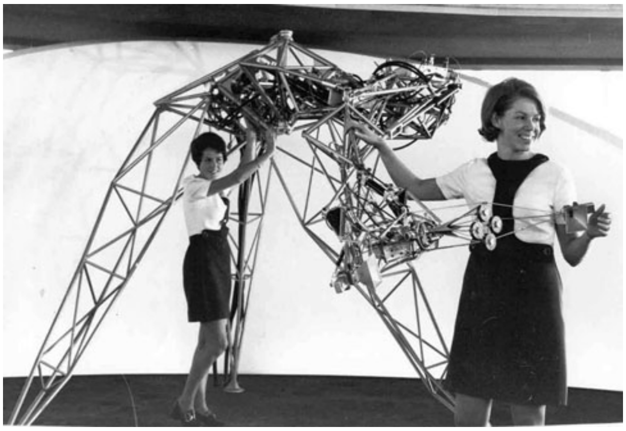

Fig. 2 Senster, installed at the Philips Evoluon, Eindhoven, 1970–1974

locations in which similar, more recent work would later arise. This explanation has some substance. A third dynamic, which played, I would argue, the most potent role, was a major paradigm shift.

4 Cybernetics, AI, and artificial life

The first generation of robotic experimentation and robotic art occurred during the middle 20th century in which Cybernetics was highly influential. Cybernetics was preoccupied with the interaction of an agent or artifact with its environment. Such agents were understood to sense and act in feedback loops to stabilize their relation to the world, seeking homeostasis. To correlate cybernetics with analog computing is superficial, but historically accurate. In the later 1960s, a new school of thought, in key ways diametrically opposed to cybernetics, emerged: symbolic AI. It rode on a different technological wave: transistor-based logic gates.

There is a profound historical rupture between cybernetics and discourses of digital computation, such as AI. We might say that the discourse of Cybernetics and the discourse of AI are opposed and complementary in the following way. Cybernetics was involved with the idea of intelligence as located in and evidenced by the integration of an agent with its world, whereas symbolic AI was concerned with the notion of intelligence as constituted by manipulation of symbolic tokens occurring in an abstract immaterial space of logico-mathematical representations, be that space in a brain or in a machine. And it is significant that these were taken to be, in principle, interchangeable.

One might say for cybernetics, intelligence existed at the interface with the world. This view is in stark opposition to the cognitivist view, which makes sensors and effectors secondary to the main event of information processing—symbol manipulation. The conception of the brain as a thinking organ at least theoretically separable from the body is a hallmark of computationalist cognitivism. Indeed, there is no more resounding reification of the Cartesian dualism than the hardware/software duality. It speaks volumes that in computer science parlance, sensors and effectors are referred to as *peripherals*. Over 30 years, roughly 1960–1990, the AI/cognitivist paradigm was dominant, but problems of integration with the real lived world, previously dismissed as future technical challenges, began to appear to be deep-seated (Dreyfus 1972, 1979).

To the extent that they are sensing and acting systems, robots constitute a kind of quasi-biology. For an artifact to behave in the world implies autonomy, and thus quasi-biology, for what has autonomy except organisms? A concern with embodied and autonomous agents was characteristic of the Artificial Life movement of the 1990s. Hence, the focus on reactive and bottom-up robotics—a robotics that found fault with the computationalist world-view and turned for inspiration to biology as a way of proceeding that which had been unfashionable for a generation. A major figure in that movement was an expatriate Australian, Rodney Brooks. The spirit of Rodney Brooks's aphorism that, "the world is its own best model," is at the very root of new ways of thinking about robotics and cognition (1990). The rise of Artificial Life in the 1990s was then both a reaction to the perceived failures of the *physical symbol system* conception of AI, and a renewed interest in biological models, which hearkened back to cybernetic ideas. In a neat dialectical synthesis, some Alife work might be conceived as cybernetics implemented in digital systems. *The notion that intelligent action necessarily and centrally involves the logical manipulation of symbolic representations is held far less surely than it once was.* In a related and abrupt turn-about, the new generation of post-cognitive cognitive science is based in a general conception that cognition only occurs in a living body (which may or may not contain a brain) and is (at least) markedly enhanced when that body is active in a pre-structured physical and social context.

The field of Artificial Life was a source of much excitement in the robotic art community. One of the most remarkable expressions of this excitement was the extraordinary *Strandbeests* of Theo Jansen.[4] These fantastical lightweight mechanical marvels are propelled in multi-legged locomotion via the force of the wind on fin-like sails. Notably, Jansen has *bred* the mechanical linkages using genetic algorithms (Fig. 3).

[4] http://www.strandbeest.com.

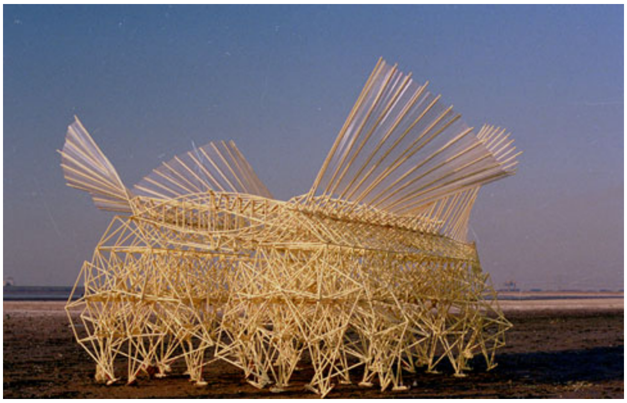

Fig. 3 Strandbeest Animarius Currens Ventosa

5 *Petit Mal,* an autobiographical interlude

My own work, *Petit Mal: an Autonomous Robotic Artwork* (built 1992–1995) arose at the confluence of embodied art practice, artificial life, and the cognitivist crisis (Fig. 4).

The primary goal of *Petit Mal* was to build a behaving machine that while entirely non-anthropomorphic and non-zoomorphic, elicits play behavior among people. Interaction is driven by curiosity and seemingly, a desire to pretend that the thing is more clever than it is. People willingly and quickly adjust their behavior and pacing to extract as much action from the device as possible, motivated entirely by pleasure and curiosity. (Interestingly, the only demographic who were unwilling to interact were adolescents). I saw the device, technically, as a demonstration of the viability of a reactive robotics strategy. The device behaves robustly, with the public (albeit in a controlled environment) for 10–12 h a day.

Fig. 4 Petit Mal—resurrected for Smile Machines, Transmediale 2006, Berlin

The context in which this work came about is significant. I had begun the project when I took up a cross-disciplinary position at Carnegie Mellon University as a professor of art and robotics. I brought to that context my experience in installation, performance, and machine sculpture, along with substantial experience in designing performative technologies and persuasive sensorial experience—and more subtly, with predictions regarding the cloud of cultural associations that might be elicited by a particular set of cues, materials, gestures, and references.

The period of development of *Petit Mal* was crucial to the development of my understanding of the hardcore engineering realities of robotics and not unrelated to the development of my critique of cognitivism. I was fortunate to have had the opportunity to move in circles with leading roboticists and to come to terms first hand with the technical realities and motivations of robotics. I began to recognize that my legacy in creating materially instantiated sensorially affective (art) work provided me with a different leverage on robotics as compared to many in the Robotics Institute, whose backgrounds were in computer science and engineering. When *Affective computing* became a buzz word in that world, my response was a forehead slapping, “Well duh!”

6 Cognitivism and robots

It is paradoxical that such an embodied and grounded pursuit as robotics should have arrived in the context of a discipline with an axiomatic commitment to the irrelevance of embodiment. The deep ontological bifurcation in the ideologies of robotics and computing extend a fault line leading back to Plato and Aristotle. We believe that either knowledge resides in abstraction, or that knowledge resides in the word. In other words, knowledge as derived or inferred from the world, or imposed upon and framing the world, based on non-material archetypes. This tension between regimes of abstraction, such as the mathematized sciences, and the realm of the senses or at least the sensorial, as exemplified by the arts, is a paradox at the core of robotics. A robot must exist and survive in the world. It must be able to manage its heterogeneous messiness and extract salience, but at its core lie the decision-making systems that inhere the legacy of platonic abstract logic. Computers, we must always remember, are in no way natural; they are implementations of a specific tradition of western mathematics and logic—the epitome of rarefied abstraction, a practice with little direct relation to the world or application for survival in the world.

Through the modern period, scientific or scientized disciplines have steadily tuned systems, epistemological and technical, for the distribution and elaboration of a worldview, which validates modes of abstract symbolic representation. Computers and computational are isomorphic with this worldview and have played a large role in its continued development. But this validation of abstract symbolic representation has serious implications. Philip Agre notes, "…the privileged status of mathematical entities in the study of cognition was already central to Descartes' theory…" and goes on to cogently observe, "a theory of cognition based on formal reason works best with objects of cognition whose attributes and relationships can be completely characterized in formal terms" (1997). Humberto Maturana puts a similar idea with typical economy: "A mathematical formalism is a conceptual and operational system that reveals the relational coherences of the space that it defines…But mathematical formalisms do not by themselves create an understanding of the phenomena that an observer helps to explain through them" (2002).

It is this commitment to abstraction that puts art and robotics together on the other side of the fence. Given its normal institutional location, robotics has had to negotiate an ongoing détente. The arts, in their commitment to unmediated (or non-semantically mediated) sensorial immediacy, are discursively snookered. A commitment to the specificities of materiality eliminates any possibility of making a persuasive case within the regime of generalized symbolic abstraction. This situation goes some way to explaining the (modern) anti-intellectualism and intellectual marginalization of art and artists.

7 Generality and specificity

One of the most unfortunate slippages in computer science discourse is around the notion of generality. While Turing's mathematical formulation of the *general-purpose machine* has undeniable force in its context, the notion has, by some sort of elision, weirdly oozed into other frames of reference. This, I would conjecture, has as much to do with marketing rhetoric sleight of hand and economies of scale in the computer industry as it does with any principled argument.

Of all the things I do in my life, only some of them map well onto sitting at a desk in front of a glowing surface, poking at buttons. Nor is this situation improved one iota when the context is miniaturized, so the buttons are smaller than my fingers and I have to put my reading glasses on to look at the screen. And if the situation is odd with respect to human users, it is decidedly weird when applied to robots. Robots are situated machines and, therefore, occupy particular and specific real-world contexts. There is no general-purpose robot. Artists and artisans understand this: there is a tool for every job and a job for every tool. To the extent that a tool is general-purpose, precisely to that extent it is horrible. A good tool makes an artisan more effective, but in the hands of a dolt, a good tool is ruined. A good tool isn't a good tool unless it is wrapped in a complementary context of skilled practice.

8 Embodiment and material instantiation

Because robots sense and act in the world, they are necessarily composed of components which connect them to their electro-physical environment, and components which translate and coordinate such activity. The pragmatic survival requirement for (power limited, time bound and resource poor) real-time computation makes robotic engineering very different to conventional computing contexts, where computational cycles and memory locations are taken to be effectively infinite and time is rubbery. In robotics, the question is not whether the task is possible, but whether it is possible before the robot falls off the cliff. AI and robotics are profoundly pragmatic pursuits, as captured in what is wittily referred to as the work ethic of, "It has to work." No amount of handwaving can obscure the fact that your robot fell down the stairs.

Material instantiation is key because mechanical components play an integral role in closed control loops characteristic of these devices. Inertia and friction lock digital systems into the temporality of real-world physics—information processing at the speed of gravity. Any robotics hobbyist knows that you can sample the state of a switch 1,000 times while it is being pressed, with unwelcome results. Thus, the *intelligence* of a robot is necessarily heterogeneous and integrated from the interactions between tightly coupled structural, mechanical, electromechanical, analog electronic, and digital aspects.[5]

This line of argument is directly related to embodied approaches to cognition, and particularly to neuroethology. In the heyday of cybernetics, Maturana, McCulloch, Pitts, and Lettvin published a groundbreaking paper, "What the frogs eye tells the frog's brain" (1959). The physical structure of a fly's eye or a bat's ear has been shown to be doing computational work. The materiality of the physical structure does real data processing long before a neuron gets a look-in (MacIver 2001). Such realities, found all over the biological world, undermine conventional

[5] I can build an analog device of a dozen components that can do a job, which would take hundreds of lines of code, without the power consumption and the complex ecology of compilers, device drivers, and operating systems. This condition already challenges assumption that the cleverness of a robot can be calculated in lines of code.

Cartesian/cognitivist assertions that *thinking* occurs in gray matter in the skull and all the rest of the body is a meat marionette with a telescope. The biomimetic and reactive turn in Robotics of the late 1980s has shown that embracing more holistic and situated models is a way out of the cognitivist box that has plagued Robotics and AI.

9 Situation, enaction, and structural coupling

In cognitive science, a second dualism, as bedeviling as the mind–body duality, is the binary of self-world. Many systems are tightly coupled to other systems, embedded in other machines, such as airplanes or automobiles. It is a question of philosophical debate whether such tight coupling constitutes situatedness or integration. For the moment, let's propose that tight coupling is constituted by permanent and actively reciprocal electronic or mechanical connection. Whether we speak in terms of Lucy Suchman's notion of Situated Cognition, or Edwin Hutchin's related Distributed Cognition, Varela and Thompson's Enactive Cognition, or for that matter current work in material anthropology and cognitive anthropology, the notion that mind ends where the brain ends is now untenable. Nor is the idea that mind stops where body stops so easily held. Intelligence can be individual, contextual, and social. The notion that we own our intelligence and our selfhoods, portably, in our skulls, might simply be a cultural construct of Victorian individualist humanism (Suchman 1987; Hutchins 1996; Varela et al. 1991).

Recently I moved into an apartment in a new city. I was profoundly aware of the miniscule details of navigation, which had to be consciously managed—the way to turn and push the front door; how much I had to lean against it to counteract its heft; the step up to the elevator; the wrinkle in the carpet; the behavior of the ancient elevator call button; the distance down the hall; the orientation of the key, and the direction it turned to let me into the apartment. The most interesting part of this experience was observing the steady process by which those procedures became unconscious, routine, and automatic. And indeed, its is just such transitions from mentally represented rules to muscular gestalts that Hubert Dreyfus took as a key argument in his phenomenologically informed refutation of the AI project (Dreyfus 1996).

Such quotidian poetry of experience is a commonly attended to in the arts. These kinds of everyday experiences demonstrate a fluid interchange between situatedness and representationalism, and the development of *intelligent action* that is not reasoned or entirely representational. How could one make a robot that could internalize such environmental learning, effectively leveraging the environment as an armature for more effective or intelligent behavior? What would it take to build a device with these sorts of learning behaviors, especially given the existing *tools of the trade*? The tools of AI have largely developed around the paradigms of reasoning and planning—the construction of internal abstract representational schemes. In this paradigm, *thinking* occurs upon these representations, and then is transferred to the world. There is good evidence, from neuroscience, ethology and psychology, that this is simply untrue, at least in many cases. As an alternative to the cognitivist paradigm, we might propose that *intelligence* exists, in an autopoietic sense, on the membrane, at the interface between organism and world, and self and other. Under such a paradigm, many of the tools of AI would have no relevance, and attempts to apply them would only result in confusion, like trying to apply a torque wrench to a cow.

A recent work that seems to embody such ideas is *Der Zermesser* (2007). By Leonhard Peschta, *Der Zermesser* is a large, but minimal geometrical robot that engages in a sensitive relation with architectural space, expressing its response as perturbations of its tetrahedral form. One might call it "architaxis." Each node is instrumented and motorized, so it can both move and adjust the lengths of is vertices. *Der Zermesser* works on a slow, non-human timescale, like a mollusk on a rock. This work is exemplary of a long-delayed reunion of esthetic systems of conventional plastic arts with interactive art. The geometric minimalism of *Der Zermesser* recalls formalist sculpture, that epitome of modernism, and indeed, it could be said to *implement* formalist esthetics as a machine, in the same way that computer hardware implements Boolean algebra. And certainly the formalisms of coding and computer engineering do have immediate sympathies with the modes and methods of formalist sculpture. The formal coherence and elegance of *Der Zermesser* is carried through to its electronic aspects—the geometrical nodes are also motor, sensory and computational nodes. The tetrahedron, the first platonic solid and the strongest polyhedron, is an icon of engineering efficiency. It is also manifested in a myriad of biological forms, evidentiary of the effectiveness of evolutionary design. *Der Zermesser* captures this doubling as an adaptive geometry (Fig. 5).[6]

10 New perspectives, new paradigms

From this perspective, the fact that our approach to building robots, at least along the high academic road, has proceeded according to a Cartesian model, is passing odd. There is a wonderful irony in the fact that the project of building intelligent embodied machines that must survive

[6] http://www.youtube.com/watch?v=2kRuxutzluE.

Fig. 5 Der Zermesser. LABoral, Gijón, Spain, 2007

in a contingent, noisy, and diverse world came to be based in a science, which asserts the irrelevance of material instantiation and situated embodiment.

I propose that the privileging of these regimes of symbolic abstraction continues to bedevil Robotics as it bedeviled AI. We find ourselves at a curious moment when computationalist theories of cognition have been hoist on their own petard as it were, and new perspectives have come forth that offer the possibility of a significant reconfiguration of notions of cognition and intelligence. While reactive approaches and Artificial Life thinking permitted some headway in the 1990s, now it is the revolution in embodied and situated cognition promising new ways forward.

How then might we start to develop a theoretical context for such a robotics? I propose that a reconsideration of mid 20th century cybernetic ideas would provide a useful basis, integrated with post-cognitivist cognitive science, emphasizing the work of the enactivist, situated, and distributed schools. From a cognitivist perspective, such ideas might seem profoundly alien, if not incomprehensible. This is the one-way mirror quality of Kuhnian paradigm shifts—cogntivism is comprehensible (if wrongheaded) from a post-cogntivist perspective, but post-cogntivism is not comprehenssible from a cogntivist perspective. With remarkable intelligence, candor, and humility, Philip Agre captured this condition in his reflections on the Klein bottle-like rhetorical closure of the culture of AI and his attempts to open it to relevant perspectives from other disciplines (Agre 1997).

11 Why cultural robotics is important

Cultural robotics, and robotics in general, is a crucible for the meeting of such profoundly different worldviews: that of the arts, committed to crafting sensorial specificity with precision and force; and that of computer science, committed to the power of generalized symbolic abstraction. To the extent that robotics is preoccupied with effective material instantiation, it is at odds with its parent discipline of computer science, and by the same token has immediate kinship with arts practices. Artworks find their validation in the socio-cultural milieu. As such, consideration of human behavior and expectations is fundamental, and not an afterthought, as implicit in the moniker *human factors*. To the extent that art practice is concerned with artifacts, which function in the socio-cultural milieu, any researcher who proposes computational devices with human interaction has something to learn from the arts.

What distinguishes robotic art from robotics proper? In my opinion, the matter hinges on instrumentality. When one builds a robot to dig a hole, the criterion for assessing the success of the device is already built into its specifications. But when one builds a performative machine for cultural purposes, the function is speculative and provocative. Because the field is technical, often one is asked about user testing and metrics for success. To me, this question has always seemed absurd. Did Shakespeare user test his sonnets? Did Goya user test his black paintings? The purpose of the act, in part, is to create a context for an unpredictable outcome. Nor is this approach exclusive to the arts. Albert Einstein said, “If we knew what it was we were doing, it would not be called research, would it?” He also asserted, “Imagination is more important than knowledge.”

From an interdisciplinary standpoint, not only is cultural robotics a highly charged test environment at the intersection of engineering and the arts, but it is also increasingly a territory of interest for critical theory—agency, embodiment, performativity are all key concepts in robotics. And to the extent that robotics aspires to create machines that in some sense behave like animals, robotics also exists at an interdisciplinary nexus between biology and engineering.

References

Agre P (1997) Towards a critical technical practice: lessons learned in trying to reform AI. In: Bowker GC et al (eds) Social science, technical systems, and cooperative work: beyond the great divide. Lawrence Erlbaum Associates, Mahwah

Braitenberg V (1986) Vehicles: experiments in synthetic psychology. MIT Press, Cambridge

Brooks R (1990) Elephants don't play chess. Robot Auton Syst 6:3–15

Brooks RA (1999) Cambrian intelligence: the early history of the new AI. MIT Press, Cambridge

Burnham J (1968) Beyond modern sculpture; the effects of science and technology on the sculpture of this century. G. Braziller, New York

Clark A (2008) Supersizing the mind: embodiment, action, and cognitive extension. Oxford University Press, Oxford

Daniels D, Schmidt BU (eds) (2008) Artists as inventors, inventors as artists. Hatje Cantz, Ostfildern

Dreyfus HL (1972) What computers can't do; a critique of artificial reason. Harper & Row, New York

Dreyfus HL (1979) What computers still can't do: a critique of artificial reason. MIT Press, Cambridge

Dreyfus H (1996) The current relevance of Merleau-Ponty's phenomenology of embodiment. Electron J Anal Philos 4(Spring):1–16. http://www.phil.indiana.edu/ejap/1996.spring/dreyfus.1996.spring.html

Hutchins E (1996) Cognition in the wild. MIT Press, Cambridge

Lettvin J, Maturana H, McCulloch W, Pitts W (1959) What the frog's eye tells the frog's brain. Proc IRE 47(11):1940–1959

MacIver MA (2001) The computational neuroethology of weakly electric fish—body modeling, motion analysis, and sensory signal estimation. University Microfilms International, Ann Arbor

Maturana HR (2002) Autopoiesis, structural coupling and cognition: a history of these and other notions in the biology of cognition. Cybern Hum Knowing 9(3/4):5–34

Maturana HR, Varela FJ (1980) Autopoiesis and cognition: the realization of the living. D. Reidel Pub. Co, Boston

Mindell DA (2002) Between human and machine: feedback, control, and computing before cybernetics. Johns Hopkins University Press, Baltimore

Newell A, Simon HA (1976) Computer science as empirical inquiry: symbols and search. Commun ACM 19(3):113–126

Penny S (2000) Agents as artworks and agent design as artistic practice. In: Dautenhahn K (ed) Human cognition and social agent technology. J. Benjamins, Philadelphia

Penny S (2008) Bridging two cultures—towards a history of the artist-inventor. In: Daniels D, Schmidt B (eds) Artists as inventors, inventors as artists. Boltzmann Institute/Pub Hatje Cantz, Ludwig, Ostfildern

Penny S (2010) Desire for virtual space: the technological imaginary in 90s media art. In: Brejzek T (ed) Space and desire: scenographic strategies in theatre, art and media. Zurich University of the Arts, ZHdK Zurich

Penny S (2011a) Trying to be calm: ubiquity, cognitivism, and embodiment. In: Ekman U (ed) Throughout. MIT Press, Cambridge (forthcoming)

Penny S (2011b) Towards a performative aesthetics of interactivity. Fibreculture. In: Ekman U (ed) Ubiquity. http://nineteen.fibreculturejournal.org/fcj-132-towards-a-performative-aesthetics-of-interactivity/

Pickering A (2010) The cybernetic brain: sketches of another future. University of Chicago Press, Chicago

Reichardt J (1968) Cybernetic serendipity: the computer and the arts: a Studio International special issue. Studio International, London

Suchman LA (1987) Plans and situated actions: the problem of human-machine communication. Cambridge University Press [Cambridgeshire], Cambridge

Varela FJ, Thompson E, Rosch E (1991) The embodied mind: cognitive science and human experience. MIT Press, Cambridge